Quelques modèles non linéaires en mécanique quantique

Mathieu Lewin

▶ **To cite this version:**

Mathieu Lewin. Quelques modèles non linéaires en mécanique quantique. Mathématiques [math]. Université Paris Dauphine - Paris IX, 2004. Français. NNT : . tel-00006306

HAL Id: tel-00006306
https://theses.hal.science/tel-00006306

Submitted on 22 Jun 2004

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.

Université Paris Dauphine
UFR Mathématiques de la Décision

Quelques modèles non linéaires en mécanique quantique

THÈSE

pour l'obtention du titre de

DOCTEUR EN SCIENCES
Spécialité Mathématiques
(arrêté du 30 mars 1992)

présentée et soutenue par

Mathieu LEWIN

JURY

Directeur de Thèse :	**Éric SÉRÉ** Professeur, Université Paris Dauphine
Rapporteurs :	**Gero FRIESECKE** Professeur, University of Warwick (UK)
	Claude LE BRIS Professeur, École Nationale des Ponts et Chaussées
	Michael LOSS Professeur, Georgia Institute of Technology, Atlanta (USA)
Examinateurs :	**Maria J. ESTEBAN** Directrice de recherche au CNRS
	Bernard HELFFER Professeur, Université Paris-Sud

21 Juin 2004

Quelques modèles non linéaires en mécanique quantique

Résumé : Cette thèse est consacrée à l'étude de trois modèles non linéaires issus de la mécanique quantique.

Dans la première partie, nous démontrons l'existence d'un minimum et d'états excités approchés pour les modèles multi-configurations décrivant la structure électronique des molécules. Ces points critiques peuvent être calculés grâce à un nouvel algorithme. Des résultats numériques sont présentés pour le premier état excité de systèmes à deux électrons.

Dans la seconde partie, nous étudions un lemme du col modélisant des réactions chimiques avec le modèle de Schrödinger indépendant du temps. Nous prouvons l'existence d'un point selle, sous l'hypothèse que les molécules à l'infini sont chargées ou polarisées.

La dernière partie est consacrée à l'étude de la polarisation du vide grâce au modèle Bogoliubov-Dirac-Fock, une théorie relativiste de champ moyen issue de l'électrodynamique quantique. Nous montrons que l'énergie étudiée possède un minimum qui est le vide polarisé.

Mots clés : méthodes variationnelles, analyse non linéaire, physique mathématique, mécanique quantique, états excités de molécules, lemme du col, électrodynamique quantique, polarisation du vide.

Some nonlinear models in quantum mechanics

Abstract : This thesis is devoted to the study of three nonlinear models from quantum mechanics.

In the first part, we prove the existence of a minimizer and of approximate excited states for the multiconfiguration methods, which aim at describing electrons in molecules. These so-defined critical points can be computed numerically by a totally new algorithm. Numerical results are provided for the first excited state of two-electron systems.

In the second part, we study a mountain pass lemma modelling adiabatic reactions in the Schrödinger time-independent framework. We prove the existence of a mountain pass point, assuming that the molecules at infinity are charged or polarized.

Our last part is devoted to the study of the polarization of the vacuum with the Bogoliubov-Dirac-Fock model, a relativistic mean-field theory deduced from quantum electrodynamics. Our energy is bounded from below and has a minimizer which can be interpreted as the polarized vacuum.

Keywords : variational methods, nonlinear analysis, mathematical physics, quantum mechanics, excited states of molecular systems, mountain pass lemma, quantum electrodynamics, polarization of the vacuum.

CEREMADE
UMR CNRS 7534
Université Paris IX - Dauphine
Place du Maréchal De Lattre de Tassigny
75 775 PARIS Cedex 16

Remerciements

Tout d'abord, je tiens à remercier très chaleureusement Éric Séré qui, durant ces trois années, a dirigé cette thèse avec beaucoup d'attention et de disponibilité. J'ai été très impressionné par sa grande culture mathématique et sa capacité à aborder les problèmes nouveaux avec patience et détermination. De nombreux résultats dans cette thèse ont été obtenus grâce à ses conseils. Enfin, la collaboration plus affirmée que nous avons entamée la dernière année a constitué une expérience scientifique extrêmement instructive.

Je remercie également mes collaborateurs. Éric Cancès a su me donner le goût aux applications en Chimie dès son cours de DEA ; notre collaboration m'a ensuite permis de mettre un pied dans un monde plus appliqué et de mieux cerner le rôle que nous, mathématiciens, pouvons jouer dans celui-ci. Hervé Galicher s'est joint à nous avec enthousiasme pour le début de sa thèse ; son implication explique en grande partie la réussite de notre travail. Enfin, merci à Christian Hainzl qui a su, lors de son post-doc au Ceremade, m'initier à des modèles physiques très complexes, avec une rare exaltation.

Je suis également très reconnaissant envers les rapporteurs et examinateurs de cette thèse, qui ont accepté de participer au jury. Tout d'abord, un grand merci à Maria Esteban qui m'a suivi et beaucoup soutenu depuis le début de ma thèse ; ma passion pour les problèmes et méthodes non linéaires provient sans aucun doute de son excellent cours de DEA que j'ai suivi à l'Université Paris Dauphine. Claude Le Bris s'est intéressé à mes travaux depuis le début et je le remercie très vivement pour ses conseils, ses questions pertinentes et son soutien. Gero Friesecke a accueilli mes résultats avec enthousiasme et nous avons eu plusieurs discussions stimulantes lors de sa visite au Ceremade. Michael Loss a accepté de rédiger un rapport sur la dernière partie ; je le remercie pour ses questions et remarques. Merci enfin à Bernard Helffer pour sa participation au jury, ses indications bibliographiques, et pour avoir accepté de prendre en charge certains de mes déplacements au sein du réseau européen *"Analysis and Quantum" HPRN-CT-2002-00277.*

Je souhaite également remercier Rafael Benguria pour son accueil très chaleureux à l'université catholique de Santiago (Chili) au mois de janvier 2004, et pour les discussions captivantes que nous avons eues durant ce séjour à propos de certaines inégalités de Lieb-Thirring. Merci également à Jean Dolbeault pour m'avoir permis de participer au projet *ECOS/CONICYT C02E06*, et pour ses multiples conseils judicieux.

Je remercie aussi Heinz Siedentop pour son accueil au LMU (Münich) en février 2004 et son intérêt pour mon travail.

Je tiens à remercier vivement Isabelle Catto pour ses encouragements et ses conseils. Je suis également très reconnaissant envers Jean-Marie Barbaroux, Éric Paturel, Stephan De Bièvre, Jan Philip Solovej, Volker Bach, Nassif Ghoussoub et Ivar Ekeland pour l'intérêt qu'ils ont porté à mon travail, et pour leurs invitations.

Pour terminer, je désire remercier mes amis et collègues, en particulier Nabile (pour ses conseils et relectures) et Benjamin (pour son aide concernant la mise en page), ainsi que Claire, Adrien, Olivia, Denis, Niousha, Florent, Maxime, Sylvain, Nora et Valérie. L'ambiance conviviale du bureau C618 me laissera un excellent souvenir.

Cette liste ne serait pas complète sans ma famille qui m'a toujours soutenu, et surtout Tabatha qui a su m'encourager durant toutes ces années.

Table des matières

Introduction et présentation des résultats

Cette thèse est consacrée à l'étude mathématique et numérique de modèles issus de la physique et de la chimie quantique.

Les problèmes mathématiques provenant de ces domaines applicatifs sont nombreux, variés, et bien souvent complexes. L'intérêt qui leur est porté est par ailleurs en plein essor depuis quelques années : les mathématiciens appliqués jouent un rôle de plus en plus décisif grâce à leurs interactions avec physiciens et chimistes. Cette thèse cherche à s'inscrire dans ce mouvement. Bien qu'abondamment utilisés dans la pratique, tous les modèles que nous étudions ici ont été très peu abordés d'un point de vue théorique. Grâce à notre analyse, nous avons tenté d'apporter une lumière nouvelle qui, nous l'espérons, permettra de clarifier et éventuellement d'améliorer les méthodes utilisées.

Les travaux présentés ici concernent la modélisation de la matière à l'échelle microscopique. Ils sont principalement théoriques, notre étude de la première partie ayant toutefois débouché sur des applications numériques intéressantes présentées au Chapitre A.II. Tous les modèles étudiés sont non linéaires.

Cette thèse est divisée en trois parties indépendantes mais fortement reliées.

La première partie est consacrée aux modèles multi-configurations (MC), qui sont utilisés pour la détermination de la structure électronique des atomes et des molécules. Nous démontrons l'existence d'un minimum de l'énergie associée, pour les modèles réels utilisés par les chimistes, améliorant les résultats antérieurs [LB94, Fri03]. Il s'agit d'un problème non linéaire, dont les équations d'Euler-Lagrange sont très complexes : c'est un système de K équations aux dérivées partielles non linéaires elliptiques couplées, posées sur l'espace tout entier $\mathbb{R}^3$. Cependant, l'originalité de notre étude concerne surtout l'existence d'états excités approchés. Alors que les études théoriques sur ce type de modèles se concentrent le plus souvent sur l'existence de minima, nous ne connaissons aucun travail permettant de relier les points critiques d'un modèle non linéaire et les états excités de la molécule (c'est-à-dire les états propres du Hamiltonien). Les points critiques que nous construisons au Chapitre A.I peuvent être calculés numériquement, et l'étude d'un algorithme totalement nouveau permettant le calcul du premier état excité de systèmes à deux électrons fait l'objet du Chapitre A.II (travail réalisé en collaboration avec Éric Cancès et Hervé Galicher, École Nationale des Ponts et Chaussées). Les résultats obtenus sont en remarquable adéquation avec les mesures expérimentales.

Dans la Partie B, nous étudions un problème d'optimisation de géométrie : les positions des noyaux de la molécule ne sont plus fixées. Nous considérons le cas d'une molécule qui possède deux configurations stables pour ses noyaux, et étudions un lemme du col entre ces deux minima, modélisant une réaction chimique adiabatique (c'est-à-dire infiniment lente). Nous commençons par décrire le comportement de ces chemins en cas de perte de

compacité. Puis, nous nous restreignons au cas particulier d'une molécule composée de deux sous-systèmes dont les géométries internes sont fixées et nous prouvons que ce comportement est impossible, démontrant ainsi l'existence du point col recherché. Ces résultats présentés au Chapitre B.I sont suivis de commentaires généraux et assez simples sur des méthodes d'optimisation de valeurs propres de matrices symétriques dans le Chapitre B.II, qui permettent de relier les travaux des Parties A et B.

Enfin, nous présentons à la Partie C l'étude, en collaboration avec Christian Hainzl et Éric Séré, d'un modèle de champ moyen en mécanique quantique relativiste, issu de l'électrodynamique quantique. Dans cette théorie, le vide réinterprété en la *mer de Dirac* joue un rôle crucial car il peut se polariser en présence d'un champ extérieur. Le modèle que nous étudions, appelé Bogoliubov-Dirac-Fock, permet de modéliser cette polarisation, qui est par ailleurs absolument nécessaire pour la construction de modèles bien fondés (c'est-à-dire d'une énergie bornée inférieurement). Le vide est représenté par un projecteur sur l'espace associé à la partie négative du spectre d'un opérateur, suivant l'interprétation de Dirac lui-même. Le vide polarisé, qui minimise l'énergie BDF, est solution d'une équation non linéaire, que nous résolvons au Chapitre C.I grâce à une méthode de type point fixe. Enfin, le Chapitre C.II est consacré à la présentation et l'interprétation de nos résultats dans un langage plus physique.

Les trois modèles que nous étudions sont non linéaires. Nous utilisons dans cette thèse des méthodes adaptées à cette difficulté, principalement initiées par P.L. Lions dans [Lio84, Lio87]. Plusieurs preuves sont basées sur des idées introduites par N. Ghoussoub dans [Gho93], et qui généralisent [Lio87]. Enfin, l'article [Fri03] de G. Friesecke sur les modèles MC nous a également inspiré, en particulier grâce à la comparaison présentée entre la méthode de concentration-compacité [Lio84] et le Théorème HVZ [Hun66, Van64, Zhi60] ou les méthodes géométriques utilisées pour les problèmes à N corps [Sim77, Sig82, HS00].

La méthode numérique utilisée au Chapitre A.II pour le calcul d'états excités de molécules est également intéressante car nous déformons des chemins en suivant ainsi le principe variationnel. Notre méthode, tout à fait générale (et inspirée de diverses idées existantes), pourrait être utilisable dans d'autres situations.

L'étude mathématique des modèles présentés nécessite quelques connaissances générales sur la modélisation de molécules en mécanique quantique. Nous n'avons pas souhaité rappeler en détails cette théorie dans cette thèse. Le lecteur pourra consulter par exemple [Lie90, Thi02, AF97, C$^+$03] et les excellents chapitres introductifs de [LB93, Can98]. Nous rappellerons dans cette introduction certains points absolument nécessaires à la compréhension de notre travail. Le lecteur aguerri pourra passer les sections 1 (consacrée à une description rapide du modèle de Schrödinger), 4.1 et 4.2 (consacrées à une description des modèles relativistes dont la QED). Nous présentons nos résultats dans les autres sections, de façon assez détaillée et, nous l'espérons, imagée.

1 La modélisation d'une molécule : le modèle linéaire de Schrödinger indépendant du temps

Dans cette section, nous présentons le modèle linéaire de Schrödinger. C'est une théorie non relativiste qui permet la description de systèmes quantiques, comme des molécules composées d'électrons et de noyaux.

Les propriétés des opérateurs de Schrödinger associés aux systèmes atomiques sont abondamment utilisées dans la seconde partie de cette thèse, consacrée à la modélisation

d'une réaction chimique. Par ailleurs, une connaissance précise du modèle linéaire est nécessaire pour décrire l'intérêt des modèles de type multi-configurations que nous étudions dans la première partie.

Dans toute la suite, nous travaillerons dans le *système des unités atomiques* pour lequel

$$m_{e^-} = e = \hbar = \frac{1}{4\pi\varepsilon_0} = 1$$

où m_{e^-} et e sont respectivement la masse et la charge de l'électron, $\hbar$ est la constante réduite de Planck, et $\frac{1}{4\pi\varepsilon_0}$ est la constante diélectrique du vide. [1]

1.1 Description du modèle

Nous voulons décrire une molécule comprenant M noyaux de charges $Z_1,...,Z_M \geq 0$, et N électrons. Pour cela, nous utilisons l'approximation de Born-Oppenheimer [BO27] dans laquelle les noyaux sont traités comme des particules ponctuelles classiques, situées en $R_1,...,R_M \in \mathbb{R}^3$. Elle se justifie par la différence de masse entre les noyaux et les électrons (un facteur de l'ordre de 10^3). Nous renvoyons par exemple à [Thi02, Section 4.6] et à [Com75, CS80, KMSW92, C$^+$03] pour une justification théorique. Nous noterons

$$R = (R_1,...,R_M) \in \Omega^M, \qquad Z = (Z_1,...,Z_M) \in (\mathbb{N})^M \quad \text{et} \quad |Z| = \sum_{m=1}^{M} Z_m,$$

avec

$$\Omega^M = \left\{(R_1,...,R_M) \in (\mathbb{R}^3)^M \mid \forall i \neq j,\, R_i \neq R_j\right\}.$$

L'état des N électrons est représenté par une fonction $\Psi(x_1,...,x_N)$ normalisée dans $L^2((\mathbb{R}^3)^N,\mathbb{C})$ (chaque x_i est dans $\mathbb{R}^3$), $|\Psi(x_1,...,x_N)|^2\, dx_1 ... dx_N$ étant interprété comme la densité de probabilité de trouver l'électron i en x_i.

A cause de l'indiscernabilité des particules, il est donc naturel de supposer que $|\Psi|^2$ soit symétrique en les variables x_i. Ensuite, il y a deux sortes de particules : les *bosons* pour lesquels Ψ est symétrique, et les *fermions* pour lesquels Ψ est antisymétrique. Comme les électrons sont des fermions, on doit supposer que Ψ est *antisymétrique* par rapport aux permutations des variables x_i :

$$\forall \sigma \in S_N,\ \Psi(x_1,...,x_N) = \varepsilon(\sigma)\Psi(x_{\sigma(1)},...,x_{\sigma(N)})\ pp.$$

Si Ψ est une fonction régulière, l'hypothèse d'antisymétrie implique que $\Psi(x_1,...,x_N) = 0$ dès que $x_i = x_j$ pour $i \neq j$. Étant donnée l'interprétation probabiliste de $|\Psi|^2$, cela correspond au fait que deux électrons ne peuvent être au même endroit. On parle de *principe d'exclusion de Pauli*.

L'espace de Hilbert qui modélise l'état de N électrons est donc le sous-espace de $L^2((\mathbb{R}^3)^N,\mathbb{C})$ constitué de ces fonctions antisymétriques, et noté dans la suite $L^2_a((\mathbb{R}^3)^N,\mathbb{C})$.

Dans cette description, nous avons négligé le spin par soucis de simplicité. Si on désire tenir compte du spin, il faut travailler dans l'espace de Hilbert $L^2_a((\mathbb{R}^3 \times \{\pm\})^N,\mathbb{C})$. Les fonctions sont alors antisymétriques par rapport aux permutations des couples de variables $(x_i,\sigma_i) \in \mathbb{R}^3 \times \{\pm\}$.

1. Dans ce système, l'unité de masse est $9{,}11 \cdot 10^{-31}$kg, l'unité de longueur est le rayon de Bohr $a_0 = 5{,}29 \cdot 10^{-11}$m, l'unité de temps est $2{,}42 \cdot 10^{-17}$s, et l'unité d'énergie est le Hartree $E = 4{,}36 \cdot 10^{-18}$J. La distance moyenne noyau-électron est de l'ordre de 1 dans l'atome d'hydrogène, et la vitesse de la lumière vaut $1/\alpha \simeq 137$ [C$^+$03].

Le Hamiltonien du système est

$$H^N(R,Z) = \sum_{i=1}^{N} \left(-\frac{1}{2}\Delta_{x_i} + V(x_i) \right) + \sum_{1 \leqslant i < j \leqslant N} \frac{1}{|x_i - x_j|} + \sum_{1 \leq l < k \leqslant M} \frac{Z_k Z_l}{|R_k - R_l|},$$

où V est le potentiel de Coulomb créé par les noyaux

$$V(x) = -\sum_{m=1}^{M} \frac{Z_m}{|x - R_m|}.$$

L'indice x_i sur Δ_{x_i} signifie que l'opérateur agit sur la $i^{\text{ème}}$ variable de $\Psi(x_1,...,x_N)$. L'opérateur $H^N(R,Z)$ est auto-adjoint, défini sur l'espace de Hilbert $L^2_a((\mathbb{R}^3)^N,\mathbb{C})$, de domaine $H^2_a((\mathbb{R}^3)^N,\mathbb{C})$ [Kat51, Kat57]. L'énergie est la forme quadratique associée

$$\mathcal{E}^N(R,\Psi) = \langle H^N(R,Z)\Psi,\Psi \rangle,$$

définie sur l'espace $H^1_a(\mathbb{R}^{3N},\mathbb{C})$. Elle contient l'énergie cinétique des électrons, l'énergie d'interaction avec les noyaux, l'énergie d'interaction entre les électrons, et finalement l'énergie d'interaction des noyaux. En mécanique classique, l'énergie cinétique d'une particule de quantité de mouvement p est $|p|^2/(2m_{e^-}) = |p|^2/2$. L'opérateur auto-adjoint associé $-\Delta/2$ s'obtient donc en effectuant le remplacement $p \leftarrow -i\nabla$.

L'évolution temporelle du système est régie par l'équation de Schrödinger

$$i\hbar \frac{\partial}{\partial t}\Psi_t = H^N(R,Z) \cdot \Psi_t.$$

Un *état stationnaire*[2] est une solution de la forme $\Psi_t = e^{-\frac{iEt}{\hbar}}\Psi$, c'est-à-dire Ψ est une fonction propre de $H^N(R,Z)$. L'étude du spectre de cet opérateur est donc primordiale pour la compréhension du système.

Nous mentionnons ici les résultats nécessaires à la compréhension des sections suivantes. Le lecteur pourra consulter le Théorème I.1 page 133 pour un énoncé des autres propriétés de $H^N(R,Z)$, et par exemple [RS78, HS00, Thi02, HS96] pour une présentation plus détaillée.

Introduisons les notations

$$E^N(R,Z) = \lambda_1^N(R,Z) = \inf_{\|\Psi\|_{L^2}=1} \langle \Psi, H^N(R,Z)\Psi \rangle,$$

et

$$E^0(R,Z) = \sum_{1 \leq l < k \leqslant M} \frac{Z_k Z_l}{|R_k - R_l|}$$

(énergie des noyaux seuls). On suppose dans toute la suite que $Z \neq (0,...,0)$. Le spectre de $H^N(R,Z)$ a l'allure générale de la Figure 1.

Il est composé d'un spectre essentiel de la forme [HS00, RS78] $\sigma_{\text{ess}}\left(H^N(R,Z)\right) = \left[\Sigma^N(R,Z);+\infty\right)$, où $\Sigma^N(R,Z)$ est un réel vérifiant $\Sigma^1(R,Z) = E^0(R,Z)$ pour $N = 1$ et $\Sigma^N(R,Z) < E^0(R,Z)$ pour $N > 1$.

2. D'après le formalisme de la mécanique quantique que le lecteur pourra trouver par exemple dans [vN32, Zei95] et [Tha92, Section 1.2], un *état physique* est un élément du quotient de la sphère de $L^2_a(\mathbb{R}^{3N},\mathbb{R})$, $SL^2_a(\mathbb{R}^{3N},\mathbb{R})/\sim$, où $\psi \sim \psi' \Leftrightarrow \exists\theta \in \mathbb{R}$, $\psi = e^{i\theta}\psi'$. L'état physique d'une solution stationnaire est donc bien constant.

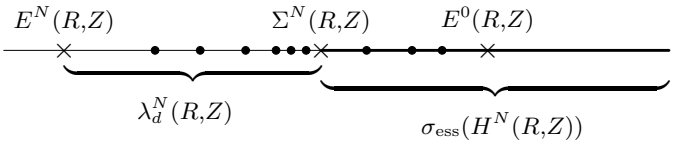


FIG. 1 – *Forme générale du spectre de* $H^N(R,Z)$

Les valeurs propres, quand elles existent, sont toutes inférieures à $E^0(R,Z)$ [RS78, HS00]. Celles qui sont strictement inférieures à $\Sigma^N(R,Z)$ sont isolées et de multiplicité finie ; elles sont en nombre fini ou dénombrable mais convergent dans ce cas vers $\Sigma^N(R,Z)$. Elles peuvent être obtenues par la formule usuelle

$$\lambda_d^N(R,Z) = \inf_{\dim(V)=d} \sup_{\substack{\Psi\in V,\\ \|\Psi\|_{L^2}=1}} \langle \Psi, H^N(R,Z)\Psi\rangle,$$

qui les fournit dans l'ordre. Lorsque qu'il n'y a qu'un nombre fini de valeurs propres sous le spectre essentiel, on a alors $\lambda_{d_0}^N(R,Z) = \Sigma^N(R,Z)$ pour un certain d_0. S'il n'y a aucune valeur propre sous $\Sigma^N(R,Z)$, alors $d_0 = 1$.

Lorsque $E^N(R,Z) = \lambda_1^N(R,Z)$ est une valeur propre strictement inférieure à $\Sigma^N(R,Z)$, on parle d'*énergie fondamentale*. Une fonction propre associée

$$H^N(R,Z)\Psi = E^N(R,Z)\Psi$$

est appelée *état fondamental*. Lorsque les $\lambda_d^N(R,Z)$, $d > 1$, sont des valeurs propres strictement inférieures à $\Sigma^N(R,Z)$, on parle d'*énergies excitées*. Les fonctions propres associées sont appelées *états excités* de la molécule. Un état fondamental ou excité est appelé *état lié*.

En vertu du principe de Hamilton, la valeur propre de plus basse énergie $E^N(R,Z)$, lorsqu'elle existe, correspond à l'état "le plus stable", dans lequel on trouvera la molécule le plus souvent (voir les remarques de [C$^+$03, Chapitre 1 - Section 6]). Toutefois, lorsque de l'énergie est fournie à la molécule, elle peut alors atteindre un *état excité*. Les différences entre deux de ces valeurs propres $\lambda_d^N(R,Z) - \lambda_{d'}^N(R,Z)$ correspondent à l'énergie apportée ou emportée par un photon lorsque le système passe d'un état excité à l'autre. Elles sont quantifiées et s'identifient au spectre de raies observé lors d'expériences de spectroscopie et que la mécanique classique ne peut décrire. [3]

Intuitivement, une molécule ne devrait pas être stable lorsque N est trop grand, les noyaux ne pouvant conserver près d'eux trop d'électrons. La traduction mathématique de

3. L'étude des valeurs propres qui sont plongées dans le spectre essentiel, dans l'intervalle $\left[\Sigma^N(R,Z); E^0(R,Z)\right]$, est bien plus délicate : celles-ci sont "moins stables" que les valeurs propres isolées, puisqu'elles peuvent disparaître subitement lorsque $H^N(R,Z)$ est légèrement perturbé. L'étude met en valeur certains phénomènes complexes de résonnance donnant naissance à des états *métastables* (c'est-à-dire ayant une vie finie, mais longue). Le lecteur intéressé pourra consulter par exemple [RS78, Section XII.6], [HS96, Chap. 16], [HS00, Section V] ainsi que [MS99, SW98, CS01] et les références citées dans ces travaux. Les valeurs propres isolées situées sous le bas du spectre essentiel $\Sigma^N(R,Z)$ ont, elles, un comportement bien plus simple puisqu'elles sont globalement stables lorsque l'opérateur est perturbé (analytiquement par exemple, voir les commentaires du Chapitre B.II).

ce fait est l'existence ou non d'un spectre discret sous $\Sigma^N(R,Z)$. Les résultats connus sont les suivants :

- *Molécules neutres et chargées positivement ($N \leq |Z|$).* Dans ce cas, Zhislin a démontré [Zhi60] l'existence d'une infinité de valeurs propres sous $\Sigma^N(R,Z)$, impliquant la stabilité des molécules neutres et chargées positivement.
- *Molécules chargées négativement ($N > |Z|$).* Dans ce cas, il existe un entier d_0 tel que $\lambda^N_{d_0}(R,Z) = \Sigma^N(R,Z)$, ce qui signifie que le spectre discret est toujours fini (voir [Zhi71, HS00] ainsi que les références citées dans [RS78, Notes of section XIII.3]).
- *Molécules comprenant un grand nombre d'électrons ($N > N_c$).* Pour chaque configuration des noyaux (R,Z), il existe un entier N_c vérifiant $|Z| \leq N_c \leq 2|Z|+M$ et tel que l'on ait $\lambda^N_1(R,Z) = \Sigma^N(R,Z)$ pour tout $N > N_c$ [Lie84, Sig82], ce qui signifie qu'il n'y a plus aucune valeur propre sous le spectre essentiel : la molécule est instable.

D'autres propriétés importantes sont mentionnées page 133, comme le Théorème HVZ [Hun66, Van64, Zhi60] qui fournit la formule $\Sigma^N(R,Z) = E^{N-1}(R,Z)$, ou la régularité et la décroissance exponentielle des fonctions propres de $H^N(R,Z)$.

Comme les fonctions d'onde à N électrons sont parfois difficiles à manipuler, on utilise souvent la *densité électronique de charge* définie par

$$\rho_\Psi(x) = N \int_{\mathbb{R}^3} \cdots \int_{\mathbb{R}^3} |\Psi(x,x_2,...,x_N)|^2 dx_2 \cdots dx_N \tag{1}$$

qui est une fonction de $L^1(\mathbb{R}^3,\mathbb{R}^+)$ d'intégrale égale à N, et *l'opérateur de densité à un corps*, défini par son noyau

$$\gamma_\Psi(x,y) = N \int_{\mathbb{R}^3} \cdots \int_{\mathbb{R}^3} \Psi(x,x_2,...,x_N)\overline{\Psi(y,x_2,...,x_N)} dx_2 \cdots dx_N, \tag{2}$$

c'est-à-dire

$$(\gamma_\Psi \varphi)(x) = \int_{\mathbb{R}^3} \gamma_\Psi(x,y)\varphi(y)\, dy$$

qui vérifie par ailleurs $0 \leq \gamma_\Psi \leq 1$, $\mathrm{tr}(\gamma_\Psi) = N$ et $\rho_\Psi(x) = \gamma_\Psi(x,x)$.

1.2 Les principaux objectifs pratiques

Le modèle linéaire de Schrödinger étant explicité, nous pouvons maintenant présenter les objectifs pratiques de cette théorie, en suivant [C$^+$03, HG96].

En pleine expansion ces dernières années, les calculs de chimie ou de physique quantique ont pour but la modélisation de la matière à *l'échelle microscopique*. Le modèle de Schrödinger possède une propriété importante : il ne contient aucun paramètre empirique, et ne dépend que de constantes physiques universelles (la masse et la charge de l'électron par exemple). On parle de modèle *ab initio*. D'autre part, la connaissance de la fonction d'onde $\Psi \in L^2_a((\mathbb{R}^3)^N,\mathbb{C})$ permet théoriquement d'accéder à toutes les propriétés physico-chimiques de la molécule : énergies, structure moléculaire, moments multipolaires, spectre Raman, et autres propriétés chimiques, mécaniques, optiques ou magnétiques.

Le principal objectif concernant le problème électronique est la connaissance de l'état fondamental, et des états excités. Si l'étude de l'état fondamental constitue la motivation première de tout modèle, le calcul des états excités est souvent aussi très important (voir des exemples au Chapitre A.II). Lorsque la position des noyaux n'est plus fixée, le modèle

de Schrödinger permet la description de réactions chimiques. La **Partie B** de cette thèse est consacrée à une étude théorique sur ce sujet.

Cependant, les avantages de la modélisation d'une molécule avec le modèle de Schrödinger s'évanouissent lorsqu'il s'agit d'effectuer des calculs réels. En effet, le modèle est posé dans l'espace $L^2_a((\mathbb{R}^3)^N,\mathbb{C})$ dont la "dimension" est très vite gigantesque : pour la molécule d'eau H_2O, on doit travailler dans $L^2_a(\mathbb{R}^{30},\mathbb{C})$! Ainsi, cette théorie ne peut en fait être utilisée que pour des systèmes très petits (6 ou 7 électrons de nos jours). Pour simuler des systèmes un peu plus grands, on utilise des modèles approchés.

Il y a deux sortes de modèles approchés, qui correspondent à deux stratégies différentes.

Dans une première approche, l'énergie est approximée et exprimée en fonction d'un objet plus simple que la fonction d'onde. C'est le cas de la *théorie de la fonctionnelle de la densité* (DFT) qui propose des fonctionnelles ne dépendant que de la densité électronique ρ_Ψ. Nous n'utiliserons pas ce type d'approximation dans cette thèse.

Une autre méthode consiste à conserver $H^N(R,Z)$, mais à réduire l'espace $L^2_a((\mathbb{R}^3)^N,\mathbb{C})$ à un sous-ensemble plus petit. C'est le cas pour les méthodes multi-configurations que nous étudions à la **Partie A**, et qui contiennent le célèbre modèle de Hartree-Fock. L'avantage de ce type d'approximation est que le minimum approché est nécessairement au dessus du minimum réel. Nous verrons que ces modèles permettent également le calcul approché des états excités, ce qui n'est probablement pas le cas de la DFT.

Ces deux approches différentes fournissent des modèles qui sont tous non linéaires. Grâce à ces techniques d'approximation du modèle linéaire, on peut maintenant simuler des sytèmes comprenant $10^3 \sim 10^4$ atomes. Enfin, dans la **Partie C**, nous introduisons un autre modèle qui permet de prendre en compte les effets relativistes.

2 Le problème électronique : méthodes de type multi-configura--tions (Partie A)

Dans cette partie, nous étudions des méthodes d'approximation pour le problème électronique, qui permettent le calcul des états fondamentaux et excités. Remarquons que comme les noyaux sont fixés, leur énergie d'interaction $E^0(R,Z)$ est constante[4]. D'autre part, nous nous restreignons à des fonctions d'onde à valeurs réelles et sans spin[5].

L'idée conductrice de ces approximations est basée sur la remarque suivante

$$L^2_a(\mathbb{R}^{3N},\mathbb{R}) \simeq \bigwedge_{i=1}^{N} L^2(\mathbb{R}^3,\mathbb{R})$$

où le membre de droite est le produit tensoriel antisymétrique. Plus précisément, si $(\varphi_i)_{i\in\mathbb{N}}$ est une base hilbertienne de $L^2(\mathbb{R}^3,\mathbb{R})$, un résultat classique permet d'affirmer que $(\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N})_{(i_1,\ldots,i_N)\in\mathbb{N}^N}$ fournit une base hilbertienne de $L^2(\mathbb{R}^{3N},\mathbb{R})$, où

$$\varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N}(x_1,\ldots,x_N) = \varphi_{i_1}(x_1)\cdots\varphi_{i_N}(x_N).$$

L'antisymétrisée de cette base

$$\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}(x_1,\ldots x_N) = \frac{1}{\sqrt{N!}} \det(\varphi_{i_k}(x_l))_{k,l}$$

4. Par soucis de clarté, nous conservons cette énergie ici, alors qu'elle sera supprimée dans la Partie A.

5. Tout notre travail peut être aisément généralisé à des fonctions d'onde à valeurs complexes et/ou avec spin.

est donc une base de $L^2_a(\mathbb{R}^{3N},\mathbb{R})$. Les fonctions $\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N}$ sont appelées *déterminants de Slater*. Une fonction d'onde quelconque $\Psi\in L^2_a(\mathbb{R}^{3N},\mathbb{R})$ s'écrit donc

$$\Psi=\sum_{I=\{i_1<i_2<\ldots<i_N\}} c_I\cdot\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N} \tag{3}$$

où la somme converge dans $L^2(\mathbb{R}^{3N},\mathbb{R})$. L'idée est maintenant de réduire l'espace en limitant le nombre de déterminants.

2.1 Le modèle Hartree-Fock et l'énergie de corrélation

Dans le modèle Hartree-Fock (HF), l'espace fonctionnel est restreint à l'ensemble des fonctions d'onde qui sont un unique déterminant de Slater [Har28, Foc30, Sla30, Löw55b]. Autrement dit, on introduit

$$\mathcal{E}^{HF}_N(\Phi)=\mathcal{E}^{HF}_N(\varphi_1,\ldots,\varphi_N)=\mathcal{E}^N(R\,,\varphi_1\wedge\cdots\wedge\varphi_N)$$

pour $\Phi=(\varphi_1,\ldots,\varphi_N)^T\in(H^1(\mathbb{R}^3))^K$ vérifiant $\int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}$, et le problème de minimisation

$$E^{HF}_N=\inf\left\{\mathcal{E}^{HF}_N(\varphi_1,\ldots,\varphi_N),\ \varphi_1,\ldots,\varphi_N\in H^1(\mathbb{R}^3),\ \int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}\right\}.$$

L'énergie $\mathcal{E}^{HF}_N$ est aisément calculée. Elle ne dépend que de l'opérateur de densité γ_Φ de la fonction d'onde $\varphi_1\wedge\cdots\wedge\varphi_N$ (voir (2)), qui est défini par son noyau

$$\gamma_\Phi(x,y)=\sum_{i=1}^N\varphi_i(x)\varphi_i(y).$$

En effet, on trouve

$$\begin{aligned}\mathcal{E}^{HF}_N(\Phi)&=\sum_{i=1}^N\int_{\mathbb{R}^3}\frac{1}{2}|\nabla\varphi_i|^2+V|\varphi_i|^2+\frac{1}{2}\iint\frac{\rho_\Phi(x)\rho_\Phi(y)}{|x-y|}dx\,dy\\&\qquad-\frac{1}{2}\iint\frac{|\gamma_\Phi(x,y)|^2}{|x-y|}dx\,dy+E^0(R,Z)\\&=\operatorname{tr}\left((-\Delta/2+V)\gamma_\Phi\right)+\frac{1}{2}\iint\frac{\rho_\Phi(x)\rho_\Phi(y)}{|x-y|}dx\,dy-\frac{1}{2}\iint\frac{|\gamma_\Phi(x,y)|^2}{|x-y|}dx\,dy+E^0(R,Z),\end{aligned}$$

ρ_Φ étant la densité de charge associée : $\rho_\Phi(x)=\gamma_\Phi(x,x)=\sum_{i=1}^N\varphi_i(x)^2$. Remarquons que γ_Φ est un projecteur de rang N. Ce résultat est général (voir plus loin) : une fonction d'onde Ψ est un déterminant de Slater si et seulement si γ_Ψ est un projecteur de rang N.

L'existence d'un minimum pour E^{HF}_N est démontrée par E.H. Lieb et B. Simon dans [LS77] et, avec une autre approche, par P.-L. Lions dans [Lio87] qui prouve également l'existence d'une infinité de points critiques. Ceux-ci, souvent interprétés dans la littérature comme des états excités de la molécule, seront discutés plus loin.

Les équations satisfaites par un point critique de l'énergie $\mathcal{E}^{HF}_N$ sont de la forme

$$\mathcal{F}_\Phi\cdot\varphi_i+\sum_{j=1}^N\lambda_{ij}\varphi_j=0, \tag{4}$$

où la matrice $\Lambda = (\lambda_{ij})$ contient les multiplicateurs d'Euler-Lagrange associés à la contrainte $\int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}$, et $\mathcal{F}_\Phi$ est l'*opérateur de Fock*

$$\mathcal{F}_\Phi = -\frac{\Delta}{2} + V + \rho_\Phi * \frac{1}{|\cdot|} - \frac{\gamma_\Phi(x,y)}{|x-y|}, \tag{5}$$

ce qui signifie

$$\mathcal{F}_\Phi \cdot \chi = \left(-\frac{\Delta}{2} + V + \rho_\Phi * \frac{1}{|\cdot|}\right)\chi - \sum_{i=1}^{N} \left((\varphi_i \chi) * \frac{1}{|\cdot|}\right)\varphi_i.$$

Comme γ_Φ est l'opérateur de densité d'ordre 1 des électrons, l'opérateur $\mathcal{F}_\Phi$ s'interprète comme l'*opérateur de champ moyen* créé par le système.

Comme $\mathcal{E}^{HF}$ est invariante[6] sous l'action du groupe $O_N(\mathbb{R})$, i.e. $\Phi \to U \cdot \Phi$ pour $U \in O_N(\mathbb{R})$, l'équation (4) peut être diagonalisée, c'est-à-dire il existe $(\varphi'_1,...,\varphi'_N)$ tels que

$$\mathcal{F}_{\Phi'} \cdot \varphi'_i = \varepsilon_i \varphi'_i. \tag{6}$$

Les φ'_i sont donc solutions d'un problème aux valeurs propres, mais dont l'opérateur dépend lui-même de Φ'. On parle d'équations auto-consistantes (en anglais on utilise le sigle SCF pour Self-Consistent Field). Notons que cette propriété d'invariance par rotation de l'énergie HF est primordiale pour obtenir des équations diagonalisées : elle est abondamment utilisée dans les preuves d'existence de [LS77, Lio87], et dans les algorithmes propres à la méthode HF (voir par exemple [CB00, C$^+$03]). De plus, on peut démontrer (on parle de principe *aufbau*) que si Φ minimise $\mathcal{E}^{HF}_N$, alors les $\varepsilon_i < 0$ sont les N plus petites valeurs propres de $\mathcal{F}_\Phi$, et que de plus $\varepsilon_N < \varepsilon_{N+1}$ [BLS94]. Dans ce cas, l'opérateur γ_Φ est le projecteur sur la somme directe des espaces propres correspondant aux N premières valeurs propres de $\mathcal{F}_\Phi$

$$\gamma_\Phi = \chi_{(-\infty;\varepsilon_N]}(\mathcal{F}_\Phi). \tag{7}$$

La différence

$$E^{corr} = E^{HF}_N - E_N > 0$$

est appelée *énergie de corrélation* [Löw59, HG96]. En effet, les déterminants de Slater sont les fonctions d'onde qui tiennent le moins compte des effets de corrélation entre les électrons : des électrons totalement indépendants devraient être représentés par une fonction d'onde de la forme $\varphi_1 \otimes \cdots \otimes \varphi_N$, mais à cause du principe de Pauli, on doit anti-symétriser cette fonction, ce qui fournit un déterminant de Slater. L'énergie de corrélation peut représenter une part non négligeable de l'énergie totale de la molécule, mais comme ce sont toujours les différences d'énergies qui importent, le modèle de Hartree-Fock fournit souvent des résultats acceptables. Cependant, au cours de certains phénomènes particuliers, cette énergie varie de manière importante, ce qui diminue fortement la précision des calculs. Des exemples typiques sont le parcours de la surface d'énergie potentielle (c'est-à-dire la résolution du problème électronique pour de nombreuses positions de noyaux), le calcul d'un état de transition lors de réactions chimiques, ou le calcul d'états excités. L'idée générale que l'on peut retenir est que la méthode HF n'est adaptée qu'à la description de systèmes *à l'équilibre* [HG96] (la même remarque est valable pour la DFT). Pour effectuer des calculs raisonnables dans d'autres situations, une description précise de la corrélation est indispensable. Les modèles multi-configurations constituent la méthode la plus adaptée pour atteindre ce but.

6. En fait $\gamma(x,y) = \Phi(x)^T \Phi(y)$ donc $\gamma_\Phi(x,y) = \gamma_{U\Phi}(x,y)$ pour tout $U \in O_N(\mathbb{R})$.

2.2 Méthodes de type multi-configurations : existence d'un minimum

2.2.1 Définition

Pour les méthodes multi-configurations (notées MC dans la suite), on limite la somme (3) à un nombre fini de déterminants, ou plus précisément on limite le nombre de fonctions (φ_i) utilisées [Löw55c, Löw59, She87, AF97, LB94, Fri03, Lew02, Lew04b]:

$$\Psi = \sum_{1\leq i_1<i_2<\ldots<i_N\leq K} c_I\, \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}. \tag{8}$$

Les variables sont donc[7]

- les *orbitales* $(\varphi_1,...,\varphi_K) \in (H^1(\mathbb{R}^3))^K$, qui vérifient la contrainte $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$;
- les coefficients $(c_{i_1,...,i_N}) \in S^{\binom{K}{N}-1}$, i.e. $\displaystyle\sum_{1\leq i_1<i_2<\ldots<i_N\leq K} (c_I)^2 = 1.$

Les contraintes permettent d'assurer que $\|\Psi\|_{L^2} = 1$. Introduisons donc la variété

$$\mathcal{M}_N^K = \left\{ (c,\Phi) \in S^{\binom{K}{N}-1} \times (H^1(\mathbb{R}^3))^K,\ \int_{\mathbb{R}^3} \Phi\Phi^T = I_K \right\}$$

où nous avons noté

$$c = (c_{i_1<\cdots<i_N}) \in \mathbb{R}^{\binom{K}{N}}, \qquad \Phi = (\varphi_1,...,\varphi_K)^T \in H^1(\mathbb{R}^3,\mathbb{R})^K$$

(nous utilisons par exemple l'ordre lexicographique pour les c_I). D'après [Löw55a], Ψ est de la forme (8) si et seulement si l'opérateur γ_Ψ est de rang fini égal à K, et dans ce cas, les φ_i engendrent $\mathrm{Im}(\gamma_\Psi)$ (c'est donc une généralisation du cas HF).

L'énergie est définie sur $\mathcal{M}_N^K$ par la formule

$$\mathcal{E}_N^K(c,\Phi) = \langle \Psi_{(c,\Phi)}, H^N(R,Z)\cdot\Psi_{(c,\Phi)}\rangle \tag{9}$$

$$\Psi_{(c,\Phi)} = \sum_{1\leq i_1<\cdots<i_N\leq K} c_{i_1\ldots i_N}\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}.$$

Comme pour HF, elle est non quadratique en les φ_i (mais elle est quadratique en c). L'énergie fondamentale approchée est alors

$$E_N^K = \inf_{\mathcal{M}_N^K} \mathcal{E}_N^K. \tag{10}$$

Il est facile de voir que l'on a $E_N^K \geq E_N$ et

$$\lim_{K\to+\infty} E_N^K = E_N,$$

ce qui montre bien que cette méthode d'approximation permet de décrire l'énergie de corrélation.

Un cas particulier de $K = N+2$ est étudié par C. Le Bris dans [LB94] où il considère des sommes de deux déterminants seulement

$$\Psi = \alpha\,\varphi_1 \wedge \cdots \wedge \varphi_{N-1} \wedge \varphi_N + \beta\,\varphi_1 \wedge \cdots \wedge \varphi_{N+1} \wedge \varphi_{N+2},$$

7. Bien remarquer que les φ_i et les c_I sont inconnus ici. Lorsque les φ_i sont fixés et que seuls les c_I sont optimisés, on parle de modèle d'Interaction de Configurations.

prouve l'existence d'un minimum pour les fonctions d'onde ayant cette forme, et démontre l'inégalité $E_N^{N+2} < E_N^N$. L'existence d'un minimum pour tout $K \geq N$ a récemment été obtenue par G. Friesecke dans [Fri03]. Sa preuve est basée sur des méthodes de localisation géométrique, un argument de type concentration-compacité de P.-L. Lions [Lio84] et l'inégalité $E_{N+1}^{K+1} < E_N^K$. Toutefois son approche, qui n'est pas du tout basée sur les équations d'Euler-Lagrange (peu étudiées jusqu'à présent à cause de leur complexité), ne permet pas *a priori* de prouver l'existence de points critiques, et donc de décrire les états excités. Dans notre travail [Lew02, Lew04b, CGL04] exposé à la Partie A, nous avons étudié ces modèles avec une méthode différente, qui repose sur les équations d'Euler-Lagrange écrites sous une forme très compacte, et permet la définition d'états excités approchés.

2.2.2 Equations d'Euler-Lagrange

Dans [Lew02, Lew04b] (voir la Partie A), nous avons écrit les équations MC sous la forme suivante

$$\begin{cases} \left(\left(-\dfrac{\Delta}{2}+V\right)\Gamma + 2W_{(c,\Phi)}\right)\cdot\Phi + \Lambda\cdot\Phi = 0 \\ H_\Phi c = \lambda\, c, \end{cases} \tag{11}$$

où Γ est une matrice $K\times K$ dépendant de c (plus précisément, c'est la matrice de γ_Ψ, écrite dans la base $(\varphi_1,...,\varphi_K)$ de $\mathrm{Im}(\gamma_\Psi)$), et $W_{(c,\Phi)}$ est une fonction matricielle $\mathbb{R}^3 \mapsto M_K(\mathbb{R})$ dépendant de c et de Φ (voir les expressions à la page 65).

La première ligne de (11) est un système de K équations aux dérivées partielles non linéaires couplées, écrit sous forme matricielle. Λ est la matrice (symétrique) des multiplicateurs associés à la contrainte $\int_{\mathbb{R}^3}\Phi\Phi^T = I_K$. La seconde équation est simplement une équation aux valeurs propres, car $\mathcal{E}_N^K$ est quadratique par rapport à c. $H_\Phi = (\langle\Phi_I, H_N\Phi_J\rangle)_{I,J}$ est la matrice de taille $\binom{K}{N}\times\binom{K}{N}$ du Hamiltonien $H^N(R,Z)$ sur l'espace $\mathrm{vect}(\Phi_I)$ engendré par les déterminants, formés à partir des φ_i.

Comme la première ligne d'équations possède une forme similaire au cas de Hartree-Fock, les chimistes utilisent le sigle MCSCF (*Multiconfiguration Self-Consistent Field*) pour désigner cette méthode, car l'opérateur (matriciel) $\left(-\frac{\Delta}{2}+V\right)\Gamma + 2W_{(c,\Phi)}$ dépend des inconnues (c,Φ). Toutefois, cette dénomination peut induire en erreur : les φ_i ne sont pas les fonctions propres d'un opérateur de champ moyen.

Pour la méthode Hartree-Fock ($K = N$), on a $\Gamma = I_N$ car γ_Ψ est un projecteur de rang N, et d'autre part le terme non linéaire $W_{(c,\Phi)}\cdot\Phi$ se simplifie de sorte que l'équation s'écrit

$$(\mathcal{F}_\Phi I_K)\cdot\Phi + \Lambda\cdot\Phi = 0.$$

On peut alors diagonaliser Λ en appliquant une rotation appropriée à Φ. Dans le cas $K > N$, on peut aussi faire agir le groupe des rotations[8] sur Φ mais à cause de la forme de l'équation (11), on pourra diagonaliser Γ **ou** Λ, mais pas les deux simultanément. Cette perte d'invariance par rotation explique la difficulté principale du modèle MC face au modèle HF. Dans nos preuves, nous diagonalisons Γ ou Λ, selon l'information désirée.

D'un point de vue numérique, alors qu'il existe des méthodes (de type point fixe) pour la résolution des équations SCF Hartree-Fock, aucune méthode spécifique n'est connue pour les méthodes MCSCF (des algorithmes de type Newton sont utilisés pour la minimisation). La forme des équations que nous avons utilisée pourrait être utile pour construire une méthode du type de celle employée pour HF .

8. L'action de $O_K(\mathbb{R})$ sur $\mathcal{M}_N^K$ est explicitée au Chapitre A.I, page 66.

2.2.3 Méthodes pratiques : existence d'un minimum

Pour les simulations numériques, on n'utilise pas le modèle MC tel qu'il vient d'être décrit car le nombre $\binom{K}{N}$ de déterminants est bien trop important. À la place, certains déterminants sont choisis (selon des considérations chimiques complexes et souvent empiriques) et l'optimisation est effectuée en fixant cette forme particulière. Comme la preuve d'existence d'un minimum de G. Friesecke [Fri03] ne semble s'adapter que au cas où tous les déterminants possibles sont pris en compte, nous avons voulu compléter ce résultat en fournissant une preuve applicable aux situations réelles.

Introduisons l'ensemble $\mathcal{A}_N^K = \{I \subset \{1,...,K\},\ |I| = N\}$. Pour un certain $\mathcal{I} \subset \mathcal{A}_N^K$, nous pouvons définir

$$\mathcal{M}_N^{K,\mathcal{I}} = \{(c,\Phi) \in \mathcal{M}_N^K,\ \forall I \notin \mathcal{I},\ c_I = 0\} \subset \mathcal{M}_N^K.$$

Dans [Lew04b], nous nous sommes limités à des méthodes que nous avons appelées *naturelles*. On dit que les φ_i sont les *orbitales naturelles* de la fonction d'onde Ψ si $\gamma_\Psi \varphi_i = n_i \varphi_i$, c'est-à-dire si les φ_i sont des fonctions propres de l'opérateur de densité à un corps[9]. Les n_i sont appelés les *nombres d'occupation*. Une méthode est dite *naturelle* lorsque pour tout $(c,\Phi) \in \mathcal{M}_N^{K,\mathcal{I}}$, il existe une rotation $U \in O_K(\mathbb{R}^3)$ telle que $(c',\Phi') = U \cdot (c,\Phi) \in \mathcal{M}_N^{K,\mathcal{I}}$ et $\Gamma' = \mathrm{diag}(n'_1,...,n'_K)$. En d'autres termes, on suppose qu'il est toujours possible d'appliquer une rotation permettant d'utiliser les orbitales naturelles, tout en restant dans l'ensemble $\mathcal{M}_N^{K,\mathcal{I}}$.

La méthode la plus utilisée en pratique, appelée CASSCF (Complete Active Space SCF) [Roo87, She87], vérifie cette hypothèse. Elle consiste à séparer les électrons en deux catégories. Les électrons de coeur (ou passifs) sont traités par un modèle Hartree-Fock, et les électrons de valence (ou actifs) sont traités par un modèle MC faisant intervenir tous les déterminants. Les fonctions d'onde utilisées sont donc de la forme

$$\Psi = \varphi_1 \wedge \cdots \wedge \varphi_{N-N_v} \wedge \left(\sum_{\substack{I=\{i_1<...<i_{N_v}\} \subset \{N-N_v+1,...,K\}, \\ |I|=N_v}} c_I\, \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_{N_v}} \right), \qquad (12)$$

N_v étant le nombre d'électrons de valence.

Notre résultat est alors le suivant (voir le Corollaire I.2 page 69)

Théorème 1 (Existence d'état fondamentaux MC). *Nous supposons $Z > N-1$. Pour tout $K \geq N$ et tout $\mathcal{I} \subset \mathcal{A}_N^K$ définissant une méthode naturelle, $\mathcal{E}_N^K$ possède un minimum sur $\mathcal{M}_N^{K,\mathcal{I}}$.*

La démonstration du Théorème 1 repose sur un résultat central que nous n'exposerons pas ici (voir le Théorème I.1 page 69), dans lequel nous montrons la compacité des suites de Palais-Smale avec information de Morse, dans l'idée de [Lio87, Gho93]. Lorsque l'on étudie un procédé de minimisation, il est possible d'obtenir une suite de Palais-Smale avec des informations supplémentaires sur la dérivée seconde grâce à la généralisation du Lemme d'Ekeland prouvée dans [BP87] (voir aussi [Gho93, FG92]).

9. Rappelons que les φ_i engendrent $\mathrm{Im}(\gamma_\Psi)$ lorsque Ψ est une somme finie de déterminants de Slater et qu'il existe toujours une rotation $U \in O_K(\mathbb{R}^3)$ de Φ qui permet de diagonaliser la matrice Γ de γ_Ψ.

2.3 Points critiques et états excités

Dans cette partie, nous présentons nos résultats concernant l'existence de points critiques pour le modèle MC, et leur interprétation en tant qu'états excités approchés de la molécule. Nous exposerons pour simplifier le cas où tous les déterminants sont pris en compte et renvoyons au Chapitre A.I pour une étude plus générale s'adaptant à des méthodes MC partielles, comme la méthode CASSCF.

2.3.1 L'abondance de points critiques

Alors que la définition d'un état fondamental approché ne pose aucun problème (il suffit de minimiser $\mathcal{E}_N^K$ sur la variété $\mathcal{M}_N^K$), il n'en est pas de même pour les états excités.

À cause de la non linéarité du modèle, les expériences montrent que les points critiques de $\mathcal{E}_N^K$ abondent sur la variété $\mathcal{M}_N^K$. Dans notre travail numérique exposé plus loin, nous avons remarqué que le lancement de l'algorithme de Newton à partir d'un point choisi au hasard induit presque toujours une convergence vers un point critique très proche. Dans cette situation où les points critiques de $\mathcal{E}_N^K$ sont en très grand nombre, il parait difficile de sélectionner ceux qui peuvent être interprétés comme des états excités approchés pour les fonctions propres du Hamiltonien $H^N(R,Z)$. Notons qu'à l'opposé, il est communément admis qu'un état excité approché doit être un état stationnaire du modèle étudié, c'est-à-dire l'un des points critiques de $\mathcal{E}_N^K$ sur $\mathcal{M}_N^K$.

Des observations similaires sont faites pour les autres modèles non linéaires (HF ou DFT). Dans [Lio87], une infinité de points critiques est exhibée pour le modèle Hartree-Fock. C'est une suite $(\varphi_1^n,...,\varphi_N^n)_{n\in\mathbb{N}}$ de solutions des équations HF qui vérifie en outre

$$\lim_{n\to+\infty} \mathcal{E}_N^{HF}(\varphi_1^n,...,\varphi_N^n) = E^0(R,Z).$$

Comme la limite vaut $E^0(R,Z)$ alors que nous savons que $\lambda_d^N(R,Z) \to_{d\to+\infty} \Sigma^N(R,Z) < E^0(R,Z)$, ces points critiques ne peuvent pas être interprétés comme des approximations des $\lambda_d^N(R,Z)$. Cela correspond par ailleurs aux constations pratiques : la méthode HF n'est pas adaptée au calcul d'états excités.

Au Chapitre A.I, Théorème I.3 (i) page 69, nous énonçons un résultat similaire à celui de [Lio87] :

Théorème 2 (Infinité de points critiques MC). *Soient $Z > N-1$ et $K \geq N$. Alors il existe une suite $(\tilde{c}^n,\tilde{\Phi}^n)$ de points critiques de $\mathcal{E}_N^K$ sur $\mathcal{M}_N^K$ qui vérifie par ailleurs*

$$\lim_{n\to+\infty} \mathcal{E}_N^K(\tilde{c}^n,\tilde{\Phi}^n) = E^0(R,Z).$$

Comme pour [Lio87], ces points ne peuvent pas être interprétés comme des états excités approchés.

2.3.2 Existence d'états excités approchés pour le modèle MC

Soient $\lambda_d^K(\Phi)$, $d = 1,..., \binom{K}{N}$, les valeurs propres de la matrice H_Φ apparaissant dans (11). De la définition de H_Φ découle immédiatement l'inégalité

$$\lambda_d(R,Z) \leq \lambda_d^K(\Phi)$$

pour tous Φ. Cette propriété a conduit les chimistes à utiliser la définition suivante pour les énergies excitées MC :

$$\mu_d^K = \inf_{\substack{\Phi\in H^1(\mathbb{R}^3)^K,\\ \int_{\mathbb{R}^3}\Phi\Phi^T = I_K}} \lambda_d^K(\Phi). \tag{13}$$

Comme on a par ailleurs

$$\lim_{K\to+\infty} \mu_d^K = \lambda_d(R,Z),$$

cette définition semble adaptée pour décrire des états excités approchés. Bien noter que pour chaque K, seulement $\binom{K}{N}$ états sont ainsi définis. Pour obtenir une approximation de la valeur propre λ_d de $H^N(R,Z)$, il est donc nécessaire d'utiliser assez de déterminants de sorte que $\binom{K}{N} \geq d$. En particulier, aucune définition d'états excités n'est fournie par les chimistes pour le modèle Hartree-Fock, puisque l'on a alors $K = N$.

Toutefois, des problèmes à la fois théoriques et pratiques importants sont la conséquence de la définition (13). Tout d'abord μ_d^K est défini comme un infimum d'une valeur propre d'une matrice symétrique dépendant d'un paramètre. Ce type de méthode ne fournit pas en général des points critiques de l'énergie $\mathcal{E}_N^K$ en cas de dégénérescence de la valeur propre à l'optimum. Nous renvoyons au Chapitre B.II dans lequel nous donnons des exemples très simples du type de comportement qui peut être obtenu dans ce cas. D'autre part, aucune méthode numérique claire n'est actuellement connue pour l'optimisation des valeurs propres et aux Chapitres A.I et A.II, nous montrons que l'algorithme actuellement utilisé pour la résolution de (13) possède des problèmes de convergence importants. Des oscillations entre des états non physiques sont parfois obtenues.

Ces difficultés de convergence sont souvent attribuées par les chimistes à un mauvais choix au départ de la forme des déterminants (c'est-à-dire des orbitales actives ou passives), comme le montrent ces extraits de modes d'emploi de logiciels de calcul (utilisant la définition (13)) :

> *Convergence of the orbital optimization procedure is normally good for CASSCF type wavefunctions, but problems can occur in calculations of excited states, especially when several states are close in energy.* (Extrait du mode d'emploi du logiciel `Molcas`)

> *Convergence of MCSCF is by no means guaranteed. Poor convergence can invariably be traced back to either a poor initial selection of orbitals, or a poorly chosen number of active orbitals. My advice is, before you even start: "Look at the orbitals ; then look at the orbitals again". Later, if you have any trouble: "Look at the orbitals some more". Even if you don't have any trouble, look at the orbitals to see if they converged to what you expected, and have reasonable occupation numbers. MCSCF is by no means the sort of "black box" that HF these days, so LOOK VERY CAREFULLY AT YOUR RESULTS.* (Extrait du mode d'emploi du logiciel `Gamess`)

Au vu de ces mises en garde traduisant les difficultés importantes rencontrées, il nous a semblé nécessaire de proposer une nouvelle définition des états excités MC, pouvant également permettre de construire une nouvelle méthode numérique. Notre résultat est le suivant (voir le Théorème I.3 (*ii*), page 69) :

Théorème 3 (Existence d'états excités MC). *Soient $Z > N-1$ et $K \geq N$. Alors il existe une suite (c^d,Φ^d) pour $d = 1,...,\binom{K}{N}$ de points critiques de $\mathcal{E}_N^K$ sur $\mathcal{M}_N^K$, d'indice de Morse au plus $d-1$, qui vérifie par ailleurs, en posant $\lambda_d^K = \mathcal{E}_N^K(c^d,\Phi^d)$,*

$$\lambda_d(R,Z) \leq \lambda_d^K \leq \mu_d^K$$

et donc

$$\lim_{K\to+\infty} \lambda_d^K = \lambda_d(R,Z).$$

Notre définition permet d'obtenir des points critiques qui peuvent être interprétés comme des états excités approchés de la molécule. Comme pour les μ_d^K utilisés par les chimistes, ceux-ci sont en nombre fini pour chaque K, et aucun état excité n'est défini pour HF.

$Sch.$	$K=+\infty$	λ_1	λ_2	$\cdots$	$\cdots$	$\cdots$	λ_d	$\cdots$	$\cdots$
MC	$\vdots$	$\uparrow$	$\uparrow$				$\uparrow$		
	K	λ_1^K	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\lambda^K_{\binom{K}{N}}$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$				
	$K=N+2$	λ_1^{N+2}	λ_2^{N+2}	$\cdots$	$\lambda^{N+2}_{\binom{N+2}{N}}$				
HF	$K=N$	λ_1^N							

FIG. 2 – *Les états excités MC et leur convergence vers le modèle linéaire de Schrödinger*

Expliquons maintenant comment ces points critiques sont construits. Pour cela, introduisons l'action suivante du groupe $\mathbb{Z}_2$ sur $\mathcal{M}_N^K$:

$$(-)\cdot(c,\Phi)=(-c,\Phi).$$

En suivant [Gho93], nous dirons qu'une fonction continue $f : S^{d-1} \to \mathcal{M}_N^K$ est $\mathbb{Z}_2$-invariante si elle vérifie $f(-x)=(-)\cdot f(x)$, c'est-à-dire

$$f(x)=(c,\Phi) \Longrightarrow f(-x)=(-c,\Phi),$$

et notons $\mathcal{C}^{d-1}$ l'ensemble de ces fonctions. Alors, λ_d^K est défini par

$$\lambda_d^K = \inf_{f\in\mathcal{C}^{d-1}} \sup_{(c,\Phi)\in f(S^{d-1})} \mathcal{E}_N^K(c,\Phi). \tag{14}$$

Autrement dit, nous effectuons un inf sup sur des ensembles qui sont des déformations de sphères de dimension fixée $d-1$, invariantes sous l'action du groupe $\mathbb{Z}_2$. Grâce à la théorie développée dans [Gho93, FG92], il est possible d'obtenir à partir de cette méthode variationnelle une suite de Palais-Smale avec une information de Morse, qui est compacte grâce au résultat démontré au Chapitre A.I page 69.

Dans nos expériences numériques, nous avons été confrontés à des cas où $\lambda_2^K=\mu_2^K$, et à d'autres où $\lambda_2^K<\mu_2^K$ (voir plus loin). Notre méthode semble donc intéressante d'un point de vue pratique. Au Chapitre B.II, nous donnons des exemples très simples (non issus de modèles chimiques) où l'inégalité stricte a lieu, même dans des cas de non dégénérescence.

2.4 Une nouvelle méthode numérique pour le calcul du premier état excité de molécules

Dans cette section, nous présentons une nouvelle méthode numérique pour le calcul du premier état excité d'une molécule, basée sur la définition (14). Ce travail, réalisé en collaboration avec Éric Cancès et Hervé Galicher du CERMICS (Ecole Nationale des Ponts et Chaussées), est présenté en détail au Chapitre A.II. Nous nous sommes limités dans un premier temps à la simulation de systèmes comprenant deux électrons. Remarquons que la définition (14) est valable pour tous les états excités (avec $d=1,...,\binom{K}{N}$), et un nombre quelconque d'électrons. Toutefois, la faisabilité de méthodes numériques pour les autres états excités ($d>2$) est moins claire car il faut déformer des surfaces de dimension $d-1$.

Commençons par fournir une autre expression de λ_2^K. Une fonction $\mathbb{Z}_2$-équivariante de $\mathcal{C}^1$ peut s'écrire sous la forme $t\in[0;2]\mapsto(c(t),\Phi(t))\in\mathcal{M}_N^K$ avec $c(1+t)=-c(t)$ et

$\Phi(1+t) = \Phi(t)$. Comme $\mathcal{E}_N^K$ est $\mathbb{Z}_2$-invariante (i.e. $\mathcal{E}_N^K(-c,\Phi) = \mathcal{E}_N^K(c,\Phi)$), nous obtenons donc aisément

$$\lambda_2^K = \inf_{(c,\Phi)\in\mathcal{M}_N^K} \left\{ \inf_{\gamma\in\Gamma_{(c,\Phi)}} \max_{t\in[0;1]} \mathcal{E}_N^K(\gamma(t)) \right\} \tag{15}$$

où

$$\Gamma_{(c,\Phi)} = \left\{ \gamma \in C^0\left([0;1],\mathcal{M}_N^K\right),\ \gamma(0) = (c,\Phi),\ \gamma(1) = (-c,\Phi) \right\}.$$

Remarquons que la méthode variationnelle de type $\inf - \max$ entre accolades dans (15) modélise un lemme du col usuel (dans l'idée de A. Ambrosetti et P.H. Rabinowitz [AR73]), entre (c,Φ) et $(-c,\Phi)$. Pour notre algorithme de calcul, nous avons simplifié le premier inf de (15) en prenant toujours $(\bar{c},\bar{\Phi})$, un minimimum global de $\mathcal{E}_N^K$ [10].

Notre nouvel algorithme pour le calcul du premier état excité peut donc être résumé ainsi (tous les détails sont donnés au Chapitre A.II) :

1. Calculer un minimum global $(\bar{c},\bar{\Phi})$ de l'énergie MC (algorithme de type Newton).
2. Chercher, grâce à un algorithme de déformation, un chemin γ sur la variété $\mathcal{M}_N^K$ dont les extrémités sont $\gamma(0) = (\bar{c},\bar{\Phi})$ et $\gamma(1) = (-\bar{c},\bar{\Phi})$, et qui minimise

$$\max \mathcal{E}_N^K(\gamma([0;1])).$$

3. Améliorer la convergence du point d'énergie maximale en utilisant un algorithme de type Newton.

En pratique, l'étape 2 au cours de laquelle les chemins sont optimisés est la plus délicate. De nombreuses méthodes numériques existent pour la résolution de ce type de problème. Les plus élaborées suivent vraiment le principe du col et déforment des chemins, modélisés par une suite de points qui sont déplacés grâce à un algorithme de type gradient. Le problème le plus important est la gestion du fait que les points ont tendance à "tomber dans les vallées". Une reparamétrisation est donc nécessaire à chaque pas de l'algorithme pour éviter ce phénomène. La méthode que nous avons utilisée est générale ; elle est décrite au Chapitre A.II.

Le lecteur remarquera que si l'étape 2 est correctement menée, l'énergie maximale sur le chemin diminue au cours des itérations. Il ne peut donc pas y avoir d'oscillations, comme celles qui sont observées avec les méthodes actuelles. Cette théorie s'adapte sans peine à des méthodes de type CASSCF. Si le choix des orbitales actives ou passives reste un peu empirique, la convergence de l'algorithme aura lieu indépendamment de ce choix.

Notre nouvelle méthode a été implémentée en Scilab (avec des modules interfacés en C), pour calculer le premier état excité de systèmes à deux électrons. Les résultats sont présentés en détail au Chapitre A.II.

Nous avons calculé le premier état excité de la molécule H_2 en fonction de la distance inter-atomique R, en utilisant la base à 10 orbitales atomiques CC-PVDZ. La Figure 3 montre l'allure du chemin optimal pour $R = 0{,}5$ Angström, et l'évolution des chemins au cours de l'algorithme. Le chemin résolvant le problème (15) possède deux "bosses". Deux points critiques d'indice de Morse égal à 1 sont donc fournis par cette méthode, mais seul celui qui possède l'énergie maximale peut être interprété comme un état excité d'après notre définition, l'autre étant une sorte d'état parasite.

La Figure 4 représente les énergies fondamentales HF et MC, ainsi que le premier état excité MC calculé par notre méthode, en fonction de la distance inter-atomique. Elle correspond remarquablement bien aux valeurs expérimentales. Notons que le modèle

10. Tout autre (c,Φ), pour lequel on sait que $\mathcal{E}_N^K(c,\Phi) < \lambda_2^K$, peut être *a priori* utilisé.

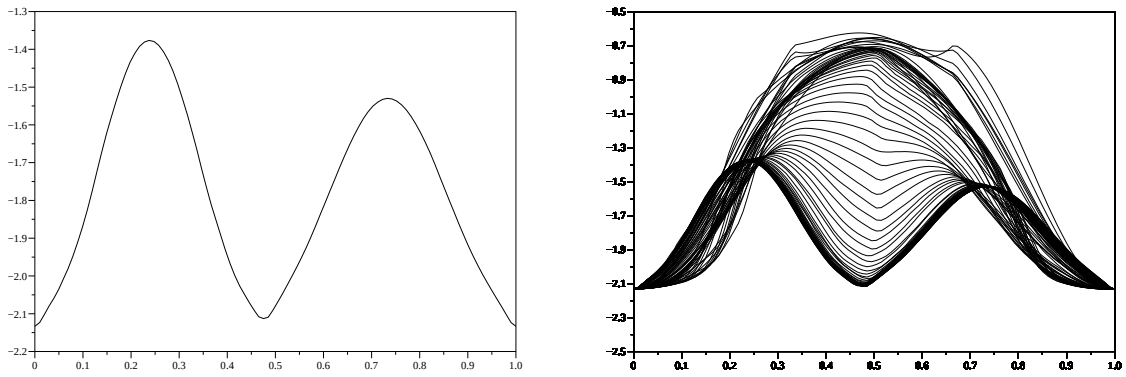


FIG. 3 – *Chemin optimal pour le calcul du premier état excité de H_2 avec $r = 0{,}5$ Angström, et évolution des chemins au cours de l'algorithme d'optimisation (à droite).*

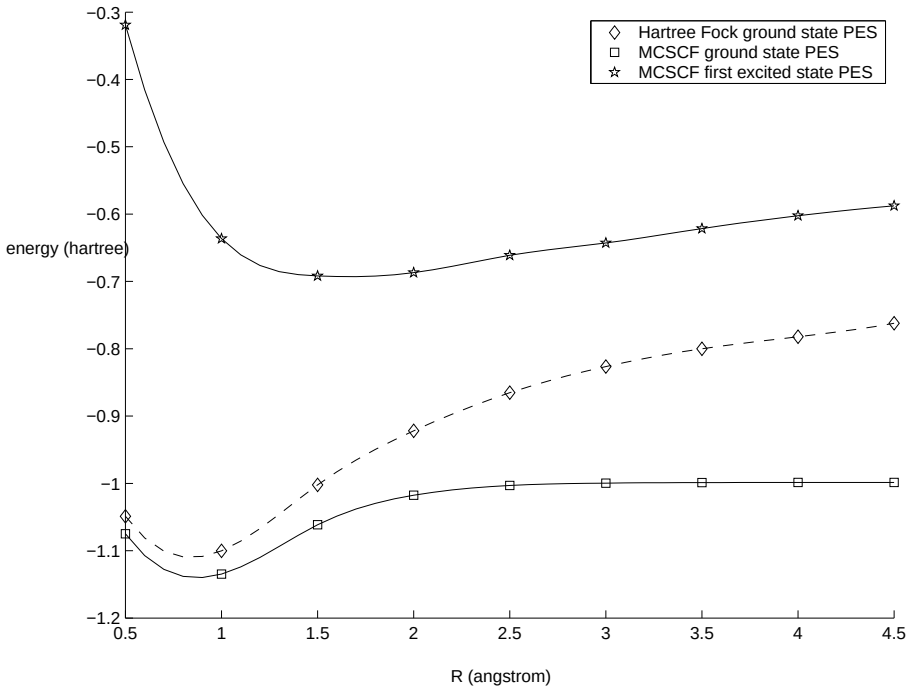


FIG. 4 – *La PES des énergies fondamentales HF et MC, et du premier état excité MC de la molécule H_2 en fonction de la distance inter-atomique.*

Hartree-Fock se comporte très mal lorsque R augmente : il n'est pas adapté à la description de l'énergie de liaison, pour laquelle la corrélation joue un rôle important. Pour R grand, les chemins optimaux finissent par posséder une seule bosse. Dans ce cas, l'algorithme se heurte parfois à une sorte de palier où la convergence est très ralentie. Pour remédier à ce problème, nous avons utilisé une technique d'allumage progressif du terme d'interaction.

Lorsque $R \in [1;2]$ (c'est-à-dire dans un voisinage du minimum de l'état excité), la solution trouvée numériquement vérifie

$$\lambda_2^K < \mu_2^K$$

(c'est-à-dire la valeur critique optimale correspond à la première valeur propre de H_Φ), améliorant donc les résultats obtenus par les méthodes classiques.

2.5 Conclusions

Dans cette première partie, nous avons étudié les modèles multi-configurations qui permettent le calcul approché du problème électronique. Nous avons démontré l'existence d'un minimum pour les méthodes utilisées en pratique, généralisant les résultats de C. Le Bris [LB94] et G. Friesecke [Fri03]. Nous avons également démontré l'existence d'une infinité de points critiques pour chaque K, dont un nombre fini seulement peut être interprété comme des états excités approchés car leurs énergies λ_d^K convergent vers les valeurs propres λ_d de $H^N(R,Z)$.

Malheureusement, nous n'avons pas pu démontrer la convergence des fonctions d'onde vers les fonctions propres de $H^N(R,Z)$, même si nous pensons qu'elle est vérifiée. Nous espérons pouvoir résoudre ce problème dans l'avenir.

Grâce à notre nouvelle définition, nous avons, dans une collaboration avec E. Cancès et H. Galicher, implémenté un nouvel algorithme permettant le calcul du premier état excité de systèmes à deux électrons. Il nous faut maintenant étendre cette étude à des systèmes plus grands, et à d'autres états excités (il faut alors déformer des surfaces de dimension ≥ 3).

De nombreuses questions intéressantes restent ouvertes concernant ces modèles, comme par exemple la vitesse de convergence vers le modèle linéaire lorsque $K \to +\infty$ (assez lente d'après les chimistes).

3 Optimisation de géométrie : la modélisation de réactions chimiques adiabatiques (Partie B)

Dans la section précédente, nous avons étudié des modèles qui permettent la résolution approchée du problème électronique. Nous voulons maintenant nous consacrer au cas où les positions des noyaux R_1,...,R_M ne sont plus fixées, et doivent être optimisées (on parle d'*optimisation de géométrie*). L'énergie d'interaction entre les noyaux $E^0(R,Z)$ n'est alors plus une constante et doit absolument être conservée. Nous revenons ici vers le modèle de Schrödinger présenté au début de cette introduction.

La première question naturelle qui apparait est celle de l'existence d'un minimum au problème d'optimisation

$$\inf_{R\in\Omega^M} E^N(R,Z), \tag{16}$$

et qui fournirait l'existence d'une configuration stable pour la molécule[11].

11. Rappelons que $\Omega^M = (\mathbb{R}^3)^M \setminus \{R_i = R_j\}$

Nous avons vu lors de la description du spectre de $H^N(R,Z)$ (page 15) que lorsque N était trop grand, il n'y avait plus d'état fondamental. On s'attend donc à un phénomène similaire lorsque cette fois $|Z|$ est trop grand : un minimum à (16) ne devrait pas exister, car la molécule a tendance à vouloir diminuer le nombre de ses noyaux. Pour contourner cette difficulté, nous allons maintenant (comme dans [LT86]) nous limiter au cas des **molécules neutres**[12] : $N = |Z|$. Dans ce cas particulier[13], il a été démontré par Lieb et Thirring dans [LT86] l'existence d'un minimum au problème d'optimisation (16).

3.1 La modélisation d'une réaction chimique adiabatique

Dans notre travail [Lew04a] présenté au Chapitre B.I, nous avons étudié le cas d'une molécule neutre qui possède **deux** positions stables R_1 et R_2 dans Ω, c'est-à-dire deux minima (éventuellement locaux) pour $R \mapsto E^N(R,Z)$. Ces deux configurations stables sont appelés *isomères* en Chimie.

La surface $\left(R,E^N(R,Z)\right)_{R\in\Omega^M}$ est appelée *Surface d'Énergie Potentielle (PES)*, ou *surface de Born-Oppenheimer*, et représente les énergies fondamentales de la molécule en fonction de la position des noyaux. Un chemin tracé sur cette surface menant de R_1 à R_2, c'est-à-dire une application continue $t \in [0;1] \mapsto R(t) \in \Omega^M$ vérifiant $R(0) = R_1$ et $R(1) = R_2$, est alors interprété comme une réaction chimique infiniment lente (ou adiabatique) au cours de laquelle la configuration R_1 est transformée en R_2. Il s'agit donc d'une *réaction d'isomérisation*. C'est le modèle le plus simple qui permet la simulation de réactions chimiques. Il ne prend pas en compte la dynamique des noyaux, et considère que les électrons sont toujours à leur état fondamental (alors qu'il arrive que ces derniers passent à un état excité au cours de réactions chimiques). C'est pourquoi ce modèle n'est adapté que pour la description de réactions très lentes.

L'hypothèse la plus classique est de supposer que le chemin emprunté par le système est celui qui nécessite le moins d'énergie, c'est-à-dire qui "monte le moins haut" sur la PES. Nous sommes donc une nouvelle fois amenés à considérer un principe du col dans l'esprit de A. Ambrosetti et P.H. Rabinowitz [AR73], et introduisons

$$c' := \inf_{R\in\mathcal{R}} \sup_{t\in[0;1]} E^N(R(t)) \tag{17}$$

$$\mathcal{R} = \left\{R \in \mathcal{C}^0([0;1],\Omega^M),\ R(0) = R_1,\ R(1) = R_2\right\}.$$

La question principale est donc maintenant celle de l'existence d'un point col d'énergie c' : nous pouvons être confrontés ici à une perte de compacité, au cours de laquelle la molécule se sépare en plusieurs sous-systèmes.

Notons tout de suite que la méthode variationnelle (17) pose d'importants problèmes car la fonction $R \mapsto E^N(R,Z)$ est une valeur propre d'un opérateur auto-adjoint : elle n'est donc pas régulière lorsque des dégénérescences ont lieu (des exemples très simples sont donnés au Chapitre B.II). La recherche d'un point critique de niveau c' n'a donc pas grand sens dans ce cadre. Au Chapitre B.I, nous proposons une méthode variationnelle pour

12. Des résultats similaires peuvent être obtenus pour des molécules chargées, mais en ajoutant des hypothèses "à l'infini".

13. Notons que ce problème est étudié pour des modèles issus de la DFT par I. Catto et P.L. Lions dans la série [CL92, CL93a, CL93b, CL93c] (dans ce cas, la force attractive est exponentiellement petite, ce qui rend l'étude d'une difficulté incomparable par rapport au cas linéaire traité dans [LT86]), et que la stabilité des molécules neutres est inconnue pour le modèle de Hartree-Fock, excepté pour des cas très particuliers [CL93c]. Elle est par contre vérifiée pour les modèles multi-configurations lorsque le paramètre K est assez grand.

remplacer (17), et qui pourrait également être intéressante d'un point de vue numérique. Nous décrivons le comportement d'une suite non compacte de chemins, puis démontrons que ce comportement est impossible dans le cas particulier d'une molécule composée de deux sous-systèmes dont les positions des noyaux sont fixes. Pour simplifier l'exposé, nous nous restreindrons dans cette introduction à ce cas particulier et renvoyons au Chapitre B.I pour une étude plus détaillée.

Pour éclaircir un peu le problème considéré, nous allons maintenant présenter un exemple très simple d'une telle réaction. Il s'agit de la molécule HCN qui existe aussi sous la forme CNH. Pour passer de HCN (minimum global) à CNH (un minimum local de l'énergie), le chemin énergétiquement le plus favorable correspondra (voir la Figure 5) :

- soit à un processus au cours duquel l'atome H "tourne" autour de la molécule CN, et passe par un point col (un point critique de l'énergie) ;
- soit à un scindement de la molécule (H part à l'infini), puis une recomposition de celle-ci dans la nouvelle configuration.

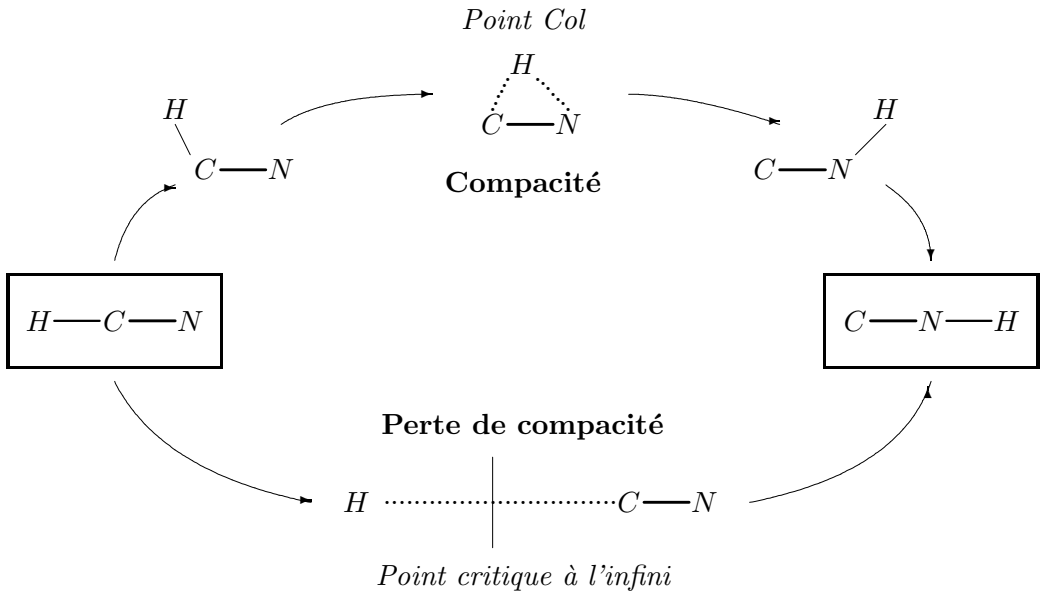


FIG. 5 – *Isomérisation de HCN.*

Sans surprise, c'est le premier cas qui est observé dans la pratique (voir la PES de l'atome H dans le champ créé par la molécule CN supposée fixe à la Figure 6). Le chemin optimal passe par un *point col*, la configuration d'énergie maximale le long de ce chemin.

Numériquement, plusieurs éléments intéressent plus particulièrement les chimistes :

- **le calcul du niveau** c', ou plus précisément la différence $c' - E^N(R_1,Z)$ qui correspond à l'énergie qu'il faut fournir à la molécule pour qu'elle puisse passer la barrière énergétique ;
- **le calcul du point col**. Cette configuration est très importante car elle fournit des informations cruciales sur le mécanisme réactionnel. Dans l'exemple précédent, ce point correspond au moment où il y a un transfert de la liaison chimique de l'hydrogène, du carbone à l'atome d'azote ;

- **le calcul du chemin optimal**, qui permet la description de tout le processus réactionnel. Nous n'aborderons pas ce problème ici[14] (voir par exemple [QH84]).

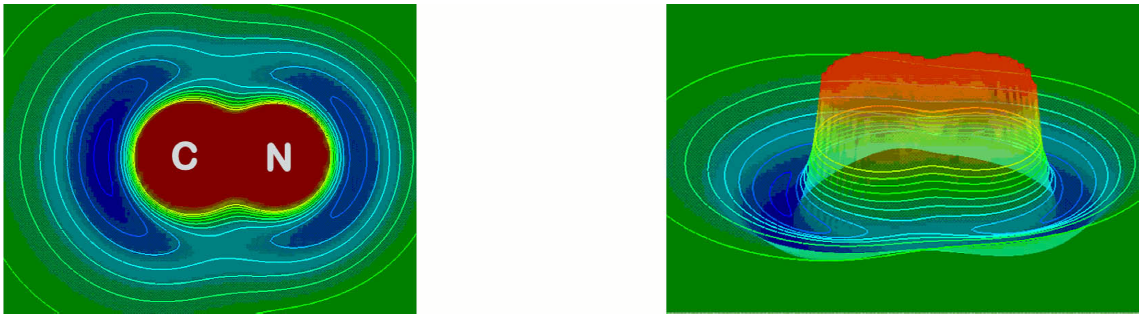


FIG. 6 – *La PES de l'atome d'hydrogène dans le champ créé par les molécules C et N, dont les noyaux sont fixés. Source :* `http://www.mcl.chem.tohoku.ac.jp/HCN/PES.html.en`

Le cas de deux molécules dont les géométries sont fixées

Comme annoncé, nous n'avons pas réussi à démontrer l'existence du point col en toute généralité et nous nous sommes restreints au cas particulier d'une molécule composée de deux sous-systèmes dont les géométries internes sont fixées, avec des hypothèses supplémentaires sur les "configurations à l'infini".

Dans [LT86], la preuve de l'existence d'un minimum à (16) est basée sur la mise en valeur d'une force attractive de l'ordre de $1/d^7$, appelée *force de Van Der Waals*[15]. Pour cela, les auteurs éliminent les termes d'ordre inférieurs grâce à une moyenne selon toutes rotations possibles des noyaux de chaque molécule. Ensuite, la force est exhibée en utilisant *la corrélation entre les électrons*. À cause de la moyenne sphérique, cela démontre l'existence d'une orientation (inconnue) de chaque molécule pour laquelle la force attractive permet la liaison des différents sous-systèmes.

Cette astuce paraît difficilement utilisable dans le cadre de l'étude du lemme du col puisque nous devons déformer des chemins : une information très précise sur les directions de descente au voisinage des "points critiques à l'infini" est nécessaire. Ces raisons justifient l'introduction d'hypothèses "à l'infini" sur les configurations des molécules.

Les paramètres sont donc

- la distance entre les deux sous-systèmes (qui sera notée d dans la suite),
- l'orientation de chaque sous-système (représentée par deux rotations $u,u' \in S\mathcal{O}_3(\mathbb{R})$),
- la fonction d'onde à N électrons.

Nous considérons donc $r = (0,r_2,...,r_m) \in \Omega^m$, $r' = (0,r'_2,...,r'_{m'}) \in \Omega^{m'}$, $z = (z_1,...,z_m)$ et $z' = (z'_1,...,z'_{m'})$. Nous notons ensuite $Z = (z,z')$, et introduisons

$$R(d,u,u') = (u \cdot r \ , \ d\vec{v} + u' \cdot r'),$$

où $\vec{v}$ est un vecteur fixé de norme 1, $d \in \mathbb{R}$, et $u,u' \in S\mathcal{O}_3(\mathbb{R})$. Nous avons utilisé la notation $u \cdot r = (u \cdot r_1,...,u \cdot r_m)$.

14. Essentiellement, le chemin optimal est composé de deux lignes de gradient menant de chaque minimum au point col.

15. Il y a différentes sortes de forces ayant ce nom. Il s'agit ici de forces à grande distance.

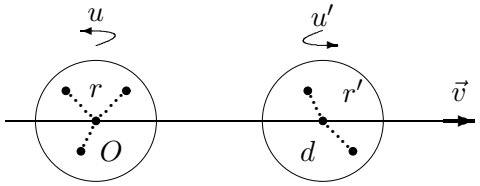


FIG. 7 – *une molécule composée de deux sous-systèmes dont les noyaux sont fixes.*

Nous supposons $N = |Z|$ et définissons maintenant

$$\mathcal{E}^N(d,u,u',\Psi) := \langle H^N(R(d,u,u'),Z)\Psi,\Psi\rangle.$$

Dans la suite, nous noterons toujours $p = (d,u,u')$ un point de $\mathbb{R}\times(S\mathcal{O}_3(\mathbb{R}))^2$ et $M = (p,\Psi)$ un état de $\mathbb{R} \times (S\mathcal{O}_3(\mathbb{R}))^2 \times SH^1_a(\mathbb{R}^{3N})$. La preuve de [LT86] fournit l'existence d'un minimum $p = (d,u,u') \in \mathbb{R} \times (S\mathcal{O}_3(\mathbb{R}))^2$ pour le problème

$$\inf_{(d,u,u')} E^N(R(d,u,u'),Z).$$

Nous supposerons que deux minima (éventuellement locaux) p_1 et p_2 existent. À une rotation près des molécules, nous pouvons prescrire $d_1 > 0$ et $d_2 > 0$. Nous introduisons alors le principe variationnel suivant, dans l'esprit de (17)

$$c' := \inf_{p\in\mathcal{P}} \max_{t\in[0;1]} E^N(R(p(t)),Z) \tag{18}$$

$$\mathcal{P} = \left\{p \in \mathcal{C}^0([0;1],]0;\infty[\times S\mathcal{O}_3(\mathbb{R})^2),\ p(0) = p_1,\ p(1) = p_2\right\}.$$

À cause des irrégularités éventuelles de $p \mapsto E^N(R(p),Z)$ en cas de dégénérescence, nous ne pouvons chercher de point critique de niveau c' pour cette fonction. Pour cette raison, nous considérons deux états fondamentaux Ψ_1 et Ψ_2 de $H^N(R(p_1),Z)$ et $H^N(R(p_2),Z)$ et introduisons alors le principe variationnel suivant

$$c = \inf_{\gamma\in\Gamma} \max_{t\in[0;1]} \mathcal{E}^N(\gamma(t))$$

où Γ est l'ensemble des fonctions continues $\gamma : [0;1] \to (0;+\infty) \times S\mathcal{O}_3(\mathbb{R})^2 \times SH^1_a(\mathbb{R}^{3N})$ telles que $\gamma(0) = M_1 := (p_1,\Psi_1)$ et $\gamma(1) = M_2 := (p_2,\Psi_2)$. En d'autres termes, nous autorisons à la fonction d'onde Ψ de varier tout le long du chemin (elle peut donc être différente d'un état fondamental). Cette fois, l'énergie $\mathcal{E}^N$ est régulière sur $(0;+\infty) \times S\mathcal{O}_3(\mathbb{R})^2 \times SH^1_a((\mathbb{R}^3)^N,\mathbb{C})$ et la recherche d'un point col de niveau c a un sens.

Il est très facile de voir que nécessairement $c' \leq c$. Notre définition serait donc d'un faible intérêt dans la pratique si $c' < c$. Heureusement, nous démontrons au Chapitre B.I (Théorème I.4, page 138) la proposition suivante[16] :

Proposition 1. *On a $c = c'$.*

16. voir également la Proposition 3 du Chapitre B.II, page 177

Notre démonstration est basée sur le fait que les fonctions d'onde sont à valeurs dans $\mathbb{C}$. Si on choisit de travailler avec des fonctions d'onde à valeurs réelles, il faut dans ce cas bien choisir les fonctions propres de départ Ψ_1 et Ψ_2 (voir la Proposition 4 du Chapitre B.II, page 178), pour que cette égalité reste vraie. Dans le Chapitre B.II, nous présentons des exemples très simples en dimension finie qui permettent de clarifier la différence entre c et c', dans les cas de $\mathbb{R}$ et $\mathbb{C}$.

Nous pensons que cette définition pourrait être intéressante d'un point de vue pratique. En effet, elle permet d'effectuer le calcul du niveau c' sans résoudre le problème électronique pour chaque position des noyaux.

3.2 La perte de compacité

Nous voulons maintenant identifier les comportements possibles en cas de défaut de compacité. Au Chapitre B.I, nous prouvons le résultat suivant (voir le Théorème I.5 page 140, et le Théorème I.4 page 138 pour le cas général)

Théorème 4. *On suppose que $N = |Z|$. Il existe deux suites $\gamma_k(t) = (p_k(t),\Psi_k(t)) \in \Gamma$ et $(\overline{p}_k,\overline{\Psi}_k) \in \gamma_k([0;1]) \subset (0;+\infty) \times S\mathcal{O}_3(\mathbb{R})^2 \times SH^1_a(\mathbb{R}^{3N},\mathbb{C})$ telles que*

$$\lim_{k\to+\infty} \max \mathcal{E}^N\left(\gamma_k([0;1])\right) = \lim_{k\to+\infty} \mathcal{E}^N(\overline{p}_k,\overline{\Psi}_k) = c$$

$$H^N(R(\overline{p}_k),Z)\cdot\overline{\Psi}_k - c\cdot\overline{\Psi}_k \to 0$$

et telles que l'on ait l'alternative suivante :

• **Premier cas :** *$(\overline{d}_k)$ est bornée et $(\overline{p}_k,\overline{\Psi}_k)$ converge vers un point critique $(\overline{p},\overline{\Psi})$ de $\mathcal{E}^N$, qui vérifie par ailleurs*

$$H^N(R(\overline{p}),Z)\cdot\overline{\Psi} = c\cdot\overline{\Psi}, \qquad c = E^N(R(\overline{p}),Z).$$

• **Second cas :** *Il existe une suite $\delta_k \to +\infty$ telle que $\delta_k \le \overline{d}_k$ et*

$$d_k(t) \le \delta_k \Longrightarrow \mathcal{E}^N(\gamma_k(t)) < c.$$

On a alors

$$c = \lim_{k\to+\infty} E^N(R(\overline{p}_k),Z) = \min\left\{E^n(r,z) + E^{n'}(r,z),\ n+n' = N\right\}. \tag{19}$$

La démonstration de ce théorème utilise abondamment la théorie développée dans [Gho93]. Un élément important est que l'on peut obtenir des points critiques d'indice de Morse au plus 1 car on déforme ici des chemins (donc de dimension 1) [Gho93, FG92]. Afin d'obtenir l'alternative du théorème, nous utilisons la notion de dualité de [Gho93] qui permet de localiser les points critiques.

Un *point col à l'infini* devrait donc correspondre à une configuration où la molécule est scindée en deux sous-systèmes infiniment éloignés les uns des autres, les électrons étant répartis entre eux et à leur état fondamental.[17] Au Chapitre B.I, nous définissons précisément la notion de point critique à l'infini dans l'esprit de Bahri [Bah89]. Cela nous permet de préciser le comportement de la fonction d'onde $\overline{\Psi}_k$ dans le second cas du Théorème 4. Pour cela, nous utilisons une décroissance exponentielle anisotrope des états

17. Ces derniers sont à leur état fondamental alors que l'on cherche *a priori* un point critique d'indice au plus 1, car les fonctions d'onde sont à valeurs complexes et $H^N(R,Z)$ est lui même réel (voir le Chapitre B.II pour plus de précisions).

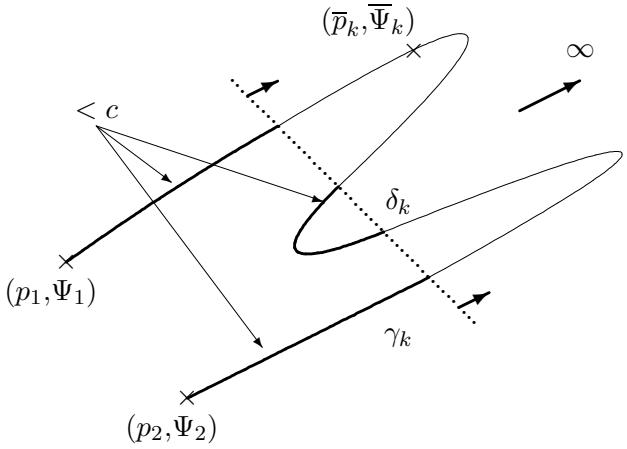


FIG. 8 – *Le second cas du Théorème 4.*

fondamentaux de $H^N(R(\overline{p}_k),Z)$ et des techniques géométriques de localisation appropriées aux problèmes à N corps inspirées de [Sim77, Sig82, HS00, Fri03]. Nous ne donnerons pas plus de détails ici et renvoyons au Chapitre B.I, Théorèmes I.4, I.2, I.3.

Remarque : il est important de noter que dans le premier cas du Théorème 4, on peut très bien avoir

$$c > \min\left\{E^n(r,z) + E^{n'}(r,z),\ n+n' = N\right\}.$$

Autrement dit, une comparaison des niveaux d'énergie est illusoire pour traiter ce type de problème, contrairement au cas de la minimisation. Il faut donc analyser le comportement des chemins dans le second cas où une perte de compacité a lieu.

3.3 Résultats d'existence d'un point col

Pour démontrer que le second cas du Théorème 4 est impossible, nous traitons deux situations

- soit (19) est atteint pour un couple (n,n') qui vérifie $(|z|-n)(|z'|-n') \neq 0$;
- soit l'unique couple minimisant (19) est $(|z|,|z'|)$.

Le cas de molécules chargées à l'infini

Dans ce premier cas, nous avons démontré le résultat suivant (voir Chapitre B.I, Théorème I.6, page 141) :

Théorème 5. *On suppose qu'il existe* $(n,n') \in \mathbb{N}^2$ *tels que* $n+n' = N$, $(|z|-n)(|z'|-n') \neq 0$ *et qui réalise le min de* (19)*. Alors le second cas du Théorème 4 n'a pas lieu, et c est donc une valeur critique de* $\mathcal{E}^N$*.*

La démonstration du Théorème 5 est assez simple (voir les détails au Chapitre B.I).

Le cas de molécules neutres mais polarisées à l'infini

Dans ce cas, nous travaillons sous l'hypothèse

(H1) $E^{|z|}(r,z)+E^{|z'|}(r',z') < E^{N_1}(r,z)+E^{N_2}(r',z')$ pour tout (N_1,N_2) vérifiant $N_1+N_2 =$

N et $(N_1 - |z|)(N_2 - |z'|) \neq 0$.

Nous allons également ajouter des hypothèses sur les configurations à l'infini. Plus précisément, nous allons supposer que chaque molécule est polarisée.

Définition 1. *Soient* $R = (R_1,...,R_M) \in \Omega^M$, $Z = (Z_1,...,Z_M)$, $\Psi \in L^2_a(\mathbb{R}^{3N},\mathbb{R})$ *et* $\rho_\Psi \in L^1(\mathbb{R}^3) \cap \mathcal{S}(\mathbb{R}^3 \setminus \{R_j\})$ *la densité de charge associée. Alors le* moment dipolaire *de la molécule est le vecteur* $P_\Psi := \int_{\mathbb{R}^3} x\rho_\Psi(x)\,dx - \sum_{j=1}^M Z_j R_j$.
Supposons que $E^N(R,Z) < \Sigma^N(R,Z)$. *Nous dirons que la molécule* (R,Z,N) *est* polarisée à son état fondamental *si* $P_\Psi \neq 0$ *pour tout état* Ψ *tel que* $H^N(R,Z)\Psi = E^N(R,Z)\Psi$, $\|\Psi\|_{L^2} = 1$.

Nous allons donc supposer

(H2) les deux molécules $(r,z,|z|)$ et $(r',z',|z'|)$ sont polarisées dans leur état fondamental ;

(H3) $E^{|z|}(r,z)$ ou $E^{|z'|}(r',z')$ est non dégénérée.

Le résultat est alors le suivant (voir Chapitre B.I, Théorème I.7, page 142)

Théorème 6. *Sous les hypothèses (H1,H2,H3), le second cas du Theorème 4 n'a pas lieu, et c est donc une valeur critique de* $\mathcal{E}^N$.

Remarquons que *(H2)* est vraie pour de nombreux cas pratiques (essentiellement, dès que les molécules ne sont pas symétriques). Quant à l'hypothèse *(H3)*, c'est une restriction purement mathématique qui est par ailleurs génériquement vraie.

La preuve du Théorème 6 est bien plus complexe que celle du Théorème 5. Elle repose en particulier sur la détermination des points critiques de l'interaction dipolaire, et de leur indice de Morse : ceux dont l'indice est au plus un ont une énergie strictement négative, qui induit une force attractive entre les molécules situées à grande distance.

3.4 Conclusions

Dans cette seconde partie, nous avons étudié la faisabilité de réactions chimiques adiabatiques. Après avoir précisé le comportement général, nous nous sommes restreints au cas d'un réarrangement de deux molécules dont les géométries internes sont fixées. Nous avons démontré l'existence d'un point col entre les deux minima, avec des hypothèses physiquement raisonnables sur les configurations à l'infini.

Malheureusement, notre méthode ne s'applique pas telle quelle à l'étude de systèmes dont les géométries internes ne sont pas fixées. De plus, même si les cas couverts sont très nombreux, nos hypothèses à l'infini restreignent tout de même le champ d'applicabilité de nos résultats. Nous aimerions donc les étendre à un cadre plus général.

D'autre part, si certaines idées sont générales, notre étude utilise pleinement la linéarité du problème en Ψ. L'étude des modèles non linéaires (DFT, HF ou MC) est clairement plus complexe.

4 La prise en compte des phénomènes relativistes : le modèle Bogoliubov-Dirac-Fock (Partie C)

Dans cette partie, nous présentons nos résultats obtenus en collaboration avec Christian Hainzl et Éric Séré [HLS04a, HLS04b], concernant la modélisation de la polarisation du vide en mécanique quantique relativiste. Pour cela, nous présentons d'abord les éléments principaux de cette théorie.

4.1 L'opérateur de Dirac

Les effets relativistes dans les molécules sont souvent négligés. Peut-être est-ce dû à la complexité des modèles, ou plutôt au fait que ces derniers présentent des difficultés théoriques et pratiques non résolues. Pourtant, ces effets ne peuvent être raisonnablement négligés dans tous les cas. Si leur influence reste faible pour les atomes légers, ils deviennent absolument nécessaires pour la modélisation d'atomes lourds. Citons l'exemple classique de l'or dont la couleur est argentée si le modèle non relativiste de Schrödinger est utilisé : sa couleur dorée est un effet purement relativiste.

Tous les maux des modèles relativistes proviennent du fait que l'opérateur de Dirac est non borné inférieurement. Cette propriété est liée à la réinterprétation de Dirac du vide, qui ne doit plus être considéré comme absent de la modélisation.

L'opérateur de Dirac libre

Nous suivrons ici principalement [Tha92, ES02, BD64, Str98, Dir33, Dir34b, Dir34a] auxquels le lecteur pourra se reporter pour une présentation plus détaillée. L'opérateur de Dirac est, comme l'opérateur de Schrödinger, obtenu à partir de l'équation d'énergie classique. Elle prend ici la forme :

$$E^2 = c^2p^2 + m^2c^4,$$

où m est la masse de l'électron, et c la vitesse de la lumière. Les remplacements $E \leftarrow i\partial_t$ et $p \leftarrow -i\nabla$ fournissent l'équation de Klein-Gordon

$$-\frac{\partial^2}{\partial t^2}\Psi(t,x) = \left(-c^2\Delta + m^2c^4\right)\Psi(t,x).$$

Toutefois, cette équation du second ordre ne décrit pas correctement l'évolution d'un système quantique qui doit être du premier ordre par rapport à t. Il faut donc chercher un opérateur auto-adjoint H vérifiant

$$H^2 = -c^2\Delta + m^2c^4 \tag{20}$$

pour l'étude de l'équation

$$i\frac{\partial}{\partial t}\Psi(t,x) = H \cdot \Psi(t,x). \tag{21}$$

La solution la plus simple serait de prendre

$$H = \sqrt{-c^2\Delta + m^2c^4}.$$

Cependant, cet opérateur est non-local, ce qui contredit le principe de causalité en mécanique relativiste, et il crée une asymétrie entre les variables spaciale x et temporelle t.

En 1928, Dirac eut l'idée de chercher H sous la forme d'un opérateur d'ordre 1

$$D_c^0 = -ic\sum_{i=1}^{3}\alpha_i\partial_{x_i} + \beta mc^2 = -ic\boldsymbol{\alpha}\cdot\nabla + \beta mc^2,$$

où nous avons noté $\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\alpha_3)$. Il est très aisé de voir que les α_i et β doivent être des matrices de dimension 4 vérifiant certaines propriétés d'anti-commutation (voir par exemple [Tha92]). Dirac a choisi

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

où

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

L'opérateur D_c^0 agit donc sur des *spineurs*

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \in L^2(\mathbb{R}^3,\mathbb{C}^4).$$

Comme nous le verrons, le fait que les fonctions d'onde soient maintenant à valeurs dans $\mathbb{C}^4$ s'interprète de manière très intéressante puisque cette représentation contient à la fois le spin et la dualité particule/anti-particule qui est caractéristique de cette théorie.

Présentons maintenant les propriétés principales de D_c^0. Tout d'abord, cet opérateur est auto-adjoint sur $L^2(\mathbb{R}^3,\mathbb{C}^4)$, de domaine $H^1(\mathbb{R}^3,\mathbb{C}^4)$ [Tha92]. Son spectre est purement continu :

$$\sigma(H_D) = (-\infty; -mc^2] \cup [mc^2; +\infty).$$

Nous noterons $P_-^0 = \chi_{(-\infty;0)}(D_c^0)$ et $P_+^0 = \chi_{(0;+\infty)}(D_c^0)$ les projecteurs sur les espaces respectivement associés au spectre négatif et positif de D_c^0.

Les valeurs propres de la matrice $c\boldsymbol{\alpha} \cdot p + \beta mc^2$ sont précisément $\pm\sqrt{c^2p^2 + m^2c^4}$. En diagonalisant cette matrice et en utilisant la transformée de Fourier inverse, on trouve

$$U_{FW} D_c^0 U_{FW}^{-1} = \begin{pmatrix} \sqrt{-c^2\Delta + m^2c^4}\, I_2 & 0 \\ 0 & -\sqrt{-c^2\Delta + m^2c^4}\, I_2 \end{pmatrix},$$

où U_{FW} est un opérateur unitaire appelé *transformation de Foldy-Wouthuysen* [Tha92, page 11]. Cela signifie, à une transformation unitaire près, P_+^0 projette sur les spineurs de la forme $(\psi_1,\psi_2,0,0)^T$ alors que P_-^0 projette sur ceux de la forme $(0,0,\psi_3,\psi_4)^T$. Les deux premières composantes de Ψ correspondent donc à une particule ayant une énergie positive, alors que les deux suivantes correspondent à une particule ayant une énergie négative, Ψ étant une combinaison de ces deux états. Chacune de ces deux particules est modélisée par une fonction d'onde à valeurs dans $\mathbb{C}^2$, donc possédant un spin à deux états. La théorie de Dirac décrit donc très naturellement le spin, alors que ce dernier n'apparait pas intrinsèquement dans la théorie non relativiste de Schrödinger. Enfin, elle met en valeur une dualité entre les états d'énergie positive et ceux d'énergie négative, que nous voulons maintenant interpréter.

La présence du spectre négatif non borné de l'opérateur D_c^0, alors que ce dernier est censé décrire l'énergie d'un électron libre, peut paraître très étrange : celui-ci devrait pouvoir "s'évanouir" dans la partie négative du spectre. C'est pourquoi Dirac a proposé l'interprétation suivante [Dir33, Dir34b, Dir34a] : si l'électron ne peut atteindre les états d'énergies négatives, c'est que ces derniers sont *déjà occupés* par des particules virtuelles. Cela implique une nouvelle conception du vide : celui-ci est en fait composé d'une infinité de particules virtuelles occupant les énergies négatives de l'opérateur de Dirac libre D_c^0, appelée *mer de Dirac*. Ainsi, le vide peut être modélisé par le projecteur P_-^0 sur le sous-espace associé à la partie négative du spectre.

Ce nouveau concept possède des conséquences très surprenantes. En effet, en fournissant suffisamment d'énergie à ces particules virtuelles, on a découvert en 1933 que l'on pouvait les faire "apparaitre". Cependant, lorsqu'un électron est créé à partir de la mer

de Dirac, un "trou" apparait dans le vide. Cette lacune peut être interprétée comme une charge positive : c'est le positron (ou anti-électron).[18] Ce dernier n'est pas moins stable dans le vide que l'électron. Cependant, lorsque les deux particules duales se rencontrent, elles s'annihilent et toute leur énergie est rayonnée. Ainsi, dans notre univers chargé d'électrons, la vie d'un positron sera en fait assez courte. La découverte des anti-particules (cinq ans après l'invention de l'opérateur D_c^0 par Dirac) constitue le succès le plus étonnant de cette théorie.

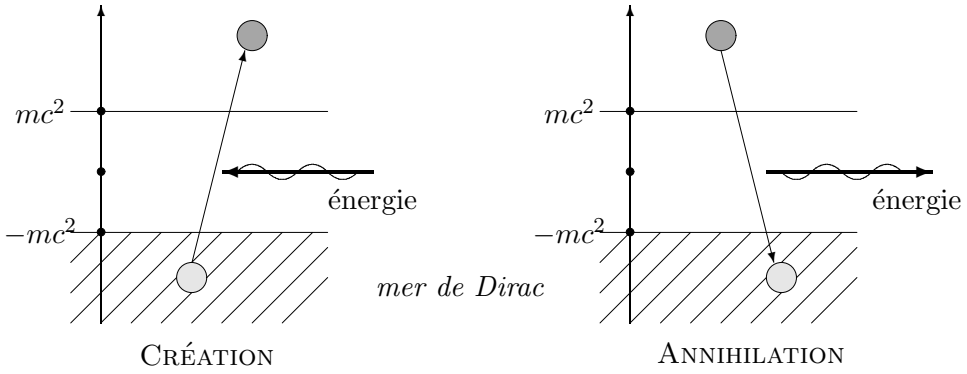


FIG. 9 – *Création et annihilation de paires électron-positron*

L'opérateur de Dirac avec potentiel extérieur

Nous avons vu que l'on devait identifier le vide avec le projecteur P_-^0. Cette vision est confirmée si l'on ajoute maintenant un potentiel extérieur, créé par exemple par un noyau de charge Z situé à l'origine. On doit alors étudier l'opérateur

$$D_c^{eV} = D_c^0 - eV = -ic\boldsymbol{\alpha}\cdot\nabla + \beta mc^2 - eV$$

où e est la charge de l'électron et V est le potentiel électrostatique créé par le noyau. Sous l'approximation de Born-Oppenheimer, on prend donc

$$V(x) = -\frac{Ze}{|x|}, \qquad Z > 0.$$

Pour la suite, nous introduisons la *constante de structure fine* $\alpha = e^2$.

L'approximation de Born-Oppenheimer pose des problèmes importants pour l'opérateur de Dirac, dus au fait que D_c^0 est un opérateur d'ordre 1. En effet, on peut donner un sens à D_c^{eV} comme opérateur auto-adjoint sur $L^2(\mathbb{R}^3,\mathbb{C}^4)$, uniquement lorsque $0 \leq \alpha Z < c$ (voir les détails dans [Tha92, page 114])[19]. Son spectre a la forme

$$\sigma(D_c^{eV}) = (-\infty; -mc^2] \cup \{\lambda_1^c \leq \cdots \lambda_d^c \leq \cdots\} \cup [mc^2; +\infty)$$

18. Ces phénomènes ont spontanément lieu dans la nature, par exemple lors de l'interaction des rayons cosmiques de haute énergie avec la matière interstellaire ou intergalactique. D'autre part, on sait maintenant créer toutes les anti-particules en laboratoire ; les positrons sont même abondamment utilisés en imagerie médicale.

19. D_c^{eV} est auto-adjoint de domaine égal à $H^1(\mathbb{R}^3,\mathbb{C}^4)$ lorsque $0 \leq e^2 Z < c\sqrt{3}/2$ et admet une unique extension auto-adjointe déterminée par les conditions de [Tha92] pour $c\sqrt{3}/2 \leq e^2 Z < c$.

où les $\lambda_d^c \in]0; mc^2[$ sont des valeurs propres de multiplicité finie, convergeant vers mc^2 lorsque $d \to +\infty$. Si $\alpha Z \to c$, la première valeur propre tend vers 0 et l'opérateur cesse d'être défini [Tha92]. Ce phénomène disparait si on considère un noyau "élargi" en prenant cette fois

$$V(x) = eZn * \frac{1}{|\cdot|}(x),$$

où n est une fonction positive d'intégrale égale à 1. Dans ce cas, l'opérateur D_c^{eV} est toujours bien défini.

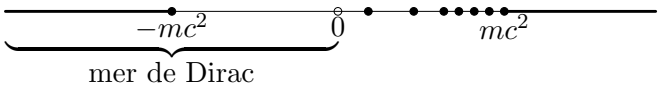


FIG. 10 – *Le spectre de l'opérateur de Dirac avec potentiel Coulombien*

Les valeurs propres de $D_c^{\alpha Z/|x|} - mc^2$ convergent vers celles de l'opérateur de Schrödinger $-\Delta/2 - \frac{\alpha Z}{|x|}$ lorsque $c \to +\infty$.[20] Cette propriété permet d'interpréter la première valeur propre comme un état fondamental et confirme l'interprétation fournie plus haut : la partie négative du spectre doit être négligée car elle correspond au vide. D'ailleurs, elle "disparait" en $-\infty$ lorsque $c \to +\infty$. Dans ce cas, il semble naturel de représenter le vide par le projecteur $P_-^{\alpha Z/|x|}$, sur le sous-espace correspondant à la partie négative du spectre de $D_c^{\alpha Z/|x|}$. Ce choix est appelé *la représentation de Furry*.

Nous pouvons remarquer que notre choix de représentation du vide est différent lorsque l'on étudie D_c^0 ou D_c^{eV}. Cela signifie que les particules virtuelles de la mer de Dirac *réagissent à un champ extérieur*, et sont modifiées pour occuper les énergies négatives du spectre d'un autre opérateur. Si la mer de Dirac associée à D_c^0 est par postulat inaccessible à l'expérience, on devrait pouvoir détecter les modifications de celle-ci en présence du champ extérieur V.

En fait, les expériences montrent que dans ce cas le vide *se polarise* (c'est-à-dire il se charge localement, mais conserve une charge globale nulle, au moins lorsque le champ n'est pas trop fort) : voir la Figure 11. Cette polarisation, très faible dans la plupart des situations pratiques, joue pourtant un rôle considérable pour la construction de modèles bien fondés, comme nous allons le voir dans un instant.

La polarisation du vide crée donc un nouveau champ, qui devrait être pris en compte dans l'opérateur $D_c^{\alpha Z/|x|}$, et le vide devrait donc être plutôt représenté par un projecteur de la forme

$$P = \chi_{(-\infty;0]}\left(D_c^{e^2 Z/|x|} + V_{\text{pol}}\right), \tag{22}$$

où V_{pol} est le potentiel créé par la polarisation du vide. Comme évidemment ce potentiel dépend lui-même du projecteur choisi, $V_{\text{pol}} = V_{\text{pol}}(P)$, l'équation (22) est une équation auto-consistante, dont l'étude fait l'objet du Chapitre C.I, et que nous préciserons plus tard. Notons que la polarisation du vide est à l'origine d'une modification du spectre de

20. Plus précisément, plusieurs λ_d^c convergent vers la même valeur propre : on parle de structure fine des raies.

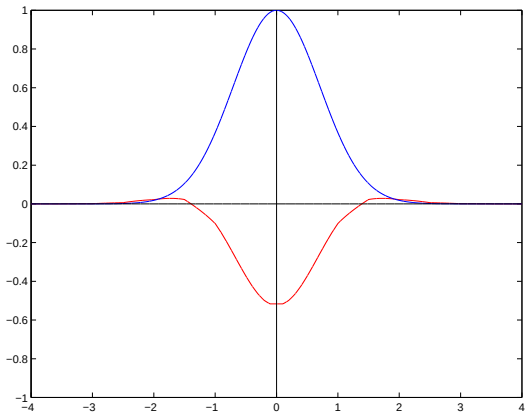


FIG. 11 – *Polarisation du vide (courbe du bas) créée par une distribution positive de charge (courbe du haut). Échelle non respectée.*

raies de l'hydrogène, appelée *structure hyperfine* [21]. Les niveaux d'énergies sont légèrement décalés (on parle de *Lamb shift*) par rapport aux λ_d^c. Au Chapitre C.II, nous montrons que notre modèle permet de retrouver ce décalage étudié dans les ouvrages de physique.

Le modèle que nous avons étudié est appelé Bogoliubov-Dirac-Fock (BDF) et il a été introduit par Chaix et Iracane dans [CI89]. Il permet la description de la polarisation du vide car il est issu de l'électrodynamique quantique (QED). Une énergie est définie et le projecteur représentant le vide est obtenu comme un minimiseur de cette énergie. Il est solution d'une équation d'Euler-Lagrange qui possède la forme (22). Mais ce modèle permet aussi la description de molécules, c'est-à-dire de systèmes comprenant des électrons. Comme nous le verrons, il devrait permettre de justifier le modèle de Dirac-Fock [ES99, ES03] qui est l'analogue relativiste de HF.

Pour expliquer comment le modèle BDF est déduit de la QED, il nous faut présenter rapidement le formalisme de cette théorie dans la section suivante. Le lecteur habitué à l'utilisation de l'espace de Fock et des opérateurs de création et d'annihilation pourra lire directement la section 4.3.

4.2 Introduction au formalisme de l'électrodynamique quantique

Dans toute la suite, nous travaillerons dans des unités telles que

$$m = c = 1.$$

Nous introduisons alors les opérateurs de Dirac dans ces nouvelles unités

$$D^0 = -i\boldsymbol{\alpha}\cdot\nabla + \beta, \qquad D^{\alpha\varphi} = -i\boldsymbol{\alpha}\cdot\nabla + \beta - \alpha\varphi.$$

Nous noterons $P^0 = \chi_{(-\infty;0]}(D^0)$ le projecteur associé à la partie négative du spectre de D^0, et $P^{\alpha\varphi} = \chi_{(-\infty;0]}(D^{\alpha\varphi})$ lorsque $0 \notin \sigma(D^{\alpha\varphi})$.

21. Plusieurs effets contribuent à la structure hyperfine du spectre des atomes et la polarisation du vide n'est que l'un d'entre eux. Elle ne représente en fait qu'environ 3% du Lamb Shift des états $2s_{1/2}$–$2p_{1/2}$ de l'hydrogène (le reste est principalement dû à l'interaction avec les photons), mais environ 90% du Lamb Shift d'un atome d'hélium muonique.

La difficile modélisation de systèmes relativistes à N particules

Nous avons vu comment interpréter le spectre de l'opérateur de Dirac décrivant un électron libre ou dans le champ créé par un noyau. L'interprétation était facilitée par le fait que le spectre comprend deux parties non bornées distinctes, séparées par un "trou spectral" dans lequel sont les valeurs propres. Cette propriété disparait complètement lorsque l'on étudie l'opérateur de Dirac pour N particules, avec $N \geq 2$,

$$\sum_{i=1}^{N} (D^{\alpha\varphi})_{x_i} + \sum_{1\leq i<j\leq N} \frac{\alpha}{|x_i - x_j|}, \tag{23}$$

car cet opérateur possède un spectre égal à toute la droite réelle. Aucune théorie satisfaisante n'est actuellement connue pour la description de systèmes à N corps avec cet opérateur. Certains auteurs ont étudié des modèles où l'opérateur $D^{\alpha\varphi}$ est remplacé par un opérateur projeté comme $P_+^{\alpha\varphi} D^{\alpha\varphi} P_+^{\alpha\varphi}$ ou $P_+^0 D^{\alpha\varphi} P_+^0$, voir par exemple [Suc80, Suc87, LSS97] et les références dans [ES02, ES99], mais qui ne semblent pas satisfaisants dans la pratique.

Le modèle de Dirac-Fock [ES99, ES03, ES01, Pat00] est lui déduit de cet opérateur tout comme Hartree-Fock est déduit de l'opérateur de Schrödinger (au moins formellement), en le restreignant aux déterminants de Slater. On obtient une énergie ne dépendant que de l'opérateur de densité γ_Φ, comme pour HF

$$\mathcal{E}_N^{DF}(\Phi) = \mathrm{tr}\,(D^{\alpha\varphi}\gamma_\Phi) + \frac{\alpha}{2}\iint \frac{\rho_\Phi(x)\rho_\Phi(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\gamma_\Phi(x,y)|^2}{|x-y|}dx\,dy. \tag{24}$$

Cette énergie est non bornée inférieurement, ce qui constitue la difficulté principale de ce modèle. L'existence de points critiques a été démontrée par M.J. Esteban et E. Séré dans [ES99], les hypothèses ayant ensuite été améliorées par E. Paturel dans [Pat00]. Ils sont solutions d'une équation SCF de la forme

$$\mathcal{F}'_\Phi \varphi_i = \varepsilon_i \varphi_i, \qquad \varepsilon_i \in (0;1) \tag{25}$$

où $\mathcal{F}'_\Phi$ est l'opérateur de champ moyen (de Fock) similaire à (5).

$$\mathcal{F}'_\Phi = D^{\alpha\varphi} + \alpha\rho_\Phi * \frac{1}{|\cdot|} - \alpha\frac{\gamma_\Phi(x,y)}{|x-y|}. \tag{26}$$

Même si cette énergie n'est pas minorée, les résultats numériques obtenus sont très satisfaisants[22]. Un objectif important des travaux de ces dernières années [CI89, CIL89, Cha90, ES03, BBHS99, BFHS04] a été de tenter d'expliquer ces succès par le fait que le modèle DF puisse être déduit de la QED, dont nous allons présenter les caractéristiques principales.

Électrodynamique quantique

La QED est une méthode très générale permettant l'étude de systèmes relativistes car

- elle réinterprète les particules d'énergies négatives en des positrons (qui sont des particules d'énergie et de charge positives)

22. Une grande prudence est toutefois nécessaire au cours des simulations numériques [Cha90].

- elle est posée sur *l'espace de Fock* qui contient les états ayant un nombre arbitraire de particules [23]
- elle tient compte de la quantification du champ électromagnétique grâce aux photons (que nous négligerons ici).

Nous survolons ici cette théorie pour expliquer comment le modèle Bogoliubov-Dirac-Fock en est déduit. Le lecteur pourra consulter l'Appendice du Chapitre C.I page 225 et par exemple [Tha92, BBHS99, HS98, CI89] pour de plus amples détails.

Commençons par définir *l'opérateur de conjugaison de charge* C qui permet de transformer un électron (virtuel) d'énergie négative en un positron

$$C \cdot \psi = i\beta\alpha_2\overline{\psi},$$

et qui vérifie $CD^{\alpha\varphi}C^{-1} = -D^{-\alpha\varphi}$ pour tout potentiel φ. Posons maintenant $\mathcal{H} = H^1(\mathbb{R}^3,\mathbb{C}^4)$ et

$$\mathcal{H}_0^+ = (1-P^0)\mathcal{H}, \qquad \mathcal{H}_0^- = P^0\mathcal{H}$$

les sous-espaces sur lesquels D^0 est de signe constant, qui sont tels que $\mathcal{H} = \mathcal{H}_0^+ \oplus \mathcal{H}_0^-$. Nous pouvons maintenant construire l'espace de Fock, à partir de ces espaces dépendant de l'opérateur libre. Pour cela, nous introduisons $\mathcal{F}_+^{(0)} = \mathcal{F}_-^{(0)} = \mathbb{C}$ et

$$\mathcal{F}_+^{(n)} = \bigwedge_{k=1}^{n} \mathcal{H}_0^+, \qquad \mathcal{F}_-^{(m)} = \bigwedge_{k=1}^{m} C\mathcal{H}_0^- \quad \text{ et } \quad \mathcal{F}^{(n,m)} = \mathcal{F}_+^{(n)} \otimes \mathcal{F}_-^{(m)},$$

décrivant les états à n électrons et m positrons. L'espace de Fock contient tous ces états possibles

$$\mathcal{F} := \bigoplus_{n,m=0}^{\infty} \mathcal{F}^{(n,m)}; \tag{27}$$

c'est un espace de Hilbert lorsqu'il est muni du produit scalaire

$$\langle \Psi,\Psi' \rangle = \overline{\Psi_{0,0}}\Psi'_{0,0} + \sum_{n,m\neq(0,0)} \langle \Psi_{n,m},\Psi'_{n,m} \rangle_{L^2(\mathbb{R}^{3(n+m)},\mathbb{C}^4)}$$

dès lors que $\Psi = (\Psi_{n,m})_{n,m\geq 0} \in \mathcal{F}$ et $\Psi' = (\Psi'_{n,m})_{n,m\geq 0} \in \mathcal{F}$. Le *vide nu* est alors

$$\Omega_0 := 1 \in \mathcal{F}^{(0,0)} = \mathbb{C}.$$

Sur $\mathcal{F}$ sont définis des opérateurs de création et d'annihilation d'électrons et de positrons. $a_0(f)$ est l'opérateur d'annihilation d'électrons; il vérifie $a_0(f)(\mathcal{F}^{(n,m)}) \subset \mathcal{F}^{(n-1,m)}$ pour $n \geq 1$ et il est défini sur $\mathcal{F}^{(n,m)}$ par

$$a_0(f)(\psi)(x_1,...,x_{n-1};y_1,...,y_m) = \sqrt{n+1}\int_{\mathbb{R}^3} \overline{(1-P^0)(f)} \cdot \psi(x,x_1,...,x_{n-1};y_1,...,y_m)\,dx,$$

et $a_0(f)(\mathcal{F}^{(0,m)}) = 0$. Son adjoint est l'opérateur de création d'électrons $a_0^*(f)$ qui vérifie $a_0^*(f)(\mathcal{F}^{(n,m)}) \subset \mathcal{F}^{(n+1,m)}$ pour $n \geq 0$ et est défini sur les produits tensoriels de $\mathcal{F}^{(n,m)}$ par

$$a_0^*(f)(\psi_+ \otimes \psi_-) = \Big((1-P^0)(f) \wedge \psi_+\Big) \otimes \psi_-$$

23. Rappelons que théoriquement des paires électrons-positrons peuvent spontanément apparaître. Normalement, on ne peut donc pas fixer le nombre de particules dans une théorie relativiste.

pour $\psi_+ \in \mathcal{F}_+^{(n)}$ et $\psi_+ \in \mathcal{F}_-^{(m)}$. Les opérateurs de création $b_0^*(f)$ et d'annihilation $b_0(f)$ de positrons sont définis de façon similaire. Tous ces opérateurs satisfont des relations d'anti-commutation (voir page 225). Notons que le vide nu Ω_0 est caractérisé (à une phase près) par les propriétés

$$\langle \Omega_0, \Omega_0 \rangle = 1, \qquad a_0(f)\Omega_0 = b_0(f)\Omega_0 = 0 \tag{28}$$

pour tout $f \in \mathcal{H}$. Enfin, l'opérateur de champ est défini par la formule

$$\Psi(f) = a_0(f) + b_0^*(f).$$

Cet opérateur doit être vu comme un opérateur d'annihilation. On notera que l'annihilation d'une particule virtuelle (i.e. $f \in \mathcal{H}_0^-$) a été réinterprétée en une création d'un positron, grâce à l'opérateur $b_0^*(f)$. Notons que

$$a_0(f) = \Psi((1 - P^0)f), \qquad b_0^*(f) = \Psi(P^0 f). \tag{29}$$

Il nous faut maintenant *implémenter* l'opérateur $D^{\alpha\varphi}$ dans $\mathcal{F}$ pour obtenir les formules usuelles dans chaque $\mathcal{F}^{(n,m)}$. Pour cela, on définit (formellement) l'opérateur second-quantifié [24]

$$\mathbb{D}^{\alpha\varphi} = \sum_{i,j \in (\mathbb{N}^*)^2} \langle f_i, D^{\alpha\varphi} f_j \rangle \Psi^*(f_i)\Psi(f_j) = \sum_{i \in \mathbb{N}^*} \Psi^*(D^{\alpha\varphi} f_i)\Psi(f_i) \tag{30}$$

où $(f_i)_{i \in \mathbb{N}^*}$ est une base hilbertienne de $\mathcal{H}$ qui est telle que $(f_i)_{i \geq 1}$ et $(f_i)_{i \leq -1}$ forment des bases de $\mathcal{H}_0^+$ et $\mathcal{H}_0^-$. Toutefois, cet opérateur n'est pas bien défini puisque même le vide Ω_0 n'est pas dans le domaine de la forme quadratique associée : $\langle \Omega_0, \mathbb{D}^{\alpha\varphi}\Omega_0 \rangle = \mathrm{tr}(P^0 D^{\alpha\varphi} P^0)$ et $P^0 D^{\alpha\varphi} P^0$ n'est pas trace-class. Il faut donc effectuer une *renormalisation* qui est réalisée en plaçant tous les opérateurs de création à gauche, en faisant comme s'ils anticommutaient. Cette opération est notée $: - :_{P^0}$. Comme

$$\Psi^*(f)\Psi(g) = a_0^*(f)a_0(f) + b_0(g)a_0(f) + a_0^*(f)b_0^*(g) + b_0(g)b_0^*(g),$$

nous obtenons donc

$$: \Psi^*(f)\Psi(g) :_{P^0} = a_0^*(f)a_0(f) + b_0(g)a_0(f) + a_0^*(f)b_0^*(g) - b_0^*(g)b_0(g),$$

$$: \mathbb{D}^{\alpha\varphi} :_{P^0} = \sum_{i,j \in (\mathbb{N}^*)^2} \langle f_i, D^{\alpha\varphi} f_j \rangle : \Psi^*(f_i)\Psi(f_j) :_{P^0}$$

et $\langle \Omega_0, : \mathbb{D}^{\alpha\varphi} :_{P^0} \Omega_0 \rangle = 0$. Les résultats de [HS98] permettent de donner un sens à $: \mathbb{D}^{\alpha\varphi} :_{P^0}$ en tant qu'opérateur essentiellement auto-adjoint sur l'espace de Fock, lorsque le potentiel φ est créé par un noyau étendu

$$\varphi = Zn * \frac{1}{|\cdot|},$$

ce que nous supposerons dans toute la suite.

Lorsque l'on tient compte de l'interaction entre les particules, une étude similaire mène à la formule suivante, pour le Hamiltonien global du système sur l'espace de Fock

$$\mathbb{H} =: \mathbb{D}^{\alpha\varphi} :_{P^0} + \sum_{i,j,k,l \in \mathbb{N}^*} \iint \frac{f_i(x)^* f_k(x) f_j(y)^* f_l(y)}{|x - y|} dx\, dy : \Psi^*(f_i)\Psi^*(f_j)\Psi(f_k)\Psi(f_l) :_{P^0} .$$

24. On obtient par exemple $\mathbb{D}^{\alpha\varphi} f_1 = (1 - P^0) D^{\alpha\varphi} f_1 + C P_-^0 D^{\alpha\varphi} f_1$, qui est un élément de $\mathcal{F}^{(1,0)} \oplus \mathcal{F}^{(0,1)}$.

Nous ne connaissons aucun résultat mathématique permettant de donner un sens à cet opérateur. Notons que : $\mathbb{D}^{\alpha\varphi} :_{P^0}$ est minoré sur son domaine dans $\mathcal{F}$. Les physiciens semblent penser que $\mathbb{H}$ est aussi borné inférieurement[25], et interprètent un minimum de $\mathbb{H}$, $\Omega_{\rm pol} \in \mathcal{F}$, comme le vide polarisé. En effet, le vide doit être l'état de plus basse énergie dans l'espace de Fock [CI89] (on parle de stabilité du vide polarisé). Aucun résultat théorique dans ce sens n'est connu.

Dans le modèle de Bogoliubov-Dirac-Fock, l'opérateur $\mathbb{H}$ est restreint à un sous-ensemble particulier de $\mathcal{F}$ contenant des états BDF, dans l'esprit des approximations de Hartree et Dirac-Fock. Nous obtenons une énergie qui est bornée inférieurement, et possède un minimum qui peut être interprété comme un vide stable parmi les états BDF.

4.3 Le modèle Bogoliubov-Dirac-Fock et l'existence d'un vide BDF-stable

Nous décrivons ici nos résultats [HLS04a, HLS04b] présentés aux Chapitres C.I et C.II, concernant la minimisation de l'énergie Bogoliubov-Dirac-Fock introduite par Chaix et Iracane dans [CI89, CIL89, Cha90].

Nous avons vu précédemment que le vide pouvait être représenté par un projecteur sur le spectre négatif d'un certain opérateur de Dirac. Notons que le choix d'un projecteur P pour représenter le vide implique le choix de deux espaces $\mathcal{H}_P^+$ et $\mathcal{H}_P^-$ vérifiant $\mathcal{H} = \mathcal{H}_P^+ \oplus \mathcal{H}_P^-$, et représentant respectivement les nouveaux espaces électronique et positronique. Pour construire l'espace de Fock, nous avions choisi $P = P^0$ et on parle dans ce cas de *vide nu* et de *particules nues*. Lorsqu'un autre P est choisi, on parle de *vide et particules habillés* [CI89, Cha90].

Pour un certain projecteur orthogonal P, les opérateurs de création et d'annihilation de particules habillées sont aisément définis, en comparaison avec la formule (29), par

$$a_P(f) = \Psi((1-P)f), \quad b_P(f) = \Psi^*(Pf). \tag{31}$$

La question principale est maintenant celle de l'existence d'un *vide habillé* associé au projecteur P dans l'espace de Fock, et qui doit être caractérisé par

$$\langle \Omega_P, \Omega_P \rangle = 1, \qquad a_P(f)\Omega_P = b_P(f)\Omega_P = 0. \tag{32}$$

La réponse est donnée par le célèbre Théorème de Shale-Stinespring [Tha92, Theorem 10.6] qui affirme qu'un tel Ω_P existe (et est alors unique à une phase près), si et seulement si $P - P^0$ est un opérateur de Hilbert-Schmidt, ce que nous noterons usuellement $P - P^0 \in \mathfrak{S}_2(\mathcal{H})$. Nous verrons que la différence $P - P^0$ joue un rôle crucial dans cette théorie car elle s'interprète comme l'opérateur de densité à un corps du vide polarisé. Ainsi, nous pouvons considérer l'ensemble

$$\left\{ e^{i\theta}\Omega_P,\ P - P^0 \in \mathfrak{S}_2 \right\} \subset \mathcal{F}$$

qui décrira les "vides possibles" que nous allons considérer.

Notons que le vide habillé Ω_P est construit dans la preuve du Théorème de Shale-Stinespring comme une rotation du vide nu dans l'espace de Fock, i.e. $\Omega_P = \mathbb{U} \cdot \Omega_0$. La rotation $\mathbb{U}$ est appelée transformation de Bogoliubov (voir [Tha92, KS77, CI89]).

25. Cela très probablement faux sans cut-off puisqu'un exemple de J.P. Solovej permet de montrer que cet opérateur n'est pas borné inférieurement lorsque $m = 0$ (l'invariance par dilatation est utilisée).

La charge du vide, supertrace

Avant de calculer l'énergie d'un vide Ω_P, nous voulons introduire le concept de charge, et de supertrace. L'*opérateur de charge* est donné par [Tha92, Formula (10.52)]

$$\mathcal{Q} = \sum_{i\in\mathbb{N}^*} \Big\{ a_0^*(f_i)a_0(f_i) - b_0^*(f_i)b_0(f_i) \Big\} = \sum_{i\geq 1} a_0^*(f_i)a_0(f_i) - \sum_{i\leq -1} b_0^*(f_i)b_0(f_i).$$

On calcule aisément la charge de Ω_P

$$\langle \Omega_P, \mathcal{Q}\,\Omega_P\rangle = \mathrm{tr}[(1-P^0)(P-P^0)(1-P^0)] + \mathrm{tr}[P^0(P-P^0)P^0].$$

Si $P-P^0$ est trace-class, cette expression est juste $\mathrm{tr}(P-P^0)$ [26]. En revanche, si $P-P^0$ n'est pas trace-class, nous devons introduire la notion suivante.

Définition 2. *Soit P un projecteur tel que P et $1-P$ sont de rang infini, et $A \in \mathfrak{S}_2$. Nous dirons que A est supertrace-class par rapport à P si $A_{--} := PAP$ et $A_{++} := (1-P)A(1-P)$ sont trace-class. Dans ce cas, on pose*

$$\mathrm{str}_P(A) = \mathrm{tr}(A_{++}) + \mathrm{tr}(A_{--}).$$

Nous noterons $\mathfrak{S}_1^P$ l'ensemble des opérateurs Hilbert-Schmidt qui sont supertrace-class par rapport à P.

Évidemment, cette notion généralise la trace puisque lorsque $A \in \mathfrak{S}_1$, alors $A \in \mathfrak{S}_1^P$ pour tout P et $\mathrm{str}_P(A) = \mathrm{tr}(A)$. Mentionnons le résultat suivant, qui est prouvé et abondamment utilisé au Chapitre C.I :

Théorème 7. *Soient P et P' deux projecteurs tels que $P-P' \in \mathfrak{S}_2$. Alors $\mathfrak{S}_1^P = \mathfrak{S}_1^{P'}$ et $\mathrm{str}_P(A) = \mathrm{str}_{P'}(A)$ pour tout $A \in \mathfrak{S}_1^P$.*

Enfin, il nous reste à démontrer que la charge de Ω_P est bien définie. Pour cela, nous prouvons au Chapitre C.I le résultat suivant :

Théorème 8. *Soient P et P' deux projecteurs tels que $Q := P-P' \in \mathfrak{S}_2$. Alors $Q \in \mathfrak{S}_1^P = \mathfrak{S}_1^{P'}$. De plus, $\mathrm{str}_P(Q)$ est un entier qui vérifie*

$$\forall k \geq 1,\ \mathrm{str}_P(Q) = \mathrm{str}_{P'}(Q) = \mathrm{tr}(Q^{2k+1}).$$

Ce résultat, conséquence de la théorie développée par Avron-Seiler-Simon dans [ASS94], est très intéressant car il prouve que la charge d'un vide Ω_P – mesurée par rapport à celle de Ω_0 – est toujours un entier (voir [Hai03] pour une interprétation).

Même si elle a été introduite dans [Tha92, Section 5.7], nous n'avons trouvé dans la littérature aucune mention de cette notion de supertrace appliquée à la charge du vide dans l'espace de Fock.

L'énergie BDF

Commençons par préciser l'ensemble des états que nous allons considérer dans $\mathcal{F}$. Nous allons simplement nous restreindre aux états de Dirac-Fock, c'est-à-dire à des déterminants de Slater d'électrons et de positrons habillés. Ces derniers sont obtenus en appliquant des opérateurs de création sur le vide Ω_P. Un état BDF est donc un état de la forme

$$\psi = a_P^*(f_1)\cdots a_P^*(f_n)\, b_P^*(g_1)\cdots b_P^*(g_m)\,\Omega_P, \tag{33}$$

26. Par exemple, si $P = P^0$, on trouve bien que la charge de Ω_0 est nulle.

où $(f_1,...,f_n) \in (\mathcal{H}_P^+)^n$ et $(g_1,...,g_m) \in (\mathcal{H}_P^-)^m$ sont tels que $\langle f_i,f_j\rangle = \delta_{ij}$, $\langle g_i,g_j\rangle = \delta_{ij}$ (n et m ne sont pas fixés dans cette théorie). L'opérateur de densité d'ordre 1 γ_ψ des particules est alors défini par son noyau

$$\gamma_\psi(x,y) = \sum_{i=1}^{n} f_i(x)\overline{f_i(y)}^T - \sum_{j=1}^{m} g_j(x)\overline{g_j(y)}^T$$

et il vérifie

$$-P \leq \gamma_\psi \leq 1 - P.$$

La densité de charge associée est bien évidemment

$$\rho_\psi(x) = \mathrm{Tr}_{\mathbb{C}^4}\left(\gamma(x,x)\right) = \sum_{i=1}^{n} |f_i(x)|^2 - \sum_{j=1}^{m} |g_j(x)|^2.$$

L'énergie d'un état BDF de la forme (33) dépend alors de l'opérateur de densité à un corps du système "particules + vide", $\Gamma = P - P^0 + \gamma_\psi$, et elle vaut

$$\langle \psi,\mathbb{H}\psi\rangle = \mathcal{E}^{BDF}(P - P^0 + \gamma_\psi), \tag{34}$$

où $\mathcal{E}^{BDF}$ est définie par

$$\mathcal{E}^{BDF}(\Gamma) = \mathrm{str}_{P^0}(D^0\Gamma) - \alpha\int \rho_\Gamma\varphi + \frac{\alpha}{2}\iint \frac{\rho_\Gamma(x)\rho_\Gamma(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\Gamma(x,y)|^2}{|x-y|}dx\,dy. \tag{35}$$

Dans cette égalité, $\rho_\Gamma(x) = \mathrm{Tr}_{\mathbb{C}^4}\left(\Gamma(x,x)\right)$. L'expression (35) est exactement semblable à l'énergie Dirac-Fock (24) ; seul l'ensemble de définition change. Cela provient du fait que, en représentant la mer de Dirac par un projecteur, on a négligé la corrélation entre les électrons virtuels. Le véritable vide polarisé (i.e. le minimiseur de $\langle\psi,\mathbb{H}\psi\rangle$ s'il existe) n'est en général pas de la forme Ω_P, avec P projecteur orthogonal.

Comme en général $P-P^0 \in \mathfrak{S}_1^{P^0}$ n'est pas trace-class, la fonction $\rho_\Gamma(x) = \mathrm{Tr}_{\mathbb{C}^4}\left(\Gamma(x,x)\right)$ n'est pas bien définie. Le problème est posé de façon cruciale dans la représentation de Furry, lorsque $P = P^{\alpha\varphi}$, puisque dans ce cas ρ_{P-P^0} diverge en tout point. Afin de donner un sens à l'expression (35), nous allons donc maintenant ajouter un cut-off dans l'espace de Fourier, c'est-à-dire, nous travaillons définitivement dans l'espace [27]

$$\mathcal{H}_\Lambda := \left\{ f \in \mathcal{H},\, \mathrm{supp}(\widehat{f}) \subset B(0,\Lambda) \right\}.$$

Dans ce cadre, la fonction ρ_Γ est bien définie par la formule

$$\rho_\Gamma(x) = \mathrm{Tr}_{\mathbb{C}^4}\left(\Gamma(x,x)\right) = \frac{1}{(2\pi)^3}\iint_{|p|,|q|\leq\Lambda} \mathrm{Tr}_{\mathbb{C}^4}\left(\widehat{\Gamma(p,q)}\right)e^{ix(p-q)}dp\,dq. \tag{36}$$

Notons qu'un cut-off ultraviolet est utilisé (même implicitement) dans tous les ouvrages de physique. Comme nous obtiendrons une contrainte de la forme $\alpha\sqrt{\log\Lambda} \leq C$ pour une certaine constante C, cela autorise un Λ très grand lorsque $\alpha \simeq 1/137$, bien au-delà de la validité des modèles physiques actuels.

Nous voulons montrer que l'énergie BDF est minorée, et admet un minimum de la forme $P - P^0$, où P est le projecteur associé au spectre négatif d'un opérateur de Dirac

27. Comme D^0 est un opérateur de multiplication dans l'espace de Fourier, $\mathcal{H}_\Lambda$ est invariant par le projecteur P^0. Nous conservons donc la notation P^0 et évitons l'introduction d'un opérateur restreint

de champ moyen. Pour cela, nous allons plutôt travailler sur l'enveloppe convexe de $\{\Gamma = P - P^0 + \gamma_\psi\}$, et introduisons l'ensemble

$$\mathcal{B}_\Lambda := \left\{\Gamma \in \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda),\ -P^0 \leq \Gamma \leq 1 - P^0,\ \rho_\Gamma \in \mathcal{C}\right\}$$

avec

$$\mathcal{C} = \left\{\rho \in L^2(\mathbb{R}^3),\ \int_{\mathbb{R}^3} \frac{|\widehat{\rho}(k)|^2}{|k|^2}\, dk < \infty\right\}, \qquad \|\rho\|_{\mathcal{C}}^2 := \int_{\mathbb{R}^3} \frac{1+|k|^2}{|k|^2} |\widehat{\rho}(k)|^2\, dk.$$

La contrainte sur ρ_Γ permet de donner un sens aux termes qui en dépendent. Notre résultat est alors le suivant (voir le Théorème I.2 page 194)

Théorème 9. *Soient $n \in \mathcal{C}$, $Z \in \mathbb{R}$, et $\varphi = Zn * \frac{1}{|\cdot|}$.*

1. *$\mathcal{E}^{BDF}$ est bien définie et continue sur*
$$\mathcal{A}_\Lambda := \left\{\Gamma \in \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda),\ \rho_\Gamma \in \mathcal{C}\right\} \supset \mathcal{B}_\Lambda.$$
2. *Si $0 \leq \alpha \leq \frac{4}{\pi}$, alors $\mathcal{E}^{BDF}$ est bornée inférieurement sur $\mathcal{B}_\Lambda$, indépendamment de Λ.*
3. *Si $0 \leq \alpha \leq \frac{4}{\pi}$ et $Z = 0$, alors $\mathcal{E}^{BDF}$ est positive sur $\mathcal{B}_\Lambda$ [BBHS99], avec 0 comme unique minimiseur.*

Comme nous cherchons un minimum de la forme $P - P^0$, nous introduisons la

Définition 3. *Un projecteur P tel que $P - P^0 \in \mathcal{B}_\Lambda$ est appelé* vide BDF-stable *si $P - P^0$ est un minimum de $\mathcal{E}^{BDF}$ sur $\mathcal{B}_\Lambda$.*

Ω_P est donc dans ce cas l'état d'énergie minimale parmi les états BDF (33). Le dernier point du Théorème 9 a déjà été prouvé dans [BBHS99]. Il signifie que P^0 est un vide BDF-stable en l'absence de potentiel extérieur, ce qui correspond à l'interprétation de Dirac. La constante $4/\pi$ est optimale dans le sens où $\mathcal{E}^{BDF}$ n'est plus minorée lorsque $\alpha > 4/\pi$ et $Z = 0$ [HRS00]. Lorsque $Z \neq 0$, il est facile de voir que 0 ne minimise plus $\mathcal{E}^{BDF}$, ce qui signifie que le vide est nécessairement polarisé.

Remarque : Dans Chaix-Iracane [CI89], formule (4.8), l'énergie (34) est développée sous la forme (rappelons que $Q = P - P^0$)

$$\mathcal{E}^{BDF}(Q+\gamma_\psi) = \mathrm{tr}\,(D_Q\gamma_\psi) + \frac{\alpha}{2}\iint \frac{\rho_\psi(x)\rho_\psi(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\gamma_\psi(x,y)|^2}{|x-y|}dx\,dy + \mathcal{E}^{BDF}(Q) \tag{37}$$

$$D_Q = D^{\alpha\varphi} + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}. \tag{38}$$

Le terme $\mathcal{E}^{BDF}(Q)$ représente l'énergie du vide polarisé, tandis que les deux termes dépendant de Q dans D_Q sont les potentiels créés par la polarisation du vide. Q s'interprète donc bien comme la matrice de densité de Ω_P. Dans [BBHS99, Formule (21)], tous ces termes dépendant de Q ont été négligés par les auteurs, qui étudient la fonctionnelle

$$\mathcal{E}'_P(\gamma) = \mathrm{tr}\,(D^{\alpha\varphi}\gamma) + \frac{\alpha}{2}\iint \frac{\rho(x)\rho(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\gamma(x,y)|^2}{|x-y|}dx\,dy,$$

sous la contrainte $-P \leq \gamma \leq 1-P$. Elle n'est pas bornée inférieurement : $\inf_P \inf_{-P\leq\gamma\leq 1-P} \mathcal{E}'_P(\gamma) = -\infty$. Pour cette raison, un principe de la forme $\sup_P \inf_{-P\leq\gamma\leq 1-P} \mathcal{E}'_P(\gamma)$ inspiré de [Mit81] est proposé dans [BBHS99, BFHS04], dont la solution est $P = P^{\alpha\varphi}$.

En fait, comme cela est expliqué dans [CI89], les termes de polarisation du vide, même s'ils sont négligeables en pratique dans beaucoup de cas, sont indispensables pour obtenir une énergie bornée inférieurement. Par ailleurs, comme nous allons le voir dans quelques instants, le minimiseur n'est pas $P^{\alpha\varphi}$, même s'il en est proche en un certain sens.

Existence d'un vide BDF-stable : l'approche de type point-fixe

On peut démontrer qu'un vide BDF-stable P satisfait l'équation d'Euler-Lagrange

$$\boxed{P = \chi_{(-\infty;0)}(D_Q) = \chi_{(-\infty;0)}\left(D^{\alpha\varphi} + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}\right).} \tag{39}$$

Il s'agit d'une équation auto-consistante similaire aux équations HF (6,7) et DF (25). Le vide doit correspondre au projecteur P associé au spectre négatif d'un opérateur de Dirac tenant compte des potentiels de polarisation, dépendant eux-mêmes de P. Cette équation peut être résolue par une méthode de type point fixe. Toutefois, le point critique trouvé n'est pas *a priori* un minimiseur de l'énergie (pour HF, une solution de (7) n'est pas forcément un minimum global). Ici, cette propriété est vraie au moins lorsque α est assez petit puisque nous prouvons au Chapitre C.I (Théorème I.3 page 196), le

Théorème 10 (BDF-Stabilité). *Soient $n \in \mathcal{C}$, $Z \in \mathbb{R}$, et P un projecteur tel que $P - P^0 \in \mathcal{B}_\Lambda$. Nous supposons qu'il existe une constante strictement positive d telle que*

$$d|D_Q| \geq |D^0| \quad \textit{avec} \quad \alpha d\frac{\pi}{4} \leq 1, \tag{40}$$

où D_Q est défini par la formule (38)*. Alors les assertions suivantes sont équivalentes*

1. *P est solution de l'équation* (39) *;*
2. *P est BDF-stable, i.e. $P - P^0$ est un minimiseur de $\mathcal{E}^{BDF}$ sur $\mathcal{B}_\Lambda$.*

La preuve de ce résultat est basée sur des idées de [BBHS99], avec des difficultés supplémentaires dues au fait que les opérateurs considérés ne sont pas trace-class. Voici maintenant notre résultat principal, exposé au Chapitre C.I (Théorème I.4 page 196),

Théorème 11 (Existence d'un vide BDF-stable). *Soit $2 - \sqrt{2} \leq b < 1$, $Z \in \mathbb{R}$ et $n \in \mathcal{C}$. Alors il existe une fonction $\Lambda \mapsto \alpha_b(\Lambda) > 0$ telle que, lorsque α vérifie*

$$4\sqrt{\pi}\alpha Z\|n\|_{\mathcal{C}} \leq b \quad \textit{et} \quad \alpha \leq \alpha_b(\Lambda)$$

avec

$$\alpha_b(\Lambda) \sim_{\Lambda\to\infty} \frac{C(1-b)}{\sqrt{\log\Lambda}}, \qquad C \geq 0{,}798,$$

alors il existe un unique vide BDF-stable P dans une certaine boule $B(P^0,R_0)$ contenant $P^{\alpha\varphi}$, qui est une solution de

$$P = \chi_{(-\infty;0)}\left(D^0 + \alpha(\rho_Q - Zn) * \frac{1}{|\cdot|} + \alpha\frac{Q(x,y)}{|x-y|}\right)$$

avec $Q = P - P^0$. De plus, P est un vide polarisé mais non chargé : $\mathrm{str}_{P^0}(P - P^0) = 0$.

Notons que l'on peut démontrer très facilement que si $\alpha \to 0$ et $Z \to \infty$ de sorte que $\alpha Z = \kappa$ est fixé tel que $4\sqrt{\pi}\kappa\|n\|_{\mathcal{C}} < 1$, alors notre solution P_α converge vers P^φ avec $\varphi = \kappa n * \frac{1}{|\cdot|}$. Ce point justifie l'utilisation fréquente de la représentation de Furry dans les calculs numériques, c'est-à-dire le choix $P = P^{\alpha\varphi}$ pour la description d'atomes lourds ($Z >> 1$).

Comme nous l'avons dit, la contrainte $\alpha \leq \alpha_b(\Lambda)$ se réduit essentiellement à $\alpha\sqrt{\log\Lambda} \leq C'$, ce qui permet un Λ très grand. Notons qu'une condition de la forme $\alpha\log\Lambda \leq (3\pi)/2$ est trouvée dans tous les ouvrages physiques (elle provient du fait que seuls les termes du premier ordre en α sont souvent considérés). Notre condition est donc meilleure que cette dernière. En fait, nous pensions que la condition portant sur le cut-off Λ était purement technique (elle permet la contraction de l'application utilisée pour le point fixe). Cette intuition est confirmée par le résultat que nous annonçons au Chapitre C.II et que nous allons présenter maintenant.

Existence d'un vide BDF-stable : la minimisation directe

L'intérêt d'une approche de type point fixe est grand car cette méthode est celle qui est utilisée numériquement. Au Chapitre C.II, dans lequel nous présentons nos résultats dans un langage plus proche de la physique, nous annonçons le résultat suivant, qui n'est pas démontré dans cette thèse et fera l'objet d'une publication ultérieure.

Théorème 12 (Existence d'un vide BDF-stable (en préparation)). *Supposons que $n \in \mathcal{C}$, $Z \in \mathbb{R}$ et que $0 \leq \alpha \leq 4/\pi$. Alors il existe un vide BDF-stable P, c'est-à-dire un minimiseur de $\mathcal{E}^{BDF}$ sur $\mathcal{B}_\Lambda$ de la forme $Q = P - P^0$ où P est solution de l'équation*

$$P = \chi_{(-\infty;0)}\left(D^0 + \alpha(\rho_Q - Zn) * \frac{1}{|\cdot|} + \alpha\frac{Q(x,y)}{|x-y|}\right).$$

Nous voyons donc que la condition liant α et Λ a disparu : Λ doit seulement être fini. Ce résultat est obtenu par une méthode de minimisation directe utilisant des techniques de type concentration-compacité [Lio84].

La charge de P n'est, en général, pas nulle (sauf pour α petit) car des paires électron-positron peuvent spontanément apparaître lorsque le champ extérieur φ est très fort. Cette charge peut être exprimée en fonction du nombre de valeurs propres qui "traversent" 0 lorsque l'on déforme P^0 en P en variant α à partir de 0 (voir [Hai03]).

4.4 Conclusions

En suivant [CI89, CIL89], nous avons introduit le modèle de Bogoliubov-Dirac-Fock qui permet la description de la polarisation du vide, représenté par un projecteur sur le sous-espace associé au spectre négatif d'un opérateur de Dirac. Nous avons prouvé que l'énergie BDF est minorée et admet un minimiseur qui correspond au vide polarisé. Ce minimiseur peut être calculé grâce à un algorithme de point fixe à partir de P^0.

L'étude de la polarisation du vide dans un champ extérieur n'est qu'une première étape : le modèle BDF permet aussi la description de molécules. Dans ce cas, il faut étudier le problème variationnel

$$\min_{\substack{\Gamma \in \mathcal{B}_\Lambda \\ \mathrm{str}_{P^0}(\Gamma) = q}} \mathcal{E}^{BDF}(\Gamma),$$

où $q \in \mathbb{N}$ est la charge du système étudié[28]. Les solutions sont de la forme $\Gamma = P' - P^0$ où P' est solution de l'équation $P' = \chi_{(-\infty;0)}(D_\Gamma - \lambda) = \chi_{(-\infty;\lambda)}(D_\Gamma)$, λ étant le multiplicateur d'Euler-Lagrange associé à la contrainte de charge. Comme $\mathrm{str}_{P^0}(\Gamma) = q$, lorsque Z n'est pas trop grand, on aura $\mathrm{str}_{P^0}(P - P^0) = 0$ où $P = \chi_{(-\infty;0)}(D_\Gamma)$, de sorte que cela implique

$$P' = P + \gamma_\psi, \qquad \gamma_\psi = \sum_{i=1}^{q} |\varphi_i\rangle\langle\varphi_i|$$

où les φ_i sont les q premières fonctions propres de D_Γ (voir la Figure 12).

Les φ_i satisfont donc l'équation

$$D_\Gamma \varphi_i = \varepsilon_i \varphi_i, \qquad \varepsilon_i \in (0;1)$$

où

$$D_\Gamma = D^{\alpha\varphi} + \alpha\rho_\psi * \frac{1}{|\cdot|} - \alpha\frac{\gamma_\psi(x,y)}{|x-y|} + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}$$

28. Rappelons que l'on ne peut pas fixer le nombre de particules, mais plutôt la charge totale. L'étude peut être également menée avec $q \in \mathbb{R}$ mais dans ce cas le minimiseur n'est plus nécessairement un projecteur.

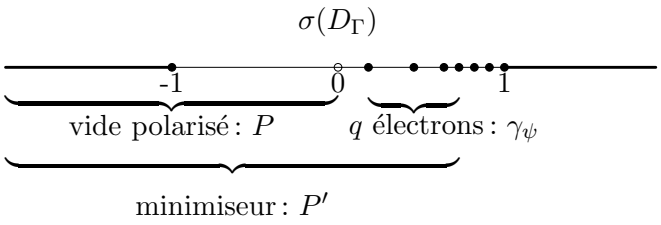


FIG. 12 – *La solution du problème de minimisation BDF*

avec $Q = P - P^0$. En d'autres termes, les φ_i satisfont des équations de Dirac-Fock perturbées par les potentiels de polarisation du vide, qui sont du second ordre en α.

Ceci permet donc de relier le modèle Dirac-Fock et la QED et explique pourquoi DF fournit de bons résultats en pratique. Notons bien que les φ_i et le vide polarisé associé P sont toutefois obtenus par un procédé de *minimisation*, un gain considérable par rapport à la théorie DF, qui ouvre une nouvelle voie pour l'étude de ces modèles.

5 Conclusions et perspectives

Nous avons déjà commenté les résultats obtenus dans chacune des trois parties. Nous souhaitons mentionner ici deux problèmes intéressants qui prolongent notre travail.

Le premier concerne les modèles multi-configurations relativistes. Comme pour le cas non-relativiste, les modèles de champ moyen de type DF ou BDF ne décrivent pas correctement la corrélation, et l'utilisation de modèles plus perfectionnés est indispensable dans certaines situations. Toutefois, aucun modèle clair de type MC tenant compte des effets relativistes n'est actuellement connu. En effet, l'utilisation de projecteurs est indispensable pour éviter un effondrement dans le spectre négatif mais, contrairement à DF, ceux-ci sont pour l'instant déterminés de manière un peu empirique [VIK98, Ind95] : comme les orbitales ne sont pas des fonctions propres d'un opérateur de champ moyen, aucun projecteur commun à toutes celles-ci n'apparait naturellement dans cette théorie.

Nous pensons que l'ajout de la polarisation du vide devrait permettre de lever cette indétermination. On peut en effet considérer un modèle où les électrons sont modélisés par plusieurs déterminants, et le vide reste lui représenté par un projecteur. Il s'agit donc d'une approximation de type CASSCF (12) car les particules virtuelles de la mer de Dirac sont traitées par un modèle de champ moyen. Cette méthode fournirait un projecteur identique pour toutes les orbitales, et pourrait être très intéressante d'un point de vue numérique. Une collaboration a été entamée sur ce sujet avec Éric Séré et Heinz Siedentop (LMU, Münich).

Un autre problème lié à notre travail de la Partie C concerne l'étude des défauts dans les cristaux. Un cristal est composé d'un réseau périodique sur lequel sont placés des noyaux chargés qui créent un potentiel électrostatique périodique V, et d'électrons qui sont soumis à ce potentiel. Le spectre de l'opérateur de Schrödinger périodique $-\Delta/2 + V$ est alors composé de bandes, chacune représentant un état électronique pour le réseau (voir par exemple [AM76]). Ainsi, un cristal comprenant k électrons par cellule du réseau en moyenne remplira totalement les k premières bandes du spectre. Évidemment, il est

faux d'imaginer les électrons comme réellement localisés dans chaque cellule, et on doit plutôt les voir comme un nuage réparti plus ou moins uniformément dans le cristal : c'est la *mer de Fermi*, qui présente des caractéristiques très similaires à la *mer de Dirac* car toutes deux sont modélisées par un intervalle du spectre continu d'un opérateur.

Si on suppose maintenant qu'il existe un défaut dans le cristal, comme par exemple un noyau plus lourd que les autres, on doit ajouter le potentiel électrostatique w créé par cette charge supplémentaire, qui satisfait $\lim_{|x|\to\infty} w(x) = 0$ (voir par exemple [Bir95, Klo91, Klo90, Out87, Dim94] pour des résultats sur ce sujet). L'étude d'un modèle autoconsistant pour la mer de Fermi soumise à ce potentiel est formellement très similaire à notre étude de la mer de Dirac dans le champ créé par un noyau à la Partie C. Une collaboration avec Éric Cancès et Éric Séré est prévue sur ce sujet.

Bibliographie

[AF97] P.W. Atkins and R.S. Friedman, *Molecular quantum mechanics*, third ed., Oxford University Press, 1997.

[AM76] N. Ashcroft and N. Mermin, *Solid state physics*, Saunders, Philadelphia, 1976.

[AR73] A. Ambrosetti and P.H. Rabinowitz, *Dual variation methods in critical point theory and applications*, J. Functional Analysis **14** (1973), 349–381.

[ASS94] J. Avron, R. Seiler, and B. Simon, *The index of a pair of projections*, J. Funct. Anal. **120** (1994), 220–237.

[Bah89] A. Bahri, *Critical points at infinity in some variational problems*, Pitman Research Notes in Mathematics Series, vol. 182, Longman Scientific and Technical Ed., 1989.

[BBHS99] V. Bach, J.M. Barbaroux, B. Helffer, and H. Siedentop, *On the stability of the relativistic electron-positron field*, Commun. Math. Phys. **201** (1999), 445–460.

[BD64] J.D. Bjorken and S.D. Drell, *Relativistic quantum mechanics*, McGraw-Hill, 1964.

[BFHS04] J.M. Barbaroux, W. Farkas, B. Helffer, and H. Siedentop, *On the Hartree-Fock equations of the electron-positron field*, preprint arXiv : math-ph/0404009, 2004.

[Bir95] M. Sh. Birman, *The discrete spectrum in gaps of the perturbed periodic Schrödinger operator. I. Regular perturbations*, Boundary value problems, Schrödinger operators, deformation quantization, Math. Top., vol. 8, Akademie Verlag, Berlin, 1995, pp. 334–352.

[BLS94] V. Bach, E.H. Lieb, and J.P. Solovej, *Generalized Hartree-Fock theory and the Hubbard model*, J. Statist. Phys. **76** (1994), no. 1-2, 3–89.

[BO27] M. Born and R. Oppenheimer, *Quantum theory of molecules*, Ann. Physics **84** (1927), 457–484.

[BP87] J. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. AMS **303** (1987), 517–527.

[C+03] E. Cancès et al., *Computational quantum chemistry: a primer*, Handbook of Numerical Analysis, vol. X, pp. 3–270, Elsevier, 2003.

[Can98] E. Cancès, *Simulation moléculaire et effets d'environnement. Une perspective mathématique et numérique*, Ph.D. thesis, École Nationale des Ponts et Chausées, 1998.

[CB00] E. Cancès and C. Le Bris, *On the convergence of SCF algorithms for the Hartree-Fock equations*, M2AN Math. Model. Numer. Anal. **34** (2000), no. 4, 749–774.

[CGL04] E. Cancès, H. Galicher, and M. Lewin, *Computing electronic structures: a new multiconfiguration approach for excited states*, preprint, 2004.

[Cha90] P. Chaix, *Une méthode de champ moyen relativiste et application à l'etude du vide de l'electrodynamique quantique*, Ph.D. thesis, 1990.

[CI89] P. Chaix and D. Iracane, *From quantum electrodynamics to mean field theory: I. the Bogoliubov-Dirac-Fock formalism.*, J. Phys. B. **22** (1989), no. 23, 3791–3814.

[CIL89] P. Chaix, D. Iracane, and P.L. Lions, *From quantum electrodynamics to mean field theory: II. Variational stability of the vacuum of quantum electrodynamics in the mean-field approximation.*, J. Phys. B. **22** (1989), no. 23, 3815–3828.

[CL92] I. Catto and P.-L. Lions, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part I.*, Comm. Part. Diff. Eq. **17** (1992), no. 7-8, 1051–1110.

[CL93a] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part II.*, Comm. Part. Diff. Eq. **18** (1993), no. 1-2, 305–354.

[CL93b] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part III.*, Comm. Part. Diff. Eq. **18** (1993), no. 3-4, 381–429.

[CL93c] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-fermi type theories. Part IV.*, Comm. Part. Diff. Eq. **18** (1993), no. 7-8, 1149–1159.

[Com75] J.M. Combes, *On the Born-Oppenheimer approximation*, International Symposium on Mathematical Problems in Theoretical Physics (Kyoto Univ., Kyoto, 1975), Springer, Berlin, 1975, pp. 467–471. Lecture Notes in Phys., 39.

[CS80] J.-M. Combes and R. Seiler, *Spectral properties of atomic and molecular systems*, Quantum dynamics of molecules (Proc. NATO Adv. Study Inst., Univ. Cambridge, Cambridge, 1979), NATO Adv. Study Inst. Ser., Ser. B: Physics, vol. 57, Plenum, New York, 1980, pp. 435–482.

[CS01] O. Costin and A. Soffer, *Resonance theory for Schrödinger operators*, Commun. Math. Phys. **224** (2001), 133–152.

[Dim94] M. Dimassi, *Développements asymptotiques pour des perturbations fortes de l'opérateur de Schrödinger périodique*, Ann. Inst. H. Poincaré Phys. Théor. **61** (1994), no. 2, 189–204.

[Dir33] P.A.M. Dirac, *Theory of electrons and positrons*, Nobel lecture delivered at Stockholm, 1933.

[Dir34a] ———, *Discussion of the infinite distribution of electrons in the theory of the positron*, Proc. Camb. Philos. Soc. **30** (1934), 150–163.

[Dir34b] ———, *Théorie du positron*, Solvay report **XXV** (1934), 203–212.

[ES99] M.J. Esteban and E. Séré, *Solutions of the Dirac-Fock equations for atoms and molecules*, Comm. Math. Phys. **203** (1999), 499–530.

[ES01] ———, *Nonrelativistic limit of the Dirac-Fock equations*, Ann. Henri Poincaré **2** (2001), no. 5, 941–961.

[ES02] ———, *An overview on linear and nonlinear Dirac equations*, Disc. and Cont. Dynam. Systems **8** (2002), no. 2, 381–397.

[ES03] ———, *Dirac-Fock models for atoms and molecules and related topics*, Proc. ICMP, 2003.

[FG92] G. Fang and N. Ghoussoub, *Morse-type information on Palais-Smale sequences obtained by min-max principles*, Manuscripta Math. **75** (1992), 81–95.

[Foc30] V. Fock, *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems*, Z. Phys. **61** (1930), 126–148.

[Fri03] G. Friesecke, *The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions*, Arch. Rat. Mech. Analysis **169** (2003), 35–71.

[Gho93] N. Ghoussoub, *Duality and perturbation methods in critical point theory*, Cambridge University Press, 1993.

[Hai03] C. Hainzl, *On the vacuum polarization density caused by an external field*, 2003.

[Har28] D. Hartree, *The wave-mechanics of an atom with a non-coulomb central field. Part I. Theory and methods.*, Proc. Camb. Phil. Soc. **24** (1928), 89–312.

[HG96] M. Head-Gordon, *Quantum chemistry and molecular processes*, J. Phys. Chem. **100** (1996), 13213–13225.

[HLS04a] C. Hainzl, M. Lewin, and E. Séré, *Existence of a stable polarized vacuum in the Bogoliubov-Dirac-Fock approximation*, Preprint ArXiv: `math-ph/0403005`, 2004.

[HLS04b] ———, *Self-consistent solution for the polarized vacuum in a no-photon QED model*, Preprint ArXiv: `physics/0404047`, 2004.

[HRS00] D. Hundertmark, N. Röhrl, and H. Siedentop, *The sharp bound on the stability of the relativistic electron-positron field in Hartree-Fock approximation*, Comm. Math. Phys. **211** (2000), no. 3, 629–642.

[HS96] P.D. Hislop and I.M Sigal, *Introduction to spectral theory with application to Schrödinger operators*, Springer (New York), 1996.

[HS98] B. Helffer and H. Siedentop, *Form perturbation of the second quantized Dirac field*, Mathematical Physics Electronic Journal **4** (1998).

[HS00] W. Hunziker and I.M. Sigal, *The quantum N-body problem*, J. Math. Phys **41** (2000), no. 6, 3348–3509.

[Hun66] W. Hunziker, *On the spectra of Schrödinger multiparticle Hamiltonians*, Helv. Phys. Acta **39** (1966), 451–462.

[Ind95] P. Indelicato, *Projection operators in Multiconfiguration Dirac-Fock calculations. Application to the ground state of heliumlike ions*, Phys. Rev. A **51** (1995), no. 2, 1132–1145.

[Kat51] T. Kato, *Fundamental properties of hamiltonian operators of Schrödinger type*, Trans. Am. Math. Soc **70** (1951), 195–221.

[Kat57] ———, *On the eigenfunctions of many particle systems in quantum mechanics*, Commun. Pure and Appl. Math. **10** (1957), 151–171.

[Klo90] F. Klopp, *Impuretés dans une structure périodique*, Journées "Équations aux Dérivées Partielles" (Saint Jean de Monts, 1990), École Polytech., Palaiseau, 1990, pp. Exp. No. XX, 12.

[Klo91] ———, *Étude semi-classique d'une perturbation d'un opérateur de Schrödinger périodique*, Ann. Inst. H. Poincaré Phys. Théor. **55** (1991), no. 1, 459–509.

[KMSW92] M. Klein, A. Martinez, R. Seiler, and X. P. Wang, *On the Born-Oppenheimer expansion for polyatomic molecules*, Comm. Math. Phys. **143** (1992), no. 3, 607–639.

[KS77] M. Klaus and G. Scharf, *The regular external field problem in quantum electrodynamics*, Helv. Phys. Acta **50** (1977), 779–802.

[LB93] C. Le Bris, *Quelques problèmes mathématiques en chimie quantique moléculaire*, Ph.D. thesis, École Polytechnique, 1993.

[LB94] ———, *A general approach for multiconfiguration methods in quantum molecular chemistry*, Ann. Inst. H. Poincaré Anal. Non linéaire **11** (1994), no. 6, 441–484.

[Lew02] M. Lewin, *The multiconfiguration methods in quantum chemistry: Palais-Smale condition and existence of minimizers*, C. R. Acad. Sci. Paris, Ser. I **334** (2002), 299–304.

[Lew04a] ———, *A mountain pass for reacting molecules*, Ann. Henri Poincaré (2004).

[Lew04b] ———, *Solutions of the multiconfiguration equations in quantum chemistry*, Arch. Rat. Mech. Anal. **171** (2004), no. 1, 83–114.

[Lie84] E.H. Lieb, *Bound on the maximum negative ionization of atoms and molecules*, Phys. Rev. A. **29** (1984), no. 6, 3018–3028.

[Lie90] ———, *The stability of matter : from atoms to stars*, Bull. Amer. Math. Soc. **22** (1990), 1–49.

[Lio84] P.L. Lions, *The concentration-compactness principle in the calculus of variations. The locally compact case, part i*, Ann. Inst. Henri Poincaré **1** (1984), no. 2, 109–149.

[Lio87] ———, *Solutions of Hartree-Fock equations for Coulomb systems*, Commun. Math. Phys. **109** (1987), 33–87.

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

[LSS97] E.H. Lieb, H. Siedentop, and J.P. Solovej, *Stability and instability of relativistic electrons in classical electromagnetic fields*, J. Stat. Phys. **89** (1997), no. 1–2, 37–59.

[LT86] E.H. Lieb and W.E. Thirring, *Universal nature of Van Der Waals forces for Coulomb systems*, Phys. Rev. A **34** (1986), 40–46.

[Löw55a] 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** (1955), no. 6, 1474–1489.

[Löw55b] ———, *Quantum theory of many-particle systems. II. Study of the ordinary Hartree-Fock approximation*, Phys. Rev. **97** (1955), no. 6, 1490–1508.

[Löw55c] ———, *Quantum theory of many-particle systems. III. extension of the Hartree-Fock scheme to include degenerate systems and correlation effects*, Phys. Rev. **97** (1955), no. 6, 1509–1520.

[Löw59] ———, *Correlation problem in many-electron quantum mechanics. I. Review of different approaches and discussion of some current ideas*, Adv. Chem. Phys. **2** (1959), 207–322.

[Mit81] M.H. Mittleman, *Theory of relativistic effects on atoms: Configuration-space hamiltonian*, Phys. Rev. A **24** (1981), no. 3, 1167–1175.

[MS99] M. Mekli and I.M. Sigal, *A time-dependent theory of quantum resonances*, Commun. Math. Phys. **201** (1999), 549–576.

[Out87] A. Outassourt, *Comportement semi-classique pour l'opérateur de Schrödinger à potentiel périodique*, J. Funct. Anal. **72** (1987), no. 1, 65–93.

[Pat00] E. Paturel, *Solutions of the Dirac-Fock equations without projector*, Ann. Henri Poincaré **1** (2000), no. 6, 1123–1157.

[QH84] W. Quapp and D. Heidrich, *Analysis of the concept of minimum energy path on the potential energy surface of chemically reacting systems*, Theoret. Chim. Acta **66** (1984), 245–260.

[Roo87] B. O. Roos, *The Complete Active Space Self-Consistent Field method and its applications in electronic structure calculation*, Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys. **69** (1987), 399–446.

[RS78] M. Reed and B. Simon, *Methods of modern mathematical physics*, vol. IV – Analysis of Operators, Academic Press, 1978.

[She87] R. Shepard, *The Multiconfiguration Self-Consistent Field method*, Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys. **69** (1987), 63–200.

[Sig82] I.M. Sigal, *Geometric methods in the quantum many-body problem. Non existence of very negative ions*, Commun. Math. Phys. **85** (1982), 309–324.

[Sim77] B. Simon, *Geometric methods in multiparticle quantum systems*, Commun. Math. Phys. **55** (1977), 259–274.

[Sla30] J.C. Slater, *A note on Hartree's method*, Phys. Rev. **35** (1930), 210–211.

[Str98] P. Strange, *Relativistic quantum mechanics*, Cambridge University Press, 1998.

[Suc80] J. Sucher, *Foundations of the relativistic theory of many-particle atoms*, Phys. Rev. A **22** (1980), no. 2, 348–362.

[Suc87] ———, *Relativistic many-electrons Hamiltonians*, Phys. Scrypta **36** (1987), 271–281.

[SW98] A. Soffer and M.I. Weinstein, *Time dependent resonance theory*, GAFA, Geom. Funct. Anal. **8** (1998), 1086–1128.

[Tha92] B. Thaller, *The Dirac equation*, Springer-Verlag, 1992.

[Thi02] W. Thirring, *Quantum mathematical physics*, vol. Atoms, Molecules and Large Systems, Springer, Second Edition 2002.

[Van64] C. Van Winter, *Theory of finite systems of particles. I. The Green function*, Mat.-Fys. Skr. Danske Vid. Selsk. **2** (1964), no. 8.

[VIK98] M.J. Vilkas, Y. Ishikawa, and K. Koc, *Quadratically convergent multiconfiguration Dirac-Fock and multireference relativistic configuration-interaction calculations for many-electrons systems*, Phys. Rev. E **58** (1998), no. 4, 5096–5110.

[vN32] J. von Neumann, *Mathematishe grundlagen der quantenmechanik*, Springer Verlag (Berlin), 1932.

[Zei95] E. Zeidler, *Applied functional analysis, applications to mathematical physics*, Springer (New York), 1995.

[Zhi60] G. M. Zhislin, *Discussion of the spectrum of Schrödinger operators for systems of many particles. (in Russian)*, Trudy Moskovskogo matematiceskogo obscestva **9** (1960), 81–120.

[Zhi71] G.M. Zhislin, *On the finiteness of the discrete spectrum of the energy operator of negative atomic and molecular ions*, Teoret. Mat. Fiz. **21** (1971), 332–341.

Première partie

Le problème électronique : méthodes de type multi-configurations

A.I

Solutions of the Multiconfiguration Equations in Quantum Chemistry

Ce chapitre reprend le texte intégral d'un article paru dans *Archive for Rational Mechanics and Analysis*, Volume 171 (2004), pages 83–114.

Il a été précédé d'une note aux *Comptes Rendus de l'Académie des Sciences, Série I*, intitulée "The multiconfiguration methods in quantum chemistry: Palais-Smale condition and existence of minimizers", Volume 334 (2002), pages 299–304.

Résumé

Les méthodes multi-configurations généralisent naturellement la théorie de Hartree-Fock, fréquemment utilisée pour la modélisation d'électrons non relativistes dans les atomes et les molécules. Avec une méthode variationnelle, nous prouvons l'existence d'un minimum de l'énergie et d'une infinité de solutions aux équations multi-configurations. Seules certaines d'entre elles, en nombre fini, peuvent être interprétées comme des états excités de la molécule. Nos résultats sont valables lorsque la charge totale des noyaux Z est supérieure au nombre d'électrons N, et s'appliquent à la plupart des méthodes utilisées dans la pratique par les chimistes. Les points critiques sont obtenus grâce à des principes de min-max. Nous utilisons une condition de type Palais-Smale avec information de Morse, et une nouvelle forme très simple des équations d'Euler-Lagrange.

Solutions of the Multiconfiguration Equations in Quantum Chemistry

Mathieu Lewin

Communicated by G. Friesecke

Abstract

The multiconfiguration methods are the natural generalization of the well-known Hartree-Fock theory for atoms and molecules. By a variational method, we prove the existence of a minimum of the energy and of infinitely many solutions of the multiconfiguration equations, a finite number of them being interpreted as excited states of the molecule. Our results are valid when the total nuclear charge Z exceeds $N-1$ (N is the number of electrons) and cover most of the methods used by chemists. The saddle points are obtained with a min-max principle; we use a Palais-Smale condition with Morse-type information and a new and simple form of the Euler-Lagrange equations.

AMS Subject Classification: 49S05, 49J35, 35Q40, 35J60, 35P30, 81Q05, 81V70, 81V45, 81V55.

Keywords: variational methods, critical points, minimax principles, Palais-Smale condition, Morse index, quantum mechanics, quantum chemistry, ground state, excited state, multiconfiguration methods, configuration-interaction method, Hartree-Fock equations.

1 Introduction

We want to prove here some results concerning the multiconfiguration (MC) methods in molecular quantum chemistry.

Under the *Born-Oppenheimer* approximation, the non-relativistic quantum energy of N electrons interacting with M static nuclei is given by

$$\begin{aligned}\mathcal{E}_N(\Psi) = \langle \Psi, \mathcal{H}\Psi\rangle &= \sum_{i=1}^{N} \int_{\mathbb{R}^{3N}} \left(\frac{1}{2}|\nabla_{x_i}\Psi(x)|^2 + V(x_i)|\Psi(x)|^2\right) dx \\ &\quad + \sum_{1\leqslant i<j\leqslant N} \int_{\mathbb{R}^{3N}} \frac{|\Psi(x)|^2}{|x_i - x_j|}\, dx \end{aligned} \tag{I.1}$$

where $x = (x_1, ..., x_N) \in (\mathbb{R}^3)^N$, and V is the purely Coulombic potential

$$V(u) = -\sum_{m=1}^{M} \frac{z_m}{|u - \bar{x}_m|},$$

$z_m > 0$ and $\bar{x}_m \in \mathbb{R}^3$ being the charge and position of each of the M nuclei.

In the formula (I.1), $\Psi \in H^1_a(\mathbb{R}^{3N}, \mathbb{R})$ is such that $\|\Psi\|_{L^2} = 1$. The subscript a on the Sobolev space H^1_a indicates that we consider functions Ψ which are antisymmetric under interchanges of variables (expression of the Pauli exclusion principle):

$$\forall \sigma \in S_N,\ \Psi(x_1 ... x_N) = \varepsilon(\sigma)\Psi(x_{\sigma(1)} ... x_{\sigma(N)}) \text{ a.e.}$$

$\mathcal{H}$ is the purely Coulombic N-body Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{N} \left(-\frac{1}{2}\Delta_{x_i} + V(x_i) \right) + \sum_{1 \leqslant i < j \leqslant N} \frac{1}{|x_i - x_j|}.$$

For the sake of simplicity, we restrict ourselves to real-valued functions and do not take the spin into account. Of course, everything in this article can be trivially adapted to the case of complex-valued (spin-dependent) functions. In what follows, we denote by $Z = \sum_{m=1}^{M} z_m$ the total nuclear charge. We refer the reader to [C$^+$03, Lie90] for a description of this model and a detailed explanation of the Born-Oppenheimer approximation.

The form of the spectrum of $\mathcal{H}$ is known (see for instance the review [HS00]): it is bounded from below, $\sigma_{ess}(\mathcal{H}) = [\Sigma; +\infty)$ where $\Sigma \leqslant 0$, and each potential eigenvalue lies in $(-\infty; 0]$. The *N-body ground state energy* is the minimum of $\sigma(\mathcal{H})$ also defined by

$$E_N = \inf\{\mathcal{E}_N(\Psi),\ \Psi \in H^1_a(\mathbb{R}^{3N}),\ \|\Psi\|_{L^2(\mathbb{R}^{3N})} = 1\}. \tag{I.2}$$

The condition $E_N < \Sigma$ translates the physical property of the nuclei being able to bind the N electrons in their vicinity. In agreement with intuition, this is false when Z is not sufficiently large. Zhislin showed in [Zhi60] (see also [Fri03a]) that when $Z > N - 1$,

$$\sigma(\mathcal{H}) = \{E_N = \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n \leqslant \cdots\} \cup [\Sigma; +\infty)$$

where $(\lambda_i)_{i \geqslant 1}$ are eigenvalues strictly below Σ and whose eigenfunctions are called *excited states*. Throughout this article, we shall always assume that $Z > N - 1$.

From the practical point of view, the determination of a minimum for (I.2) is the main goal we need to achieve in order to understand the molecular structure, but the computation of excited states is also important: they occur when the molecule gets an excess of energy, that could then be lost in a number of ways. Chemical reactions also often ensue after an initial electronic transition. Since the model contains no empirical parameter (this is an *"ab initio"* theory), it might be thought that very accurate results could be obtained.

Unfortunately, the numerical computation of a minimum (and *a fortiori* excited states) is extremely difficult, due to the excessive dimension of the space $\mathbb{R}^{3N}$ on which the wavefunctions are defined. This is why chemists have introduced some simplification of (I.2). One of them is based on a reduction of the space $H = H^1_a(\mathbb{R}^{3N})$ where the minimization has to be done. The idea is to choose a sequence of subsets $M^1 \subset \cdots \subset M^K \subset \cdots \subset H$ such that, intuitively, $M^K \to H$, and then to replace H by M^K in (I.2). The main advantage of this approach is that the approximate ground state energy is always greater than the true. Nevertheless, the main difficulty is that the Schrödinger linear problem (I.2) often becomes strongly nonlinear (when the M^i are not linear subspaces of H) and

so the methods used to study it are inefficient. The multiconfiguration methods belong to this type of approximation.

In the *multiconfiguration (MC) method of rank K* (see [LB94, Fri03a] for a mathematical set-up and for instance [She87] for chemical and numerical aspects), the set of admissible wavefunctions is limited to the Ψ which are a linear combination of Slater determinants built with K functions $\varphi_1, ..., \varphi_K$ in $H^1(\mathbb{R}^3)$ with $\int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}$

$$\Psi = \sum_{I=\{i_1<i_2<...<i_N\}\subset\{1...K\}} c_I \cdot |\varphi_{i_1}...\varphi_{i_N}\rangle \tag{I.3}$$

where

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

and

$$\sum_I {c_I}^2 = 1.$$

When there is no possible confusion, we shall abbreviate $\Phi_I = |\varphi_{i_1}...\varphi_{i_N}\rangle$ for $I = \{i_1, ..., i_N\}$ and write

$$\Psi = \sum_I c_I \Phi_I.$$

The set of the functions Ψ that can be written on the form (I.3) is characterized in terms of the rank of the *one-body density operator* Γ_Ψ: this is the operator acting on $L^2(\mathbb{R}^3)$ with kernel

$$\gamma_\Psi(x,y) = N \int_{\mathbb{R}^{3(N-1)}} \Psi(x, x_2...x_N)\Psi(y, x_2...x_N) dx_2...dx_N,$$

that is to say

$$\Gamma_\Psi(\varphi)(x) = \int_{\mathbb{R}^3} \gamma_\Psi(x,y)\varphi(y)\, dy.$$

It can easily be proved (see [Löw55, Fri03a] and Lemma I.1) that Ψ is of the form (I.3) if and only if $rank(\Gamma_\Psi) \leqslant K$ and $Range(\Gamma_\Psi) \subset span(\varphi_i)$. The *total electronic charge density* is defined on $\mathbb{R}^3$ by $\rho_\Psi(x) = \gamma_\Psi(x,x)$ and we have $Tr(\Gamma_\Psi) = \int_{\mathbb{R}^3} \rho(x)\, dx = N$.

If we introduce $rank(\Psi) := rank(\Gamma_\Psi)$, the *$N$-body ground state energy of rank K* is thus defined by

$$E_N^K = \inf\{\mathcal{E}_N(\Psi),\ \Psi \in H^1_a(\mathbb{R}^{3N}),\ \|\Psi\|_{L^2(\mathbb{R}^{3N})} = 1,\ rank(\Psi) \leqslant K\}. \tag{I.4}$$

Since every wave function Ψ can be expressed as a (infinite) linear combination of Slater determinants (see Lemma I.1), we have

$$\lim_{K\to+\infty} E_N^K = E_N.$$

When $K = N$, we obtain the well-known *Hartree-Fock* (HF) model [Lie90, LS77, Lio87, Gho93]: Ψ is a unique Slater determinant and Γ_Ψ a projector of rank N. The existence of a minimum for E_N^N has been proved for the first time by E.H. Lieb and B. Simon in [LS77]. In [Lio87], P.-L. Lions proved the existence of infinitely many solutions of the Hartree-Fock equations, often interpreted as excited states in the literature (this point is discussed further on). The difference $E_N^N - E_N \geqslant 0$

between the Hartree-Fock and the "exact" energy is called the *electron correlation energy* [LB94, LB99, LB93, Löw59]. It can change quite significantly in many elementary chemical processes: typical examples are transition states in chemical reactions, bond breaking and excited states (see for instance [AF97, HG96, WM80]). In such situations, the HF method fails and chemists have to use the more accurate MC.

In practice, chemists rarely work with an expansion of wave functions containing all Slater determinants built with $\varphi_1, ..., \varphi_K$ like in formula (I.3). They often select some determinants (*configurations*), chosen for instance thanks to the known symmetry of the molecule or some chemical intuition, and optimize the energy with this form fixed [Roo87, She87, Wer87, WM80]. In the sequel, we shall call this approach *"partial multiconfiguration"*. From the mathematical point of view, these methods are more difficult to deal with since the form may change when another basis of $Range(\Gamma_\Psi)$ is used.

For algebraic reasons (see [And63, Col63, Löw55, Fri03a] and Appendix), there do not exist N-body wave functions of rank $K = N + 1$ and when $N = 2$, all the Ψ have an even rank. A partial multiconfiguration method of rank $K = N + 2$ was studied by C. Le Bris [LB94] who considered the minimization over doubly excited configurations

$$\Psi = \alpha|\varphi_1...\varphi_N\rangle + \beta|\varphi_1...\varphi_{N-2}\varphi_{N+1}\varphi_{N+2}\rangle. \tag{I.5}$$

He proved the existence of a minimum when $Z > N - 1$ and obtained the inequality $E_N^{N+2} < E_N^N$ (see [Fri03b] for the generalization $E_N^{K+2} < E_N^K$, $K \geqslant N$).

The existence of a minimum for E_N^K provided $K \geqslant N$ and $Z > N - 1$ was recently proved by G. Friesecke in [Fri03a]. He used a very interesting proof based on geometric localization methods, P.-L. Lions' concentration-compactness ideas and the inequality $E_{N+1}^{K+1} < E_N^K$. The latter inequality allowed the author to notice an interesting connection between Lions' method and the celebrated HVZ Theorem. Unfortunately, we think that his method cannot easily be generalized to obtain saddle points. We do not know if it can be adapted to treat the case of partial MC.

To our knowledge, the problem of existence of excited states for multiconfiguration methods and their relation with the eigenvalues of $\mathcal{H}$ have never been tackled from the mathematical point of view, although this is one of the main uses of MC in today's molecular quantum chemistry. Our approach, based on ideas of [Lio87], allows us to prove the existence of a minimum and of saddle points for some partial MC, including the "full" MC studied by G. Friesecke, the doubly excited configurations of C. Le Bris, and the other main methods used by chemists. More precisely, for the "full" method,

1. For each $K \geqslant N$, we prove that there exists a sequence of distinct critical points of rank K, whose energy converges to 0;

2. For each $K \geqslant N$, we construct explicitly $\binom{K}{N}$ critical points of rank K with a nonlinear min-max method that is related to the algorithm used by chemists. Their energy converges when $K \to +\infty$ to the "true" critical values of $\mathcal{E}_N$ (the eigenvalues $(\lambda_i)_{i\geqslant 1}$ of the Hamiltonian $\mathcal{H}$).

The solutions of the multiconfiguration equations constructed in the first part are the analogue of P.-L. Lions' result for HF [Lio87]. Our impression is that they cannot really be interpreted as excited states of the molecule (for instance, their energy tends to 0 while $\lambda_i \to \Sigma$ as $i \to +\infty$). The solutions constructed in the second part are better candidates

in view of their behaviour when $K \to +\infty$. For each $K \geqslant N$, there is a finite number of such points and only one type (the minima) for the HF case $K = N$. This corresponds to chemists' experience, which suggests that many determinants have to be included in the model to obtain accuracy for a fixed excited state of the molecule.

As for HF in [LS77, Lio87], we study the energy in the "one-body space" (as a function of c_I and φ_i), in opposition to G. Friesecke's approach in the "N-body space" $H^1_a(\mathbb{R}^{3N})$. Our method is based on a Palais-Smale condition with Morse-type information, and on the Euler-Lagrange equations which were often neglected in the previous studies because of their apparent complexity. We express them in a very simple and useful way (I.9) that clarifies the connection with the HF model and could also become relevant for numeric simulations.

Finally, let us mention that some of the results proved here were outlined in [Lew02].

In the next section, we introduce some notation and definitions in order to state properly the main results of this paper; they can be found in section 3, with some comments on the proofs and on the relation with the algorithms. Section 4 is devoted to the proof of our results.

Acknowledgement: I would like to thank Éric Séré for his constant attention and helpful advice. I am also indebted to Gero Friesecke and the referee for valuable remarks.

2 General setting of the model

2.1 Expression of the energy in one-body space

We need some notation. Let Ψ be a wave function of rank K

$$\Psi = \sum_{I=\{i_1<i_2<\ldots<i_N\}\subset\{1\ldots K\}} c_I \cdot |\varphi_{i_1}\ldots\varphi_{i_N}\rangle.$$

If we use the convention

$$\begin{aligned} \alpha_{i_1\ldots i_N} &= 0 \quad \text{if } \#\{i_1\ldots i_N\} < N \\ &= \frac{\varepsilon(\sigma)}{\sqrt{N!}} c_{\{i_{\sigma(1)}<\ldots<i_{\sigma(N)}\}} \text{ otherwise,} \end{aligned}$$

(σ is the permutation of $[1;N]$ such that $i_{\sigma(1)} < \ldots < i_{\sigma(N)}$), we get

$$\Psi = \sum_{i_1\ldots i_N} \alpha_{i_1,\ldots,i_N} \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N}.$$

We are now able to express the energy as a function of the c_I and φ_i:

$$\begin{aligned} \mathcal{E}_N(\Psi) = N \sum_{1\leqslant k_2,\ldots,k_N\leqslant K} \int_{\mathbb{R}^3} \frac{1}{2} \left| \sum_{i=1}^{K} \alpha_{i,k_2\ldots k_N} \nabla\varphi_i \right|^2 + V \left(\sum_{i=1}^{K} \alpha_{i,k_2\ldots k_N} \varphi_i \right)^2 \\ + \frac{N(N-1)}{2} \sum_{1\leqslant k_3,\ldots,k_N\leqslant K} \iint_{\mathbb{R}^6} \frac{\left(\sum_{1\leqslant i,j\leqslant K} \alpha_{i,j,k_3\ldots,k_N} \varphi_i(x)\varphi_j(y) \right)^2}{|x-y|} dxdy. \quad \text{(I.6)} \end{aligned}$$

This can be written on the compact form

$$\mathcal{E}_N(\Psi) = \left\langle \left(\left(-\frac{\Delta}{2} + V \right) \Gamma + W_\Phi \right) \cdot \Phi \, , \, \Phi \right\rangle_{(L^2(\mathbb{R}^3))^K} \tag{I.7}$$

where $\Phi = (\varphi_1, ..., \varphi_K)^T \in (H^1(\mathbb{R}^3))^K$ and the $K \times K$ matrices Γ and W_Φ are defined by

$$\Gamma_{i,j} = N \sum_{k_2 ... k_N} \alpha_{i,k_2,...k_N} \alpha_{j,k_2,...k_N}$$

$$(W_\Phi)_{i,j}(x) = \frac{N(N-1)}{2} \sum_{k_3 ... k_N} \sum_{k,l} \alpha_{i,k,k_3...k_N} \alpha_{j,l,k_3...k_N} \left((\varphi_k \varphi_l) * \frac{1}{|r|} \right)(x).$$

In (I.7), we have used the notation $\langle \Phi, \Phi' \rangle_{(L^2(\mathbb{R}^3))^K} = \sum_{i=1}^K \int_{\mathbb{R}^3} \varphi_i \varphi_i'$.

With $c = (c_I) \in \mathbb{R}^{\binom{K}{N}}$ (lexicographical order), we introduce

$$\mathcal{M}_N^K = \left\{ (c, \Phi) \in \mathbb{R}^{\binom{K}{N}} \times (H^1(\mathbb{R}^3))^K, \ \sum_I {c_I}^2 = 1, \ \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij} \right\}$$

and $\mathcal{E}_N^K$, which is defined on $\mathcal{M}_N^K$ by the formula

$$\mathcal{E}_N^K(c, \Phi) = \left\langle \left(\left(-\frac{\Delta}{2} + V \right) \Gamma + W_\Phi \right) \cdot \Phi \, , \, \Phi \right\rangle_{(L^2(\mathbb{R}^3))^K}. \tag{I.8}$$

We shall denote by $d\mathcal{E}_N^K$ its first derivative, and $d_\Phi^2 \mathcal{E}_N^K$ its second derivative with respect to Φ, on the Riemannian manifold $\mathcal{M}_N^K$. Let us notice that we now have

$$E_N^K = \inf_{\mathcal{M}_N^K} \mathcal{E}_N^K.$$

The problem is to prove the existence of a minimum and saddle points for $\mathcal{E}_N^K$ on $\mathcal{M}_N^K$. Such a point (c, Φ) will be a solution of the *multiconfiguration equations* which take the form

$$\left(\left(-\frac{\Delta}{2} + V \right) \Gamma + 2W_\Phi \right) \cdot \Phi + \Lambda \cdot \Phi = 0 \tag{I.9}$$

$$H_\Phi \cdot c = \beta \cdot c. \tag{I.10}$$

The first equation (I.9) is a system of K nonlinear partial differential equations – the Euler-Lagrange equations for the φ_i, Λ being the Lagrange multipliers matrix associated with the constraints $\int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}$. The second equation (I.10) is an eigenvalue problem – the Euler-Lagrange equations for the c_I, with multiplier $\beta = \mathcal{E}_N^K(c, \Phi)$. $H_\Phi = (\langle \Phi_I, \mathcal{H} \Phi_J \rangle)_{I,J}$ is the $\binom{K}{N} \times \binom{K}{N}$ matrix of the Hamiltonian $\mathcal{H}$ on $span(\Phi_I)$.

We shall see that the compact form of (I.9) will largely simplify the study. A formalism to treat the case of partial MC will be introduced later.

We shall compare some critical values of $\mathcal{E}_N^K$ on $\mathcal{M}_N^K$ to the eigenvalues of the Hamiltonian $\mathcal{H}$:

$$\lambda_d = \min_{dim(V)=d} \max_{\substack{\Psi \in V, \\ \|\Psi\|_{L^2} = 1}} \langle \mathcal{H}\Psi, \Psi \rangle. \tag{I.11}$$

Recall that since $Z > N - 1$, the $(\lambda_d)_{d \geqslant 1}$ are eigenvalues of finite multiplicity of $\mathcal{H}$ below the bottom of the essential spectrum Σ.

2.2 About the orthogonal invariance

Let us now give a precise statement for some facts mentioned in the Introduction, concerning the role of Slater determinants in $H^1_a(\mathbb{R}^{3N})$ and the basic properties of rank K wave functions. We recall that Γ_Ψ is the one-body density operator defined previously.

Lemma I.1 (Basic properties of rank K wave functions [Löw55]).

1. *Let $(\varphi_i)_{i\in\mathbb{N}}$ be an orthonormal basis of $L^2(\mathbb{R}^3)$ (resp. $H^1(\mathbb{R}^3)$). Then $(|\varphi_{i_1}...\varphi_{i_N}\rangle)_{i_1<...<i_N}$ is an orthonormal basis of $L^2_a(\mathbb{R}^{3N})$ (resp. $H^1_a(\mathbb{R}^{3N})$).*
2. *Let Ψ be a wave function of rank K (i.e. $rank(\Gamma_\Psi) = K$), and $(\varphi_i)_{i=1}^K$ be an orthonormal basis of $Range(\Gamma_\Psi)$. Then Ψ can be expanded as a linear combination of Slater determinants built with $(\varphi_i)_{i=1}^K$. Conversely, every wave function which is a linear combination of Slater determinants is of finite rank.*

Assertion *2.* of this lemma corresponds to the *Löwdin Expansion Theorem* [Löw55], while assertion *1.* is just the expression of the classical

$$L^2_a(\mathbb{R}^{3N}) \simeq \bigwedge_{i=1}^N L^2(\mathbb{R}^3).$$

When $(\varphi_i)_{i=1}^K$ is an orthonormal basis of $Range(\Gamma_\Psi)$, each φ_i is, by definition, an *orbital of* Ψ. If moreover they are eigenfunctions of Γ_Ψ, they are called the *natural orbitals of* Ψ. The associated eigenvalues are then the *occupation numbers*. By Löwdin's Theorem, each Ψ of finite rank can be expanded as a linear combination of Slater determinants involving its natural orbitals.

If we now take a wave function of rank K

$$\Psi = \sum_I c_I \Phi_I = \sum_{k_1,...,k_N=1}^K \alpha_{k_1,...k_N}\ \varphi_{k_1} \otimes \cdots \otimes \varphi_{k_N},$$

let us see how $c = (c_I)$ changes when another basis (φ'_i) of $Range(\Gamma_\Psi)$ is used. Assuming that $\varphi'_i = \sum_{j=1}^K u_{i,j}\varphi_j$ with $U = (u_{i,j}) \in \mathcal{O}_K(\mathbb{R})$ (the set of $K \times K$ orthogonal matrices), we obtain

$$\begin{aligned}
\alpha'_{i_1,...i_N} &= \langle \Psi, \varphi'_{i_1} \otimes \cdots \otimes \varphi'_{i_N} \rangle_{L^2(\mathbb{R}^{3N})} \\
&= \sum_{j_1,...,j_N=1}^K u_{i_1,j_1} \cdots u_{i_N,j_N} \cdot \langle \Psi, \varphi_{j_1} \otimes \cdots \otimes \varphi_{j_N} \rangle_{L^2(\mathbb{R}^{3N})} \\
&= \sum_{j_1,...,j_N=1}^K u_{i_1,j_1} \cdots u_{i_N,j_N} \cdot \alpha_{j_1...j_N}. \qquad \text{(I.12)}
\end{aligned}$$

This defines a group action of $\mathcal{O}_K(\mathbb{R})$ on $\mathcal{M}_N^K$:

$$U \cdot (c, \Phi) = (c', U \cdot \Phi).$$

We have the following

Lemma I.2. *We assume that $(c', \Phi') = U \cdot (c, \Phi)$. With obvious notations, we then have :*

$$\Gamma' = U\Gamma U^T \quad \text{and} \quad W'_{\Phi'} = UW_\Phi U^T.$$

In particular $\mathcal{E}_N^K$ is invariant under the action of $\mathcal{O}_K(\mathbb{R})$ on $\mathcal{M}_N^K$.

Proof – It is easy to see that Γ is the matrix of Γ_Ψ in the basis $(\varphi_i)_{i=1}^K$. Its transformation law is thus obvious. For W_Φ, the computation is very easy but tedious. □

It is easy to see that the action of $\mathcal{O}_K(\mathbb{R})$ on $\mathcal{M}_N^K$ is not free when $K > N$. Therefore we believe that the quotient $\mathcal{M}_N^K/\mathcal{O}_K(\mathbb{R})$ is not smooth (that is to say, it is not a manifold) and that is why we shall not really use it here. Nevertheless, this action of $\mathcal{O}_K(\mathbb{R})$ will play an important role in the proof of Theorem I.1 since we shall use it to diagonalize matrices.

2.3 Formalism for partial multiconfiguration methods

We introduce

$$\mathcal{A}_N^K = \{I \subset \{1, ..., K\} : |I| = N\}.$$

In a partial MC method of rank K, a set $\mathcal{I} \subset \mathcal{A}_N^K$ is chosen and the energy $\mathcal{E}_N^K$ is restricted to the submanifold

$$\mathcal{M}_N^{K,\mathcal{I}} = \{(c, \Phi) \in \mathcal{M}_N^K : \forall I \notin \mathcal{I},\ c_I = 0\} \subset \mathcal{M}_N^K.$$

The ground state energy associated with the method is then defined by

$$E_N^{K,\mathcal{I}} = \inf_{\mathcal{M}_N^{K,\mathcal{I}}} \mathcal{E}_N^K.$$

Critical points of $\mathcal{E}_N^K$ on $\mathcal{M}_N^{K,\mathcal{I}}$ are now solutions of the following *partial multiconfiguration equations*

$$\left(\left(-\frac{\Delta}{2} + V\right)\Gamma + 2W_\Phi\right) \cdot \Phi + \Lambda \cdot \Phi = 0 \tag{I.13}$$

$$H_\Phi^{\mathcal{I}} \cdot c_{\mathcal{I}} = \beta \cdot c_{\mathcal{I}}. \tag{I.14}$$

where $c_{\mathcal{I}} = (c_I)_{I \in \mathcal{I}}$ and $H_\Phi^{\mathcal{I}} = (\langle \Phi_I, \mathcal{H}\Phi_J \rangle)_{I,J \in \mathcal{I}}$ is the $|\mathcal{I}| \times |\mathcal{I}|$ matrix of the Hamiltonian $\mathcal{H}$ on $span(\Phi_I, I \in \mathcal{I})$.

Compared to the "full" MC, the new difficulty is that one cannot choose any basis $(\varphi_i)_{i=1}^K$ of $Range(\Gamma_\Psi)$ for we need to preserve the form of the wavefunction. We introduce

$$\mathcal{O}_{\mathcal{I}} = \left\{U \in \mathcal{O}_K(\mathbb{R}) : U \cdot \mathcal{M}_N^{K,\mathcal{I}} \subset \mathcal{M}_N^{K,\mathcal{I}}\right\}.$$

Below, we shall restrict ourselves to the following methods:

Definition I.1. *We say that $\mathcal{I} \subset \mathcal{A}_N^K$ defines a* natural multiconfiguration method of rank K *if*

1. *there exists $(c, \Phi) \in \mathcal{M}_N^{K,\mathcal{I}}$ which is of rank K*
2. *for all $(c, \Phi) \in \mathcal{M}_N^{K,\mathcal{I}}$ there exists a $U \in \mathcal{O}_{\mathcal{I}}$ such that Γ', associated with $(c', \Phi') = U \cdot (c, \Phi)$, is a diagonal matrix. In other words, every wavefunction of the method can be expressed in terms of its natural orbitals.*

Let us notice that most of the methods used by chemists are "natural methods" (the natural orbitals have an important physical meaning). Here are some examples:

• (M_1) **Full MC.** It is clear that the general multiconfiguration method explained before is a natural method with $\mathcal{I} = \mathcal{A}_N^K$.

• (M_2) **Two nonvanishing determinants have at least two distinct orbitals.** Suppose that $\mathcal{I}$ satisfies

$$\#(I \cap J) = 1 \Longrightarrow I \notin \mathcal{I} \text{ or } J \notin \mathcal{I}.$$

Then it can easily be seen that Γ_Ψ is always diagonal and this method is thus natural. As an example, we can quote the method studied by C. Le Bris in [LB94]

$$\Psi = \alpha|\varphi_1...\varphi_N\rangle + \beta|\varphi_1...\varphi_{N-2}\varphi_{N+1}\varphi_{N+2}\rangle.$$

See the Appendix for a study of the role of those methods in the cases $N = 2$ and $K = N + 2$.

• (M_3) **Complete Active Space method.** The most common method used by chemists is probably the Complete Active Space method [Roo87, She87]. In this method, the electrons are divided into two groups : the *inactive* (or *core*) electrons that are supposed to be correctly described by a Hartree-Fock type model, and the *active* (or *valence*) electrons that are supposed to be the most important for estimating the correlation energy. They are thus described by a "Full" MC. Hence, an integer $N_v \in [0; N]$ is chosen and we take

$$\mathcal{I}_{CAS} = \{I : \{1, ..., N - N_v\} \subset I\}.$$

In other words, the considered wavefunctions have the form

$$\begin{aligned}\Psi &= |\varphi_1...\varphi_{N-N_v}\rangle \wedge \left(\sum_{\substack{I = \{i_1 < ... < i_{N_v}\} \subset \{N - N_v + 1, ..., K\}, \\ |I| = N_v}} c_I |\varphi_{i_1}, ...\varphi_{i_{N_v}}\rangle \right) \\ &:= \sum_{\substack{I \subset \{N - N_v + 1, ..., K\}, \\ |I| = N_v}} c_I |\varphi_1, ..., \varphi_{N-N_v}, \varphi_{i_1}, ..., \varphi_{i_{N_v}}\rangle\end{aligned}$$

$N_v = N$ corresponds to the "full" MC method of rank K, while $N_v = 0$ is the Hartree-Fock model. Typically, $N_v \in [2; 10]$ in the common chemical programs.

It is easy to see that Γ has the form

$$\Gamma = \begin{pmatrix} I_{N-N_v} & 0 \\ 0 & \Gamma_v \end{pmatrix}$$

so with an appropriate

$$U = \begin{pmatrix} I_{N-N_v} & 0 \\ 0 & U_v \end{pmatrix}$$

that does not change the form of the wavefunction, we can diagonalize Γ. This is thus a natural MC method.

3 Main results and comments

3.1 Our results

We can now state our main results.

The first is a rather technical result which is simply the compactness of Palais-Smale sequences, but with the additional difficulty that the sequence may lose rank.

Theorem I.1 (Palais-Smale condition with Morse-type information). *Assume $Z > N-1$, $K \geqslant N$ and suppose that $\mathcal{I}$ defines a natural MC method of rank K. Let $(c^n, \Phi^n)_{n\in\mathbb{N}}$ be a sequence in $\mathcal{M}_N^{K,\mathcal{I}}$, with associated wavefunctions $(\Psi^n)_{n\in\mathbb{N}}$, such that*

1. *$(\mathcal{E}_N^K(c^n,\Phi^n))_{n\in\mathbb{N}}$ is bounded*
2. *$d_\Phi \mathcal{E}_N^K(c^n,\Phi^n) \to 0$ in $(L^2(\mathbb{R}^3))^K$, and $\frac{\partial \mathcal{E}_N^K}{\partial c_I}(c^n,\Phi^n) \to 0$ for all $I \in \mathcal{I}$.*
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, $d_\Phi^2(\mathcal{E}_N^K)(c^n,\Phi^n)$ has at most j eigenvalues below $-\delta^n$.*

Then, there exists a $K' \in [N;K]$ such that, up to subsequences and after a rotation $U_n \in \mathcal{O}_K(\mathbb{R})$ on each (c^n,Φ^n):

1. *$\forall n \in \mathbb{N}$, $\Gamma^n = diag(\gamma_1^n, ..., \gamma_K^n)$.*
2. *$\forall i = 1, ..., K'$, $\gamma_i^n \to \gamma_i > 0$ and $\varphi_i^n \to \varphi_i$ strongly in $H^1(\mathbb{R}^3)$.*
3. *$\forall i = K'+1, ..., K$, $\gamma_i^n \to 0$ and $\sqrt{\gamma_i^n}\varphi_i^n \to 0$ strongly in $H^1(\mathbb{R}^3)$.*
4. *We define $\mathcal{I}' = \mathcal{I} \cap \mathcal{A}_N^{K'}$, $c' = (c_I)_{I\in\mathcal{A}_N^{K'}}$ and $\Phi' = (\varphi_1, ..., \varphi_{K'})$. Then (c', Φ') is a critical point of $\mathcal{E}_N^{K'}$ on $\mathcal{M}_N^{K',\mathcal{I}'}$ and is a solution of the partial multiconfiguration equations (I.13,I.14) with a $\Lambda > 0$.*
5. *The sequence $\Psi^n \to \Psi$ strongly in $H_a^1(\mathbb{R}^{3N})$, where Ψ is the wavefunction associated to (c', Φ').*
6. *If $\mathcal{I} = \mathcal{A}_N^K$ (Full MC), then $K' \in \{K-1, K\}$.*

Remark I.1. *The last claim $K' \in \{K-1, K\}$ is also true for some partial MC methods, including example (M_3) quoted in section 2.3. See the proof in [Fri03b, LB94]. In general, this depends on the method.*

From the above result we obtain the existence of a minimizer:

Corollary I.2 (Ground State for Multiconfiguration Methods). *Let $Z > N-1$. For all $K \geqslant N$ and any $\mathcal{I}$ that defines a natural method of rank K, there exists a minimizer of $\mathcal{E}_N^K$ on $\mathcal{M}_N^{K,\mathcal{I}}$.*

We now state our result concerning excited states:

Theorem I.3 (Excited States for Multiconfiguration Methods). *Assume $Z > N-1$.*

(i) For all $K \geqslant N$ and any $\mathcal{I}$ that defines a natural method of rank K, there exists a sequence $(\tilde{c}^i, \tilde{\Phi}^i)_{i\geqslant 1}$ of distinct points of $\mathcal{M}_N^{K,\mathcal{I}}$, which are solutions of equation (I.9) with

some $\tilde{\Lambda}_i \geqslant 0$. *In addition, if* $(\tilde{\Psi}^i)_{i\geqslant 1}$ *is the associated sequence of wavefunctions, as* i *goes to infinity,*

$$\mathcal{E}_N(\tilde{\Psi}^i) < 0, \quad \mathcal{E}_N(\tilde{\Psi}^i) \to 0, \quad \tilde{\Lambda}_i \to 0,$$

$$\sqrt{\tilde{\Gamma}^i} \cdot (\nabla \tilde{\Phi}^i) \to 0 \quad \textit{in } (L^2(\mathbb{R}^3))^K, \quad \textit{and} \quad \nabla \tilde{\Psi}^i \to 0 \quad \textit{in } L^2_a(\mathbb{R}^{3N}).$$

(ii) For all $K \geqslant N$ *and any* $\mathcal{I}$ *that defines a natural method of rank* K*, there exists* $|\mathcal{I}|$ *points* $(c_d^K, \Phi_d^K)_{1\leqslant d\leqslant |\mathcal{I}|}$ *of* $\mathcal{M}_N^{K,\mathcal{I}}$*, which are solutions of equation (I.9) with some* $\Lambda_d^K \geqslant 0$*, and such that, if* $(\Psi_d^K)_{1\leqslant d\leqslant |\mathcal{I}|}$ *is the associated sequence of wave functions, we have for all* $d \geqslant 1$

$$\mathcal{E}(\Psi_d^K) \geqslant \lambda_d.$$

If moreover $\mathcal{I} = \mathcal{A}_N^K$ *(Full MC), then we have (with* d *fixed)*

$$\mathcal{E}(\Psi_d^K) \to \lambda_d \textit{ as } K \to +\infty.$$

We recall that λ_d is the d^{th} eigenvalue of $\mathcal{H}$ defined in (I.11). The two parts of Theorem I.3 may be interpreted as follows. For a fixed method (i.e. fixed K and $\mathcal{I}$), there are many solutions to the multiconfiguration equations: in the first part (i), we show how an infinity can be exhibited. But, just a finite number of them, defined in part (ii), is related to the true eigenfunctions of the Hamiltonian in the N-body space. Hence, to obtain accuracy for the d^{th} excited state of the molecule, we need to work with a lot of Slater determinants so that $d \leqslant |\mathcal{I}|$.

Remark I.2. *When* $\mathcal{I} = \mathcal{A}_N^K$*, we believe that the sequence* $(\Psi_d^K)_{K\geqslant N}$ *in Theorem I.3-ii) is precompact in* $H^1_a(\mathbb{R}^{3N})$ *and converges, up to a subsequence, to an eigenfunction* Ψ_d *of* $\mathcal{H}$ *associated to the eigenvalue* λ_d*, but we were unable to prove it. This is true in the case of minimizing sequences (*$d = 1$*), see [Fri03a]. We hope to come back to this point in the future.*

Finally, let us specify the properties of the orbitals, when they are solutions of the multiconfiguration equations.

Proposition I.3. *Let* $(c, \Phi) \in \mathcal{M}_N^K$ *be such that* $\Gamma > 0$ *and* Φ *is a solution of*

$$\left(\left(-\frac{\Delta}{2} + V\right)\Gamma + 2W_\Phi\right) \cdot \Phi + \Lambda \cdot \Phi = 0. \tag{I.15}$$

Then each φ_i *is in* $\bigcap_{2\leqslant p<3} W^{2,p}(\mathbb{R}^3)$ *and is real-analytic on* $\Omega := \mathbb{R}^3 \setminus \{\bar{x}_1 ... \bar{x}_M\}$*. If moreover* $\Lambda > 0$*, then it also has an exponential fall-off.*

Hence, the wavefunctions constructed by our method are real-analytic on Ω^N and have an exponential fall-off. This regularity of the multiconfiguration optimal wavefunctions could be the main reason for the rather slow convergence of these methods when $K \to +\infty$, since the true eigenfunctions (Ψ_i) of $\mathcal{H}$ are known to have a cusp at electrons coalescence points, i.e. $(x_1, ..., x_N) \in \Omega^N$ with $\prod_{i\neq j}(x_i - x_j) = 0$ [Kat57, HOS94].

3.2 Strategy of proof

In this subsection, we make a few comments on the proof of the above results. For the sake of simplicity, we suppose here $\mathcal{I} = \mathcal{A}_N^K$ (full method).

In [Lio87], P.L. Lions uses second order information on the almost critical sequences in order to prove the convergence of minimizing or min-maxing sequences in the Hartree-Fock case. Such sequences can be obtained by using for instance the Borwein-Preiss minimization principle (which is an improvement of Ekeland's variational principle since the perturbations are quadratic) in the case of a minimum, and the later work of Fang & Ghoussoub [FG92, Gho93] that deals with suitable min-max. Theorem I.1 is thus the first step towards the existence of minima (Corollary I.2) and saddle points (Theorem I.3).

Therefore, our proof of Theorem I.1 is largely inspired of the study of the Hartree-Fock functional in [Lio87]. There are two main new difficulties.

The first is that we have lost the "crucial orthogonal invariance of the energy" (see [LB94]). It is linked to the fact that, in the Hartree-Fock theory, $\Gamma = I_N$ and so it may be supposed that Γ and Λ are both diagonal, which simplifies the arguments since the problem splits into N less coupled scalar equations. For the general MC theory, we have seen (Lemma I.2) that if $\Phi' = U \cdot \Phi$ where $U \in \mathcal{O}_K(\mathbb{R})$, then $\Gamma' = U\Gamma U^T$, $W'_{\Phi'} = UW_\Phi U^T$. This implies $\Lambda' = U\Lambda U^T$ in (I.9). Hence, we may suppose either Γ or Λ diagonal, according to what we need, but we cannot assume that they are simultaneous diagonal.

The second difficulty is that the occupation numbers (the eigenvalues $(\gamma_i^n)_{i=1}^K$ of Γ^n), which were all equal to 1 in HF, may go to 0 in a MC method; this corresponds to a lost of rank. Since we just know that the $\sqrt{\gamma_i^n}\varphi_i^n$ are bounded in $H^1(\mathbb{R}^3)$ (see Lemma I.4), this could obstruct the convergence of the (φ_i^n) when $\gamma_i^n \to 0$.

The scheme of our proof is the following:

Step 1 : We first show *3.*, thanks to the information on the first derivative with regard to c of the Palais-Smale sequence, and define K'. This allows us to work in $\mathcal{M}_N^{K'}$ and assume $\Gamma^n \geqslant \gamma I$ where $\gamma > 0$. In this step, we diagonalize Γ^n.

Step 2 : We use the information on the first and second derivatives of the sequence to prove that the Lagrange multipliers cannot tend to 0. More precisely, we show that there exists a $\lambda > 0$ such that $\Lambda^n \geqslant \lambda I$. In this step, we diagonalize Λ^n.

Step 3 : We use the Euler-Lagrange equation and the nondegeneracy information of step 2 on the sequence of multipliers (Λ^n) to prove that the sequence is relatively compact.

Using Theorem I.1, it is then easy to prove the existence of a minimum, as it is stated in Corollary I.2

Let us now give a sketch of the proof of Theorem I.3.

As in [Lio87, Gho93], we use the fact that $\mathcal{E}_N^K$ is even in c and Φ in order to define appropriate min-max methods to obtain critical points. If a group G acts on two topological spaces X and Y, we recall that a function $\varphi : X \to Y$ is *G-equivariant* if $\varphi(g \cdot x) = g \cdot \varphi(x)$ for all $g \in G$ and $x \in X$. We denote by $\mathcal{C}_G(X, Y)$ the set of all continuous G-equivariant functions. In our case, we shall consider, as in [Lio87, Gho93, Rab86], min-maxing methods of the form:

$$\min_{f \in \mathcal{C}_G(S^{d-1}, \mathcal{M}_N^K)} \max_{(c,\Phi) \in f(S^{d-1})} \mathcal{E}_N^K(c, \Phi) \tag{I.16}$$

where $G = \mathbb{Z}_2 \simeq \{\pm 1\}$ acts obviously on the Euclidian sphere S^{d-1} of $\mathbb{R}^d$. It remains to find an appropriate action of $\mathbb{Z}_2$ on $\mathcal{M}_N^K$. Actually, we shall define two distinct actions and prove (i) and (ii) separately.

Suppose that we want to approximate the "true" eigenvalues of the Hamiltonian λ_d defined in (I.11). Let us denote by

$$\begin{array}{rccc} \Pi_N^K : & \mathcal{M}_N^K & \longrightarrow & H_a^1(\mathbb{R}^{3N}, \mathbb{R}) \\ & (c, \Phi) & \longmapsto & \Psi = \sum_I c_I \Phi_I \end{array}$$

the natural projection from the one-body space into the N-body space. One can prove that

$$\lambda_d = \min_{g\in\Theta_d} \max_{\Psi\in g(S^{d-1})} \langle \mathcal{H}\Psi, \Psi\rangle$$

where Θ_d is the collection of all odd continuous map from S^{d-1} into the sphere of $H^1_a(\mathbb{R}^{3N})$. To be able to compare our min-max with λ_d, we shall choose an action of $\mathbb{Z}_2 \simeq \{\pm 1\}$ on $\mathcal{M}^K_N$ such that $g = \Pi^K_N \circ f \in \Theta_d$ for all $\mathbb{Z}_2$-equivariant function $f : S^{d-1} \to \mathcal{M}^K_N$, that is to say $\Pi^K_N \circ f$ is odd. In standard theories, compact symmetric sets are often used with $(-)\cdot(c,\Phi) = (-c,-\Phi)$, a group action that does not have this property when N is odd.

The following group action has the requested property

$$(-)\cdot_c (c,\Phi) = (-c,\Phi).$$

Nevertheless, it allows to construct only a finite number of critical points. To see this point, let us introduce the projection

$$P : \begin{array}{ccc} \mathcal{M}^K_N & \longrightarrow & S^{\binom{K}{N}-1} \\ (c,\Phi) & \longmapsto & c \end{array}.$$

Then if $f : S^{d-1} \to \mathcal{M}^K_N$ is equivariant for this group action, $P\circ f : S^{d-1} \to S^{\binom{K}{N}-1}$ is odd. By the Borsuk-Ulam Theorem, we therefore obtain $\mathcal{C}_G(S^{d-1}, \mathcal{M}^K_N) = \varnothing$ when $d > \binom{K}{N}$ and so only $\binom{K}{N}$ min-max will be really defined by a formula like (I.16).

This point being explained, we consider the following two group actions of $\mathbb{Z}_2$ on $\mathcal{M}^K_N$:

$$(-)\cdot_c (c,\Phi) = (-c,\Phi) \tag{I.17}$$
$$(-)\cdot_\Phi (c,\Phi) = (c,-\Phi). \tag{I.18}$$

A function $f : S^{d-1} \longrightarrow \mathcal{M}^K_N$ is said to be $(\mathbb{Z}_2)_c$-equivariant when

$$f(x) = (c,\Phi) \implies f(-x) = (-c,\Phi)$$

and $(\mathbb{Z}_2)_\Phi$-equivariant when

$$f(x) = (c,\Phi) \implies f(-x) = (c,-\Phi).$$

We remark that $\mathcal{E}^K_N$ is both $(\mathbb{Z}_2)_c$ and $(\mathbb{Z}_2)_\Phi$-invariant on $\mathcal{M}^K_N$, that is to say

$$\mathcal{E}^K_N(c,\Phi) = \mathcal{E}^K_N(-c,\Phi) = \mathcal{E}^K_N(c,-\Phi).$$

We can now define:

$$\lambda^K_d = \inf_{f\in\mathcal{C}_{(\mathbb{Z}_2)_c}(S^{d-1},\mathcal{M}^K_N)} \max_{(c,\Phi)\in f(S^{d-1})} \mathcal{E}^K_N(c,\Phi) \tag{I.19}$$

for $1 \leqslant d \leqslant \binom{K}{N}$, and

$$\mu^K_d = \inf_{f\in\mathcal{C}_{(\mathbb{Z}_2)_\Phi}(S^{d-1},\mathcal{M}^K_N)} \max_{(c,\Phi)\in f(S^{d-1})} \mathcal{E}^K_N(c,\Phi) \tag{I.20}$$

for all $d \geqslant 1$.

This argument can be easily adapted to the case of partial multiconfiguration methods. Since for all $\mathcal{I} \subset \mathcal{A}_N^K$, $\mathcal{M}_N^{K,\mathcal{I}}$ is globally invariant under the two group actions, we may thus define similarly

$$\lambda_d^{K,\mathcal{I}} = \inf_{f \in \mathcal{C}_{(\mathbb{Z}_2)_c}(S^{d-1}, \mathcal{M}_N^{K,\mathcal{I}})} \max_{(c,\Phi) \in f(S^{d-1})} \mathcal{E}_N^K(c,\Phi) \tag{I.21}$$

for $1 \leqslant d \leqslant |\mathcal{I}|$, and

$$\mu_d^{K,\mathcal{I}} = \inf_{f \in \mathcal{C}_{(\mathbb{Z}_2)_\Phi}(S^{d-1}, \mathcal{M}_N^{K,\mathcal{I}})} \max_{(c,\Phi) \in f(S^{d-1})} \mathcal{E}_N^K(c,\Phi) \tag{I.22}$$

for all $d \geqslant 1$.

In section 4, we show that the variational principles underlying $(I.21)$ and $(I.22)$ yield solutions to the multiconfiguration equations, and that they allow to prove respectively (ii) and (i) of Theorem I.3.

Remark I.3. *The sequence $(\mu_d^{K,\mathcal{I}})_{d \geqslant 1}$ is the natural generalization of the one that was used in the proof for $K = N$ (HF) in [Lio87, Gho93] (see also [ES99] for the relativistic case), and its study will use exactly the same ideas.*

Remark I.4. *We suppose here $\mathcal{I} = \mathcal{A}_N^K$. A basis $\mathcal{B} = (\varphi_i)_{i \geqslant 1}$ of $H^1(\mathbb{R}^3)$ being fixed, it is also natural to introduce, for all $K \geqslant N$ and $d \geqslant 1$ such that $\binom{K}{N} \geqslant d$*

$$\lambda_{\mathcal{B},d}^K = \min_{V \subset V_{\mathcal{B}}^K,\ dim(V) = d} \max_{\substack{\Psi \in V, \\ \|\Psi\|_{L^2} = 1}} \langle \mathcal{H}\Psi, \Psi \rangle$$

where $V_{\mathcal{B}}^K = span(|\varphi_{i_1}, ..., \varphi_{i_N}\rangle,\ 1 \leqslant i_k \leqslant K)$. This method is called by chemists Full Configuration-Interaction*. It is very easy to see that*

$$\lambda_d \leqslant \lambda_d^K \leqslant \lambda_{\mathcal{B},d}^K$$

and we shall prove in section 4.4 that $\lim_{K \to +\infty} \lambda_{\mathcal{B},d}^K = \lambda_d$.

3.3 About the algorithms

We now make some comments on the connection between our min-max defining $\lambda_d^{K,\mathcal{I}}$ and the algorithms used by chemists. Although there are a lot of different approaches, the general form of the numerical algorithms used to calculate the $(d-1)^{th}$ excited state with a fixed geometry can be summarized as follows (see for instance [She87, Wer87, WM80]):

1. Start with some $(c_{\mathcal{I}}, \Phi)$ obtained for instance from a previous Hartree-Fock or Configuration-Interaction calculation.

2. Compute the matrix $H_\Phi^{\mathcal{I}}$ of the quadratic form $\mathcal{E}_N$ associated to $\mathcal{H}$ on the subspace $Span(\Phi_I,\ I \in \mathcal{I})$ where $\Phi_I = |\varphi_{i_1}, ..., \varphi_{i_N}\rangle$ if $I = \{i_1, ..., i_N\}$.

3. Find $c'_{\mathcal{I}}$ as the d^{th} eigenvector of this matrix.

4. This $c'_{\mathcal{I}}$ being fixed, minimize the energy with regard to Φ to obtain a new Φ'.

5. Replace $(c_{\mathcal{I}}, \Phi)$ by $(c'_{\mathcal{I}}, \Phi')$ and return to step 2.

According to chemists, the aim of such an algorithm is to solve a problem of the form

$$\tilde{\lambda}_d^{K,\mathcal{I}} = \inf_{\substack{\Phi=(\varphi_i)_{i=1}^K,\\ \int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}}} \min_{\substack{V\subset V^{K,\mathcal{I}}(\Phi),\\ dim(V)=d}} \max_{\substack{\Psi\in V,\\ \|\Psi\|_{L^2}=1}} \langle \mathcal{H}\Psi,\Psi\rangle \tag{I.23}$$

$$= \inf_{\substack{\Phi=(\varphi_i)_{i=1}^K,\\ \int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}}} \lambda_d(\Phi) \tag{I.24}$$

where $V^{K,\mathcal{I}}(\Phi) = Span(\Phi_I,\ I\in\mathcal{I})$ and $\lambda_d(\Phi)$ is the d^{th} eigenvalue of the matrix $H_\Phi^{\mathcal{I}} = (\langle\Phi_I,\mathcal{H}\Phi_J\rangle)_{I,J\in\mathcal{I}}$. In other words, *"the MCSCF energy results from minimizing the appropriate eigenvalue of the Hamiltonian matrix with respect to orbital variations"* [She87].

All the sets in the "min" part of (I.23) can be written $f(S^{d-1})$ with a $f\in\mathcal{C}_{(\mathbb{Z}_2)_c}(S^{d-1},\mathcal{M}_N^{K,\mathcal{I}})$ and so we easily obtain

$$\lambda_d \leqslant \lambda_d^{K,\mathcal{I}} \leqslant \tilde{\lambda}_d^{K,\mathcal{I}}.$$

However, it is not obvious that the min-max method (I.23) yields the existence of critical points. Let us notice that this may be viewed as a special case of the following abstract problem: let X be a Banach space, and $A: X\to\mathcal{S}_N(\mathbb{R})$ that is bounded and smooth ($\mathcal{S}_N(\mathbb{R})$ is the set of all $N\times N$ symmetric real matrices). We denote by $\lambda_d(x)$ the d^{th} eigenvalue of $A(x)$ and consider

$$\lambda_d = \inf_{x\in X}\lambda_d(x)$$

$$= \inf_{x\in X}\min_{dim(V)=d}\max_{\substack{v\in V,\\ \|v\|=1}}\langle A(x)v,v\rangle.$$

We introduce

$$F(x,v) = \langle A(x)v,v\rangle$$

and want to know if this method gives a critical point of F on $X\times S^{N-1}$.

To see that critical points may not exist when degeneracies of eigenvalues occur, we consider a simple example due to Rellich and adapted by Reed and Simon [RS78]:

$$A(x) = \begin{pmatrix} \sin(x_1) & \sin(x_2)\\ \sin(x_2) & -\sin(x_1)\end{pmatrix}$$

for $x=(x_1,x_2)\in X=\mathbb{R}^2$. We have

$$\lambda_2(x) = \sqrt{\sin(x_1)^2+\sin(x_2)^2}$$

and so

$$\lambda_2 = 0 = \lambda_2(0).$$

But $\frac{\partial F}{\partial x_1}(0,v) = \frac{\partial F}{\partial x_2}(0,v) = 0$ if and only if $v=0$, and this method does not give a critical point of F on $\mathbb{R}^2\times S^1$.

If $X=\mathbb{R}$ and A is analytic, a simplification appears due to a theorem of Rellich (see [RS78]), and critical points exist even when the optimal eigenvalue is degenerate. We now

show that the algorithm described previously may possibly not converge. Consider for instance

$$A(x) = \sin(x) \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

for $x \in X = \mathbb{R}$. We have

$$\lambda_2(x) = \sqrt{2}|\sin(x)|$$

and so

$$\lambda_2 = 0 = \lambda_2(0).$$

If we denote v_1 and v_2 two eigenvectors of $A'(0) = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ such that $\langle v_i, v_j \rangle = \delta_{ij}$, it is then easy to see that $(0, v)$ with $v = (v_1 + v_2)/\sqrt{2}$ is a critical point of F on $\mathbb{R} \times S^1$.

But for any initial guess $x_0 \in [\pi/2; \pi/2]$ such that $x_0 \neq 0$, the algorithm does not converge and oscillates between $x = \pi/2$ and $-\pi/2$. This shows that even if critical points exist, degeneracies may obstruct convergence.

In the case of the MC method, oscillations are frequently obtained when using the algorithm described before.

All these difficulties explain why we focused on the more "theoretical" min-max $\lambda_d^{K,\mathcal{I}}$ that takes into account the strong nonlinearity of $\mathcal{E}_N^K$.

4 Proofs

This section is devoted to the proof of our results.

4.1 Some preliminaries

In this section, we give some basic properties that will be used in the proof of our results.

4.1.1 Properties of $\mathcal{E}_N^K$

Lemma I.4. *(i) When it is defined by the formula (I.8), $\mathcal{E}_N^K$ is continuous on $\mathcal{M}_N^K$, and weakly lower semi-continuous on $\mathbb{R}^{\binom{K}{N}} \times (H^1(\mathbb{R}^3))^K$.*
(ii) (Kato's inequality). Given $\varepsilon > 0$ there exists a constant C_ε such that

$$\mathcal{E}_N^K(c, \Phi) \geqslant (1 - \varepsilon) \sum_{i=1}^{K} \frac{\gamma_i}{2} \|\nabla \varphi_i\|^2_{L^2(\mathbb{R}^3)} - C_\varepsilon \tag{I.25}$$

for all $(c, \Phi) \in \mathcal{M}_N^K$ with $\Gamma = diag(\gamma_1, ..., \gamma_K)$. In particular, $\mathcal{E}_N^K$ is bounded from below on $\mathcal{M}_N^K$. If $(c^n, \Phi^n) \in \mathcal{M}_N^K$ is such that $\mathcal{E}_N^K(c^n, \Phi^n)$ is bounded, then $(\sqrt{\Gamma^n} \cdot \Phi^n)$ is bounded in $(H^1(\mathbb{R}^3))^K$.

Proof – The continuity of $\mathcal{E}_N^K$ on $\mathcal{M}_N^K$ is obvious. To prove its weakly lower semi-continuity, we consider a sequence (c^n, Φ^n) in $\mathbb{R}^{\binom{K}{N}} \times (H^1(\mathbb{R}^3))^K$, such that $c^n \to c$ and $\Phi^n \rightharpoonup \Phi$ weakly in $(H^1(\mathbb{R}^3))^K$. We introduce $L = \liminf_{n \to +\infty} \mathcal{E}_N^K(c^n, \Phi^n)$ and suppose (by extracting

a subsequence if necessary) that $\Phi^n \to \Phi$ strongly in $(L^2_{loc}(\mathbb{R}^3))^K$ and a.e., and that $\mathcal{E}^K_N(c^n, \Phi^n) \to L$. Since for all $k_2 ... k_N$,

$$\sum_{i=1}^{K} \alpha^n_{i,k_2...k_N} \varphi^n_i \rightharpoonup \sum_{i=1}^{K} \alpha_{i,k_2...k_N} \varphi_i$$

weakly in $H^1(\mathbb{R}^3)$ and strongly in $L^2_{loc}(\mathbb{R}^3)$, we have

$$\int_{\mathbb{R}^3} V \left(\sum_{i=1}^{K} \alpha^n_{i,k_2...k_N} \varphi^n_i \right)^2 \to \int_{\mathbb{R}^3} V \left(\sum_{i=1}^{K} \alpha_{i,k_2...k_N} \varphi_i \right)^2$$

$$\int_{\mathbb{R}^3} \left| \sum_{i=1}^{K} \alpha_{i,k_2...k_N} \nabla \varphi_i \right|^2 \leqslant \liminf_{n \to +\infty} \int_{\mathbb{R}^3} \left| \sum_{i=1}^{K} \alpha^n_{i,k_2...k_N} \nabla \varphi^n_i \right|^2 .$$

But Fatou's lemma yields

$$\iint_{\mathbb{R}^6} \frac{\left(\sum_{i,j=1}^{K} \alpha_{i,j,k_3...,k_N} \varphi_i(x) \varphi_j(y) \right)^2}{|x-y|} dxdy$$

$$\leqslant \liminf_{n \to +\infty} \iint_{\mathbb{R}^6} \frac{\left(\sum_{i,j=1}^{K} \alpha^n_{i,j,k_3...,k_N} \varphi^n_i(x) \varphi^n_j(y) \right)^2}{|x-y|} dxdy$$

which easily concludes the proof of (i).

The proof of (I.25) is standard: by Kato's inequality, one easily shows

$$\langle \mathcal{H}\Psi, \Psi \rangle \geqslant (1-\varepsilon) \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \Psi|^2 - C_\varepsilon$$

for all $\Psi \in H^1_a(\mathbb{R}^{3N})$ such that $\|\Psi\|_{L^2} = 1$ (see [Fri03a]). Given a $(c, \Phi) \in \mathcal{M}^K_N$ such that $\Gamma = diag(\gamma_1 ... \gamma_K)$, we may apply this inequality to the associated wavefunction and obtain

$$\mathcal{E}^K_N(c, \Phi) \geqslant (1-\varepsilon) \sum_{i=1}^{K} \frac{\gamma_i}{2} \|\nabla \varphi_i\|^2_{L^2(\mathbb{R}^3)} - C_\varepsilon$$

which obviously implies that $\mathcal{E}^K_N$ is bounded from below.

Finally, if we take an arbitrary sequence (c^n, Φ^n) in $\mathcal{M}^K_N$ such that $\mathcal{E}^K_N(c^n, \Phi^n)$ is bounded, we can find for each n, a $U_n \in \mathcal{O}_K(\mathbb{R})$ such that $U_n \Gamma^n U_n^T = diag(\gamma^n_1 ... \gamma^n_K)$. We then use the latter inequality and the fact that $\mathcal{O}_K(\mathbb{R})$ is compact to obtain $(\sqrt{\Gamma^n} \cdot \Phi^n)$ bounded in $(H^1(\mathbb{R}^3))^K$. □

Remark I.5. *As it has been pointed out by G. Friesecke in [Fri03a], $\mathcal{E}_N$ is not weakly lower semi-continuous on $H^1_a(\mathbb{R}^{3N})$. This shows the benefit to work in the one-body space $\mathcal{M}^K_N$.*

4.1.2 About the inequality $Z > N - 1$

As in [LS77, Lio87, Gho93, Fri03a], our results are stated with $Z > N-1$ (i.e. for neutral molecules and positive ions). The reason is that we shall often use the following

Lemma I.5. *For all $k \geqslant 1$ and all $0 \leqslant m < Z$, there exists a k-dimensional subspace $V_{k,m}$ of $H^1(\mathbb{R}^3)$ and a $\delta_{k,m} > 0$ such that*

$$\int_{\mathbb{R}^3} \frac{1}{2}|\nabla\varphi|^2 + \left(V + \mu * \frac{1}{|r|}\right)\varphi^2 \leqslant -\delta_{k,m} < 0$$

for all $\varphi \in V_{k,m}$ with $\|\varphi\|_{L^2(\mathbb{R}^3)} = 1$, and all bounded non-negative measure μ on $\mathbb{R}^3$ that satisfies $\mu(\mathbb{R}^3) \leqslant m$.

Moreover, $V_{k,m}$ can be chosen such that $V_{k,m} = span(\varphi_1, ..., \varphi_k)$ where the $(\varphi_i)_{i=1}^k$ are radially symmetric functions in $\mathcal{D}(\mathbb{R}^3)$, with disjoint supports.

In particular, the Hamiltonian

$$H_\mu = -\frac{1}{2}\Delta + V + \mu * \frac{1}{|r|}$$

admit at least k eigenvalues strictly below $-\delta_{k,m}$.

Proof – See the proof of a similar lemma in [Lio87, Gho93]. □

4.2 Proof of Theorem I.1

Our method of proof is in the spirit of [Lio87], with the additional difficulties quoted before.

Let (c^n, Φ^n) be as in the Theorem. Since the method is natural, we may suppose that for all n, $\Gamma^n = diag(\gamma_1^n ... \gamma_K^n)$, with $0 \leqslant \gamma_i^n \leqslant 1$ and $\gamma_1^n \geqslant ... \geqslant \gamma_K^n$ (we recall that $\Gamma^n = I_N$ if $K = N$). We may also assume that $c_I^n \to c_I$ for all $I \in \mathcal{I}$. Thus each γ_i^n converges to some $\gamma_i \geqslant 0$.

By lemma I.4, for all $i = 1...K$, $\gamma_i^n \int_{\mathbb{R}^3} |\nabla\varphi_i^n|^2$ is bounded, so either $\gamma_i = 0$ or $(\varphi_i^n)_{n\in\mathbb{N}}$ is bounded in $H^1(\mathbb{R}^3)$.

Step 1 : Let us suppose for instance that $\gamma_K^n \to 0$ (this cannot happen when $K = N$ since $\gamma_i^n = 1$ for all $i = 1, ..., N$ and $n \in \mathbb{N}$). Since $(\varphi_K^n)_{n\in\mathbb{N}}$ is bounded in $L^2(\mathbb{R}^3)$, we may assume –by extracting subsequences if necessary– that $\sqrt{\gamma_K^n}\varphi_K^n \to 0$ in $L^2(\mathbb{R}^3)$ and $\sqrt{\gamma_K^n}\varphi_K^n \rightharpoonup 0$ weakly in $H^1(\mathbb{R}^3)$. We now show that this last convergence is strong.

Let $I \in \mathcal{I}$ be such that $K \in I$. Since $\mathcal{E}$ is quadratic with regard to the (c_I), we have

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

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

$$(c_I^n)^2 \int_{\mathbb{R}^3} \frac{\varphi_K^n(x)^2}{|x|}dx \leqslant \int_{\mathbb{R}^3} \frac{(\sqrt{\gamma_K^n}\varphi_K^n(x))^2}{|x|}\,dx \to 0.$$

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

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

Multiplying by c_I^n, we obtain

$$\left(\sum_{i\in I}\int_{\mathbb{R}^3}|\nabla\varphi_i|^2\right)(c_I^n)^2 \to 0.$$

as $n\to+\infty$. Since we have

$$\gamma_K^n = \sum_{I,\ K\in I}(c_I^n)^2 = \sum_{I\in\mathcal{I},\ K\in I}(c_I^n)^2,$$

this shows that $\sqrt{\gamma_K^n}\varphi_K^n \to 0$ strongly in $H^1(\mathbb{R}^3)$.

Let $K' \geqslant N$ be such that $\gamma_{K'+1}=0$ and $\gamma_{K'}\neq 0$. By the same argument, one has $\sqrt{\gamma_i^n}\varphi_i^n \to 0$ strongly in $H^1(\mathbb{R}^3)$ for $i=K'+1,...,K$.

We now introduce $\mathcal{I}' = \mathcal{I}\cap\mathcal{A}_N^{K'}$,

$$\tilde{c}_I^n = \frac{c_I^n}{\sqrt{\sum_{J\subset\{1...K'\}}(c_J^n)^2}} \quad \text{for all} \quad I\in\mathcal{A}_N^{K'},$$

$$\text{and} \quad \tilde{\Phi}^n = (\varphi_1^n...\varphi_{K'}^n)^T$$

so that $(\tilde{c}^n,\tilde{\Phi}^n)\in\mathcal{M}_N^{K',\mathcal{I}'}$. If $(\tilde{\Psi}^n)$ is the corresponding sequence of wave functions, the previous arguments imply that we have $\Psi^n = \tilde{\Psi}^n + r^n$ where $r^n\to 0$ in $H_a^1(\mathbb{R}^{3N})$. The convergence of $(\varphi_i^n)_{i=1}^{K'}$ will imply those of $\tilde{\Psi}^n$ and then Ψ^n. The new sequence $(\tilde{c}^n,\tilde{\Phi}^n)_{n\in\mathbb{N}}$ satisfies the assumptions 1), 2), 3) of Theorem I.1 on $\mathcal{M}_N^{K',\mathcal{I}'}$ and also $\tilde{\Gamma}^n \geqslant \gamma I_{K'}$ with $\gamma=\gamma_{K'}>0$. 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 $K'\times K'$ symmetric matrix Λ^n such that

$$\frac{\partial \mathcal{E}_N^K}{\partial \Phi}(c^n,\Phi^n) = \left(\left(-\frac{\Delta}{2}+V\right)\Gamma^n + 2W_{\Phi^n}\right)\cdot\Phi^n + \Lambda^n\cdot\Phi^n \to 0 \tag{I.26}$$

in $(L^2(\mathbb{R}^3))^K$. We now prove that there exists a $\lambda>0$ such that $\Lambda^n\geqslant\lambda I_{K'}$. Let U_n be an orthogonal matrix such that $\check{\Lambda}^n := U_n\Lambda U_n^T = diag(\lambda_1^n...\lambda_{K'}^n)$. We work in this step with $(\check{c}_n,\check{\Phi}_n) = U_n\cdot(c_n,\Phi_n)$. The second order condition of Theorem I.1 implies, for all $i=1,...,K'$ and all φ in a closed subspace F_n of $H^1(\mathbb{R}^3)$ of codimension at most $K+j$

$$\frac{\check{\gamma}_{ii}^n}{2}\int_{\mathbb{R}^3}|\nabla\varphi|^2 + \int_{\mathbb{R}^3}\left(\check{\gamma}_{ii}^n V + \check{\gamma}_{ii}^n\rho_i^n * \frac{1}{|r|} + \lambda_i^n + \check{\gamma}_{ii}^n\varepsilon_i^n\right)\varphi^2 - N(N-1)D_i^n \geqslant 0$$

where $\varepsilon_i^n\to 0$ and

$$\rho_i^n = \frac{N(N-1)}{\check{\gamma}_{ii}^n}\sum_{k_3...k_N}\left(\sum_k \check{\alpha}_{i,k,k_3...,k_N}^n\check{\varphi}_k^n\right)^2$$

$$D_i^n = \sum_{k_3\dots k_N} \iint_{\mathbb{R}^6} \frac{\left(\varphi(x)\sum_k \check{\alpha}^n_{i,k,k_3\dots,k_N}\check{\varphi}^n_k(x)\right)\left(\varphi(y)\sum_k \check{\alpha}^n_{i,k,k_3\dots,k_N}\check{\varphi}^n_k(y)\right)}{|x-y|}dxdy.$$

Hence we obtain (it is easy to see that $D_i^n \geqslant 0$)

$$\frac{1}{2}\int_{\mathbb{R}^3}|\nabla\varphi|^2 + \int_{\mathbb{R}^3}\left(V + \rho_i^n * \frac{1}{|r|} + \frac{\lambda_i^n}{\check{\gamma}^n_{ii}} + \varepsilon_i^n\right)\varphi^2 \;\geqslant\; 0 \tag{I.27}$$

for all $\varphi \in F_n$.

We now use the method of [Lio87] (see also [Gho93]) which is based on the following simple lemma

Lemma I.6. *Let A be a self-adjoint operator on a Hilbert space H and let H_1, H_2 be two subspaces of H such that $H = H_1 \oplus H_2$, $dim(H_1) = k < +\infty$ and $P_2AP_2 \geqslant 0$ (P_2 is the orthogonal projection onto H_2). Then A has at most k negative eigenvalues.*

Proof – See [Lio87, Gho93]. □

By Lemma I.6, we deduce that the Schrödinger operator $H_i^n := -\Delta/2 + V + \rho_i^n * \frac{1}{|r|}$ has at most $K + j$ eigenvalues strictly less than $-\frac{\lambda_i^n}{\check{\gamma}^n_{ii}} - \varepsilon_i^n$. On the other hand, since we have

$$\int_{\mathbb{R}^3}\rho_i^n = \frac{N(N-1)}{\check{\gamma}^n_{ii}}\sum_{k,k_3\dots k_N}(\check{\alpha}^n_{i,k,k_3\dots,k_N})^2 = N - 1 < Z,$$

we can use Lemma I.5 to find a $\delta > 0$ that does not depend on i such that H_i^n admits at least $K + j$ eigenvalues strictly below $-\delta$. We deduce that

$$\forall n,\ \frac{\lambda_i^n}{\check{\gamma}^n_{ii}} + \varepsilon_i^n \geqslant \delta$$

and thus we obtain from $\check{\Gamma}^n = U_n\Gamma^n U_n^T \;\geqslant\; \gamma I_{K'}$ (and for n large enough)

$$\lambda_i^n \geqslant \check{\gamma}^n_{ii}(\delta/2) \geqslant \gamma(\delta/2) = \lambda > 0.$$

This shows that $\check{\Lambda}^n \;\geqslant\; \lambda I_{K'}$ and so $\Lambda^n \;\geqslant\; \lambda I_{K'}$. But (I.26) yields

$$\mathcal{E}_N^K(c^n, \Phi^n) + \langle W_{\Phi^n}\cdot\Phi^n, \Phi^n\rangle + Tr(\Lambda^n) \to 0$$

and so (Λ^n) is bounded. We finally obtain (up to subsequences) $\Lambda^n \to \Lambda > 0$.

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

Passing to the limit in (I.26), we get

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

This implies

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

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

Step 4 : It remains to show that, when $\mathcal{I}=\mathcal{A}_N^K$, $K'=K$ or $K-1$. This is obtained by the same proof as in [LB94, Fri03b] to show the inequality $E_N^{K+2}<E_N^K$ and we refer the reader to these papers. Let us notice that this uses the analyticity of the orbitals, independently proved in Proposition I.3. This ends the proof of Theorem I.1. □

4.3 Proof of Corollary I.2

Let (c^n,Φ^n) be a minimizing sequence of $E_N^{K,\mathcal{I}}$. We use the minimization principle of Borwein and Preiss (see [BP87, Gho93]) to obtain a new minimizing sequence $(\tilde{c}^n,\tilde{\Phi}^n)\in\mathcal{M}_N^{K,\mathcal{I}}$ such that

1. $\|\tilde{\Phi}^n-\Phi^n\|_{(H^1(\mathbb{R}^3))^K}\to 0$ and $|\tilde{c}_I^n-c_I^n|\to 0$ when $I\in\mathcal{I}$
2. there exists a sequence $(\hat{c}^n,\hat{\Phi}^n)\in\mathcal{M}_N^{K,\mathcal{I}}$ such that

$$\|\hat{\Phi}^n-\Phi^n\|_{(H^1(\mathbb{R}^3))^K}\to 0,\quad |\hat{c}_I^n-c_I^n|\to 0\ \text{ when } I\in\mathcal{I},\ \text{ and}$$

$$\forall n\in\mathbb{N},\qquad \mathcal{E}_N^K(\tilde{c}^n,\tilde{\Phi}^n)+\frac{1}{n}\mathcal{Q}_n(\tilde{c}^n,\tilde{\Phi}^n)=\min_{\mathcal{M}_N^{K,\mathcal{I}}}\left\{\mathcal{E}_N^K+\frac{1}{n}\mathcal{Q}_n\right\}$$

where

$$\mathcal{Q}_n(c,\Phi)=\frac{1}{2}\left(\|c-\hat{c}_n\|^2+\|\Phi-\hat{\Phi}_n\|^2\right).$$

This new minimizing sequence $(\tilde{c}^n,\tilde{\Phi}^n)$ now satisfies the assumptions of Theorem I.1, with $j=0$. □

4.4 Proof of Theorem I.3

We recall that we shall prove (i) and (ii) by using the variational principles defining respectively the sequences $(\mu_d^{K,\mathcal{I}})_{d\geqslant 1}$ and $(\lambda_d^{K,\mathcal{I}})_{1\leqslant d\leqslant|\mathcal{I}|}$ – formulas (I.22) and (I.21).

Step 1 : Existence. We shall use here a Theorem of G. Fang and N. Ghoussoub [FG92, Gho93] that enables to obtain Palais-Smale sequences satisfying the assumptions of Theorem I.1. Let us now present this result.

Let X be a complete $\mathcal{C}^2$-Riemannian manifold, and G a compact Lie group acting freely and differentiably on X and on $\mathbb{R}^d$ for some $d\geqslant 1$. Let $\mathcal{F}$ be a G-homotopic family of dimension d, that is to say a set of the form:

$$\mathcal{F}=\{f(D),\ f\in\mathcal{C}_G(D,X)\}$$

where D in a fixed G-invariant compact subset of $\mathbb{R}^d$, and $\mathcal{C}_G(D,X)$ is the set of all G-equivariant continuous functions $f : D \to X$. We now have the following result, which is a simplified version of the original one in [FG92, Gho93]

Theorem I.4 (Fang-Ghoussoub [FG92, Gho93]). *Let φ be a G-invariant $\mathcal{C}^2$ functional on X with $d\varphi$ and $d^2\varphi$ Hölder-continuous on X, and consider*

$$c = \inf_{A\in\mathcal{F}} \max_{x\in A} \varphi(x).$$

Then for every min-maxing sequence $(A_n)_n$ in $\mathcal{F}$, there exist sequences $(x_n)_n$ in X and $(\delta_n)_n$ in $\mathbb{R}^+$ with $\lim_{n\to+\infty}\delta_n = 0$ such that

1. *$x_n \in A_n$ for each n,*
2. *$\lim_{n\to+\infty} \varphi(x_n) = c$,*
3. *$\lim_{n\to+\infty} d\varphi(x_n) = 0$,*
4. *for each n, $d^2\varphi(x_n)$ has at most d eigenvalues below $-\delta_n$.*

Now we easily see that the min-max (I.21,I.22) are defined with $(\mathbb{Z}_2)_{c/\Phi}$-homotopic classes of dimension d for each d (take $X = \mathcal{M}_N^K$ and $D = S^{d-1}$). Applying this result, we obtain a sequence (c^n, Φ^n) that satisfies the assumptions of Theorem I.1 with $j = d$. Therefore, it converges up to a subsequence, in the sense of Theorem I.1, to some critical point (c', Φ') of $\mathcal{E}_N^{K'}$ on $\mathcal{M}_N^{K',\mathcal{I}'}$ with $K' \leqslant K$. One may add any $(\varphi_{K'+1}, ..., \varphi_K)$ to obtain solutions of (I.9) with a $\Lambda \geqslant 0$.

Step 2 : Study of $(\mu_d^{K,\mathcal{I}})_{d\geqslant 1}$. Our proof uses ideas of [Lio87, Gho93]. Since in this step, $K \geqslant N$ and $\mathcal{I}$ are fixed, we shall forget these subscripts. We show that $-\infty < \mu_d \leqslant \mu_{d+1} < 0$ for all $d \geqslant 1$ and that $\lim_{d\to+\infty}\mu_d = 0$.

The fact that $-\infty < \mu_d \leqslant \mu_{d+1}$ is obvious. Let be $n_d = K \cdot d$. For each $d \geqslant 1$, by lemma I.5, there exists a n_d-dimensional subspace V_d of $H^1(\mathbb{R}^3)$ and a $\delta > 0$ such that

$$\int_{\mathbb{R}^3} \frac{1}{2}|\nabla\varphi|^2 + \left(V + m * \frac{1}{|r|}\right)\varphi^2 \leqslant -\delta < 0$$

for all $\varphi \in V_d$ with $\|\varphi\|_{L^2(\mathbb{R}^3)} = 1$, and all bounded non-negative measure m on $\mathbb{R}^3$ which satisfies $m(\mathbb{R}^3) = N - 1 < Z$. Let $(\varphi_1, ..., \varphi_{n_d})$ be a L^2-orthogonal basis of V_d, of radially symmetric functions in $\mathcal{D}(\mathbb{R}^3)$ with disjoints supports and such that $\int_{\mathbb{R}^3}(\varphi_i)^2 = 1$. For the sake of simplicity, we may suppose that $\{1, ..., N\} \in \mathcal{I}$. We consider :

$$\begin{array}{rccc} f: & S^{d-1} & \longrightarrow & \mathcal{M}_N^{K,\mathcal{I}} \\ & (\alpha_1, ..., \alpha_d) & \longmapsto & \left(\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix}, \begin{pmatrix}\alpha_1\varphi_1 + \cdots + \alpha_d\varphi_d\\ \vdots \\ \alpha_1\varphi_{d(K-1)+1} + \cdots + \alpha_d\varphi_{Kd}\end{pmatrix}\right) \end{array}$$

f is a continuous $(\mathbb{Z}_2)_\Phi$-equivariant function and

$$\begin{aligned}\mathcal{E}_N^K(f(\alpha_1,...,\alpha_d)) &= \\ &= \mathcal{E}_N(|\alpha_1\varphi_1+\cdots+\alpha_d\varphi_d,...,\alpha_1\varphi_{d(N-1)+1}+\cdots+\alpha_d\varphi_{Nd}\rangle) \\ &= \mathcal{E}_N\left(\sum_{i_1=1}^{d}\cdots\sum_{i_N=(N-1)d+1}^{Nd}\alpha_{i_1}\cdots\alpha_{i_N-(N-1)d}|\varphi_{i_1},...,\varphi_{i_N}\rangle\right) \\ &= \sum_{i_1=1}^{d}\cdots\sum_{i_N=(N-1)d+1}^{Nd}(\alpha_{i_1}\cdots\alpha_{i_N-(N-1)d})^2\mathcal{E}_N(|\varphi_{i_1},...,\varphi_{i_N}\rangle) \\ &= \sum_{i_1=1}^{d}\cdots\sum_{i_N=(N-1)d+1}^{Nd}(\alpha_{i_1}\cdots\alpha_{i_N-(N-1)d})^2\mathcal{E}^{HF}(\varphi_{i_1},...,\varphi_{i_N}).\end{aligned}$$

since the (φ_i) have disjoint supports.

We recall that

$$\begin{aligned}\mathcal{E}^{HF}(\varphi_1,...,\varphi_N) &= \sum_{i=1}^{N}\int_{\mathbb{R}^3}\left(\frac{1}{2}|\nabla\varphi_i|^2+V(\varphi_i)^2\right) \\ &+\frac{1}{2}\int_{\mathbb{R}^3}\left(\left(\sum_{j=1}^{N}(\varphi_j)^2\right)*\frac{1}{|r|}\right)\left(\sum_{i=1}^{N}(\varphi_i)^2\right)-\frac{1}{2}\sum_{i,j=1}^{N}\int_{\mathbb{R}^3}\left((\varphi_i\varphi_j)*\frac{1}{|r|}\right)(\varphi_i\varphi_j),\end{aligned}$$

so that

$$\begin{aligned}\mathcal{E}^{HF}(\varphi_1,...,\varphi_N) &= \sum_{i=1}^{N}\int_{\mathbb{R}^3}\left(\frac{1}{2}|\nabla\varphi_i|^2+\left(V+m_i*\frac{1}{|r|}\right)(\varphi_i)^2\right) \\ &-\frac{1}{2}\sum_{i\neq j=1}^{N}\int_{\mathbb{R}^3}\left((\varphi_i\varphi_j)*\frac{1}{|r|}\right)(\varphi_i\varphi_j)\end{aligned}$$

where

$$m_i=\sum_{j=1,\ j\neq i}^{N}(\varphi_j)^2 \quad\text{and}\quad \int_{\mathbb{R}^3}m_i=N-1.$$

Since $\int_{\mathbb{R}^3}\left((\varphi_i\varphi_j)*\frac{1}{|r|}\right)(\varphi_i\varphi_j)\geqslant 0$ for all $i\neq j$, we finally obtain

$$\begin{aligned}\mathcal{E}_N^K(f(\alpha_1,...,\alpha_d)) &\leqslant \sum_{i_1=1}^{d}\cdots\sum_{i_N=1}^{d}(\alpha_{i_1}\cdots\alpha_{i_N})^2(-N\delta) \\ &= -N\delta<0\end{aligned}$$

for all $(\alpha_1,...,\alpha_d)\in S^{d-1}$, which shows that $\mu_d\leqslant\max_{f(S^{d-1})}\mathcal{E}_N^K<0$.

To show that $\lim_{d\to+\infty}\mu_d=0$, we classically consider a sequence (V_d) of finite dimensional subspaces of $(H^1(\mathbb{R}^3))^K$ such that $dimV_d=d$ and $\bigcup_d V_d$ is dense in $(H^1(\mathbb{R}^3))^K$, and denote $W_d=V_d^\perp$. Since for all $d\geqslant 1$, $\mu_d<0$, there exists a $(\mathbb{Z}_2)_\Phi$-equivariant continuous f_d such that

$$\mu_d\leqslant\max_{f_d(S^{d-1})}\mathcal{E}_N^K<\frac{\mu_d}{2}.$$

By the Borsuk-Ulam Theorem, there exists $(c_d, \Phi_d) \in f_d(S^{d-1})$ such that $\Phi_d \in W_{d-1}$ and

$$\mathcal{E}_N^K(c_d, \Phi_d) \leqslant \frac{\mu_d}{2} < 0.$$

Since (c_d) is bounded and $\Phi_d \rightharpoonup 0$ weakly in $(H^1(\mathbb{R}^3))^K$, we obtain by lemma I.4

$$\liminf_{d\to+\infty} \mathcal{E}_N^K(c_d, \Phi_d) = 0$$

which implies $\lim_{d\to+\infty} \mu_d = 0$.

For each $d \geqslant 1$, there exists (c^d, Φ^d) and $\Lambda^d \geqslant 0$ such that $\mathcal{E}_N^K(c^d, \Phi^d) = \mu_d$ and

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

This yields

$$\left\langle W_{\Phi^d} \cdot \Phi^d \,,\, \Phi^d \right\rangle + tr(\Lambda_d) = -\mathcal{E}_N^K(c^d, \Phi^d) = -\mu_d \longrightarrow_{d\to+\infty} 0^+$$

and so

$$\left\langle W_{\Phi^d} \cdot \Phi^d \,,\, \Phi^d \right\rangle \to 0 \quad \text{and} \quad \Lambda_d \to 0 \tag{I.28}$$

as $d \to +\infty$. By Lemma I.4, $(\sqrt{\Gamma^d}\Phi^d)_d$ is bounded in $(H^1(\mathbb{R}^3))^K$, so one may suppose that $c_I^d \varphi_i^d \rightharpoonup \varphi_{I,i}$ weakly in $H^1(\mathbb{R}^3)$, strongly in $L^2_{loc}(\mathbb{R}^3)$ and a.e., for all $i = 1, ..., K$ and all I such that $i \in I$. But $\varphi_{I,i} = 0$ by (I.28) and this implies

$$\left\langle \left(-\frac{\Delta}{2} + V\right)\Gamma^d \cdot \Phi^d \,,\, \Phi^d \right\rangle \longrightarrow 0$$

as $d \to +\infty$. We now easily obtain $\sqrt{\Gamma^d}(\nabla\Phi^d) \to 0$ strongly in $(L^2(\mathbb{R}^3))^K$, thanks to the weakly lower semi-continuity of this last expression.

This ends the proof of (i).

Step 3 : Study of $(\lambda_d^{K,\mathcal{I}})_{1\leqslant d\leqslant|\mathcal{I}|}$. We first choose a sequence (Ψ_d) of eigenvectors of $\mathcal{H}$ associated with the eigenvalues (λ_d) such that $\int_{\mathbb{R}^{3N}} \Psi_i\Psi_j = \delta_{ij}$ and

$$\begin{aligned} \lambda_d = \mathcal{E}_N(\Psi_d) &= \max_{\substack{\Psi \in span(\Psi_1 ... \Psi_d),\\ \|\Psi\|_{L^2}=1}} \mathcal{E}_N(\Psi) \\ &= \min_{\substack{\Psi \in span(\Psi_1 ... \Psi_{d-1})^\perp,\\ \|\Psi\|_{L^2}=1}} \mathcal{E}_N(\Psi). \end{aligned}$$

We denote by $V_d = span(\Psi_1 ... \Psi_d)$ and $W_d = (V_d)^\perp$ its orthogonal complements in $H_a^1(\mathbb{R}^{3N})$.

In the sequel, we fix a $d \geqslant 1$ and assume that K and $\mathcal{I}$ are such that $d \leqslant |\mathcal{I}| \leqslant \binom{K}{N}$. Let $f : S^{d-1} \to \mathcal{M}_N^{K,\mathcal{I}}$ be a $(\mathbb{Z}_2)_c$-equivariant continuous function. We have

$$\max_{f(S^{d-1})} \mathcal{E}_N^K = \max_{\Pi_N^K \circ f(S^{d-1})} \mathcal{E}_N$$

and since $\Pi_N^K \circ f$ is odd, $\Pi_N^K \circ f(S^{d-1}) \cap W_{d-1} \neq \varnothing$. This implies

$$\max_{\Pi_N^K \circ f(S^{d-1})} \mathcal{E}_N \geqslant \min_{\substack{\Psi \in (V_{d-1})^\perp, \\ \|\Psi\|_{L^2}=1}} \mathcal{E}_N(\Psi) = \lambda_d$$

for all $(\mathbb{Z}_2)_c$-equivariant continuous f, and so

$$\lambda_d \leqslant \lambda_d^{K,\mathcal{I}}. \tag{I.29}$$

We now assume $\mathcal{I} = \mathcal{A}_N^K$ and forget the subscript $\mathcal{I}$. We fix a basis $\mathcal{B} = (\varphi_i)_{i\in\mathbb{N}^*}$ of $H^1(\mathbb{R}^3)$ such that $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$. We recall that every wave function Ψ can be expanded as a combination of Slater determinants built with (φ_i) :

$$\Psi = \sum_{1\leqslant i_1<\ldots<i_N} c_I |\varphi_{i_1}\ldots\varphi_{i_N}\rangle$$

where the sum converges in $H_a^1(\mathbb{R}^{3N})$. We then introduce

$$\begin{aligned}\mathcal{P}^K(\Psi) &= \frac{\displaystyle\sum_{1\leqslant i_1<\ldots<i_N\leqslant K} c_I|\varphi_{i_1}\ldots\varphi_{i_N}\rangle}{\sqrt{\displaystyle\sum_{1\leqslant i_1<\ldots<i_N\leqslant K} (c_I)^2}} \quad \text{when} \sum_{1\leqslant i_1<\ldots<i_N\leqslant K} (c_I)^2 \neq 0 \\ &= 0 \quad \text{otherwise,}\end{aligned}$$

so that $rank(\Psi) \leqslant K$ and $Range(\Gamma_\Psi) \subset span(\varphi_i)_{i=1}^K$.

We thus have $\mathcal{P}^K(\Psi) \to \Psi$ in $H_a^1(\mathbb{R}^{3N})$ as $K \to +\infty$. We now introduce

$$V_d^K = span(\mathcal{P}^K(\Psi_1), \cdots, \mathcal{P}^K(\Psi_d)) \text{ and } W_d^K = (V_d^K)^\perp.$$

We recall (see a previous remark) that $V_{\mathcal{B}}^K = span(|\varphi_{i_1}, ..., \varphi_{i_N}\rangle,\ 1 \leqslant i_k \leqslant K)$ and so V_d^K is a subspace of $V_{\mathcal{B}}^K$. We have the following lemma

Lemma I.7. *One has $dim(V_d^K) = d$ for K large enough, and*

$$\max_{\substack{\Psi \in V_d^K, \\ \|\Psi\|_{L^2}=1}} \mathcal{E}_N(\Psi) \to \lambda_d$$

as $K \to +\infty$.

Proof – We introduce $G^K = Gram(\mathcal{P}^K(\Psi_i))$, i.e. $G_{ij}^K = \int_{\mathbb{R}^{3N}} \mathcal{P}^K(\Psi_i)\mathcal{P}^K(\Psi_j)$. $G^K \to I_d$, so $dim(V_d^K) \to d$. The end of the proof is easy, thanks to the continuity of $\mathcal{E}_N$. □

We now assume that K is large enough so that $dim(V_d^K) = d$. Let $(\psi_1', ..., \psi_d')$ be an basis of V_d^K such that $\int_{\mathbb{R}^{3N}} \psi_i'\psi_j' = \delta_{ij}$, and $c^1, ..., c^d \in S^{\binom{K}{N}-1}$ be such that

$$\psi_k' = \sum_{1\leqslant i_1<\ldots<i_N\leqslant K} c_I^k |\varphi_{i_1}\ldots\varphi_{i_N}\rangle.$$

When $(\alpha_1, ..., \alpha_d) \in S^{d-1}$, we have

$$\sum_{k=1}^d \alpha_k\psi_k' = \sum_{1\leqslant i_1<\ldots<i_N\leqslant K} \left(\sum_{k=1}^d \alpha_k c_I^k\right) |\varphi_{i_1}\ldots\varphi_{i_N}\rangle$$

and $\| \sum_{k=1}^{d} \alpha_k \psi'_k \|_{L^2(\mathbb{R}^{3N})} = 1$.

We thus introduce the continuous $(\mathbb{Z}_2)_c$-equivariant function

$$\begin{array}{rccc} f: & S^{d-1} & \longrightarrow & \mathcal{M}_N^K \\ & (\alpha_1, ..., \alpha_d) & \longmapsto & \left(\sum_{k=1}^{d} \alpha_k c^k, \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_K \end{pmatrix} \right) . \end{array}$$

By the definition of λ_d^K, we obtain

$$\lambda_d^K \leqslant \lambda_{\mathcal{B},d}^K \leqslant \max_{f(S^{d-1})} \mathcal{E}_N^K = \max_{\Pi_N^K \circ f(S^{d-1})} \mathcal{E}_N = \max_{\substack{\Psi \in V_d^K, \\ \|\Psi\|_{L^2} = 1}} \mathcal{E}_N(\Psi).$$

Hence we have proved, by (I.29),

$$\lambda_d \leqslant \lambda_d^K \leqslant \lambda_{\mathcal{B},d}^K \leqslant \max_{\substack{\Psi \in V_d^K, \\ \|\Psi\|_{L^2} = 1}} \mathcal{E}_N(\Psi)$$

and so, by Lemma I.7,

$$\lim_{K \to +\infty} \lambda_d^K = \lambda_d \quad \text{and} \quad \lim_{K \to +\infty} \lambda_{\mathcal{B},d}^K = \lambda_d.$$

□

4.5 Proof of Proposition I.3

The assertion $\varphi_i \in \bigcap_{2 \leqslant p < 3} W^{2,p}(\mathbb{R}^3)$ is obtained by standard arguments. To prove the analyticity of the orbitals, we introduce $\varphi_{i,j} = (\varphi_i \varphi_j) * \frac{1}{|r|}$. Equations I.9 can be written on the form

$$\begin{cases} \gamma_i \left(-\dfrac{\Delta}{2} + V \right) \varphi_i + \displaystyle\sum_j \sum_{k<l} \beta_{k,l}^{i,j} \varphi_{k,l} \varphi_j + \sum_j \lambda_{i,j} \varphi_j = 0, & i = 1, ..., K \\ -\Delta \varphi_{k,l} = 4\pi \varphi_k \varphi_l, & 1 \leqslant k, l \leqslant K \end{cases}$$

with $\gamma_i > 0$.

Since $(\varphi_i, \varphi_{k,l})_{i,k,l}$ is a solution of an analytic non-linear elliptic system of partial differential equations, we may use for instance the results of [Mor58] or the method of [Kat96] to obtain the analyticity of the φ_i and $\varphi_{k,l}$ on Ω.

We now assume $\Lambda > 0$ and prove that the electronic density $\rho = \sum_{i=1}^{K} \gamma_i (\varphi_i)^2$ has an exponential fall-off, which will end the proof. To obtain an inequation for ρ, we multiply (I.15) by Φ to give

$$-\frac{1}{2} \sum_{i=1}^{K} \gamma_i (\Delta \varphi_i) \varphi_i + V\rho + 2(W_\varphi \cdot \Phi, \Phi) + (\Lambda \Phi, \Phi) = 0$$

where $(\ ,\ )$ denotes the usual scalar product of $\mathbb{R}^K$. This yields

$$-\frac{1}{4} \Delta \rho + V\rho + (\Lambda \Phi, \Phi) = -2(W_\varphi \cdot \Phi, \Phi) - \frac{1}{2} \sum_{i=1}^{K} \gamma_i |\nabla \varphi_i|^2$$

and since $W_\Phi(x)$ is a non negative matrix for all x,

$$-\frac{1}{4}\Delta\rho + V\rho + (\Lambda\Phi, \Phi) \leqslant 0.$$

Now, since Λ and Γ are positive, there exists an $\alpha > 0$ such that $\Lambda \geqslant \alpha\Gamma$, and so

$$(\Lambda\Phi, \Phi) \geqslant \alpha(\Gamma\Phi, \Phi) = \alpha\rho$$

which yields

$$-\frac{1}{4}\Delta\rho + V\rho + \alpha\rho \leqslant 0.$$

For $\rho \geqslant 0$ in $H^1(\mathbb{R}^3)$, it is now standard that this implies, thanks to the maximum principle,

$$\rho(x) \leqslant C_\beta e^{-\sqrt{4\beta}|x|}$$

for all $0 < \beta < \alpha$. □

Appendix

In order to emphasize the key role of the partial method (M_2) quoted in section 2.3, we want to tackle here the classification of wave functions of rank K when $N = 2$ and $K = N+2$. Existence and non-existence of those wave functions have been largely studied in the 50-60's (see [Löw55, And63, Col63, CY00]). Their general form is known by chemists in the case $N = 2$ [Col63, CY00] but we do not know if the case $K = N + 2$ has never been written down. We shall however give here the proof for both cases, since they are quite similar.

Proposition I.8.

(N=2 – Coleman [Col63, CY00]). *For all 2-electrons wavefunctions Ψ of finite rank, there exists a $p \in \mathbb{N}^*$ and $(\varphi_i)_{i=1}^{2p}$ with $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$ such that*

$$\Psi = \sum_{i=1}^{p} c_i |\varphi_{2i-1}\varphi_{2i}\rangle. \tag{I.30}$$

In particular, all such Ψ have an even rank.

(K=N+2). *For all N-electrons wavefunctions Ψ of rank $K \leqslant N + 2$, there exists $(\varphi_i)_{i=1}^{N+2}$ with $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$ and a $p \geqslant 1$ with $2p \leqslant N + 2$, such that*

$$\Psi = \Psi' \wedge |\varphi_{2p+1}...\varphi_{N+2}\rangle \quad \textit{where} \quad (\star\Psi') = -\sum_{i=1}^{p} c_i |\varphi_{2i-1}\varphi_{2i}\rangle. \tag{I.31}$$

In particular, there does not exist N-electrons wave functions of rank $N + 1$.

Let us first recall the standard notations used in this result. The bilinear operator

$$\wedge : H^1_a(\mathbb{R}^{3P}) \times H^1_a(\mathbb{R}^{3(N-P)}) \longrightarrow H^1_a(\mathbb{R}^{3N})$$

is defined by

$$|\varphi_1, ..., \varphi_P\rangle \wedge |\varphi_{P+1}, ..., \varphi_N\rangle = |\varphi_1, ..., \varphi_N\rangle,$$

whereas

$$(\star\Psi') = -\sum_{i=1}^{p} c_i |\varphi_{2i-1}\varphi_{2i}\rangle$$

means

$$\Psi' = \sum_{i=1}^{p} c_i |\varphi_1\varphi_2, ..., \varphi_{2i-3}\varphi_{2(i-1)}\varphi_{2i+1}\varphi_{2(i+1)}, ..., \varphi_{2p-1}\varphi_{2p}\rangle.$$

Hence, when $N = 2$ or $K = N+2$, every wavefunction can be expressed in a form such that two determinants have at least two distinct orbitals, a method quoted as an example (M_2) in section 2.3. This allows to limit significantly the number of determinants.

In (I.31), the case $p = 1$ corresponds to the Hartree-Fock case (Ψ is a Slater determinant), and the case $p = 2$ corresponds to doubly excited configurations, also studied by C. Le Bris in [LB94]

$$\Psi = \alpha|\varphi_1...\varphi_N\rangle + \beta|\varphi_1...\varphi_{N-2}\varphi_{N+1}\varphi_{N+2}\rangle.$$

Proof – Let us first consider the case $N = 2$. A basis $(\varphi_i)_{i=1}^{K}$ of $Range(\Gamma_\Psi)$ being fixed, we introduce the antisymmetric matrix $A = (\alpha_{ij})$ such that $\Psi = \sum \alpha_{ij}\varphi_i \otimes \varphi_j$. We have

$\Gamma = 2AA^T$, so $K = rank(\Gamma) = rank(A)$ is even. By formula (I.12), $A' = UAU^T$ when $\Phi' = U \cdot \Phi$ and thus, we may find a $U \in \mathcal{O}_K(\mathbb{R})$ such that

$$A' = \begin{pmatrix} A_1 & & \\ & \diagdown & \\ & & A_p \end{pmatrix} \quad \text{with } A_i = \begin{pmatrix} 0 & \alpha_i \\ -\alpha_i & 0 \end{pmatrix},\ \alpha_i \in \mathbb{R}^*$$

which easily ends the proof of this first point.

The dual case $K = N + 2$ is very similar. Given a wave function $\Psi = \sum_I c_I |\varphi_{i_1} ... \varphi_{i_N}\rangle$, we introduce $\beta_{ij} = (-1)^{i+j} c_I$ for $\{1...K\} \setminus I = \{i < j\}$, and the antisymmetric matrix $B = (\beta_{ij})$ (so $\beta_{ii} = 0$ and $\beta_{ji} = -\beta_{ij}$). Then, it is easy to see that $\Gamma = N(I_K - BB^T)$ and $B' = det(U)UBU^T$ when $\Phi' = U \cdot \Phi$. Now there exists a $U \in \mathcal{O}_K(\mathbb{R})$ such that

$$B' = \begin{pmatrix} B_1 & & & \\ & \diagdown & & \\ & & B_p & \\ & & & 0 \end{pmatrix} \quad \text{with } B_i = \begin{pmatrix} 0 & \beta_i \\ -\beta_i & 0 \end{pmatrix},\ \beta_i \in \mathbb{R}^*$$

which easily ends the proof of Proposition I.8. □

Bibliography

[AF97] P.W. Atkins and R.S. Friedman. *Molecular Quantum Mechanics. Third Edition.* Oxford University Press, 1997.

[And63] T. Ando. Properties of fermions density matrices. *Rev. Mod. Phys.*, 35(3):690–702, 1963.

[BP87] J. Borwein and D. Preiss. A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions. *Trans. AMS*, 303:517–527, 1987.

[C$^+$03] E. Cancès et al. *Computational quantum chemistry: a primer*, volume X of *Handbook of Numerical Analysis*, pages 3–270. Elsevier, 2003.

[Col63] A.J. Coleman. Structure of fermions density matrices. *Rev. Mod. Phys.*, 35(3):668–689, 1963.

[CY00] A.J. Coleman and V.I. Yukalov. *Reduced Density Matrices: Coulson's Challenge.* Springer Verlag, 2000.

[ES99] M.J. Esteban and E. Séré. Solutions of the Dirac-Fock equations for atoms and molecules. *Commun. Math. Phys.*, 203:499–530, 1999.

[FG92] G. Fang and N. Ghoussoub. Morse-type information on Palais-Smale sequences obtained by min-max principles. *Manuscripta Math.*, 75:81–95, 1992.

[Fri03a] G. Friesecke. The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Arch. Rat. Mech. Analysis*, 169:35–71, 2003.

[Fri03b] G. Friesecke. On the infinitude of nonzero eigenvalues of the single-electron density matrix for atoms and molecules. *Proc. Roy. Soc. London Ser. A*, 459:47–52, 2003.

[Gho93] N. Ghoussoub. *Duality and perturbation methods in critical point theory.* Cambridge University Press, 1993.

[HG96] M. Head-Gordon. Quantum chemistry and molecular processes. *J. Phys. Chem.*, 100:13213–13225, 1996.

[HOS94] M. and T. Hoffmann-Ostenhof and H. Stremnitzer. Local properties of Coulombic wave functions. *Commun. Math. Phys.*, 163:185–215, 1994.

[HS00] W. Hunziker and I.M. Sigal. The quantum N-body problem. *J. Math. Phys*, 41(6):3348–3509, 2000.

[Kat57] T. Kato. On the eigenfunctions of many particle systems in quantum mechanics. *Commun. Pure and Appl. Math.*, 10:151–171, 1957.

[Kat96] K. Kato. New idea for proof of analyticity of solutions to analytic non linear equations. *SUT Journal of Math.*, 32(2):157–161, 1996.

[LB93] C. Le Bris. *Quelques problèmes mathématiques en chimie quantique moléculaire*. PhD thesis, École Polytechnique, 1993.

[LB94] C. Le Bris. A general approach for multiconfiguration methods in quantum molecular chemistry. *Ann. Inst. H. Poincaré Anal. Non linéaire*, 11(6):441–484, 1994.

[LB99] C. Le Bris. Mathématiques appliquées et chimie quantique. *Matapli - Bulletin de la Société de Mathématiques Appliquées et Industrielles*, 58:25–36, 1999.

[Lew02] M. Lewin. The multiconfiguration methods in quantum chemistry: Palais-Smale condition and existence of minimizers. *C. R. Acad. Sci. Paris, Ser. I*, 334:299–304, 2002.

[Lie90] E.H. Lieb. The stability of matter : from atoms to stars. *Bull. Amer. Math. Soc.*, 22:1–49, 1990.

[Lio87] P.L. Lions. Solutions of Hartree-Fock equations for Coulomb systems. *Commun. Math. Phys.*, 109:33–87, 1987.

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

[Löw55] 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):1474–1489, 1955.

[Löw59] P.O. Löwdin. Correlation problem in many-electron quantum mechanics. I. Review of different approaches and discussion of some current ideas. *Adv. Chem. Phys.*, 2:207–322, 1959.

[Mor58] C. B. Morrey Jr. On the analyticity of the solutions of analytic non-linear elliptic systems of partial differential equations. *Amer. J. Math.*, 80:198–218, 1958.

[Rab86] P.H. Rabinowitz. *Minimax methods in critical point theory with applications to differential equations*. Number 65 in Conference Board of the Mathematical Sciences, Regional Conference series in Mathematics. American Mathematical Society, 1986.

[Roo87] B. O. Roos. The Complete Active Space Self-Consistent Field method and its applications in electronic structure calculation. *Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys.*, 69:399–446, 1987.

[RS78] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, volume IV – Analysis of Operators. Academic Press, 1978.

[She87] R. Shepard. The multiconfiguration self-consistent field method. *Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys.*, 69:63–200, 1987.

[Wer87] H.-J. Werner. Matrix-formulated direct Multiconfiguration Self-Consistent Field and Multiconfiguration Reference Configuration-Interaction methods. *Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys.*, 69:1–62, 1987.

[WM80] H.-J. Werner and W. Meyer. A quadratically convergent multiconfiguration-self-consistent field method with simultaneous optimization of orbitals and CI coefficients. *J. Chem. Phys.*, 73(5):2342–2356, 1980.

[Zhi60] G. M. Zhislin. Discussion of the spectrum of Schrödinger operators for systems of many particles. (in Russian). *Trudy Moskovskogo matematiceskogo obscestva*, 9:81–120, 1960.

CEREMADE, CNRS UMR 7534,
Université Paris IX Dauphine,
Place du Maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France.
lewin@ceremade.dauphine.fr

(Accepted May 19, 2003)

A.II

Computing Electronic Structures: a new Multiconfiguration Approach for Excited States

Ce chapitre reprend le texte intégral d'un article écrit en collaboration avec Éric Cancès et Hervé Galicher.

Résumé

Nous présentons une nouvelle méthode pour le calcul d'états excités de molécules, basée sur un travail récent de l'un d'entre nous [Lew04] concernant les modèles multiconfigurations. Contrairement aux méthodes précédentes, notre algorithme converge toujours vers un état stationnaire du modèle, qui peut être interprété comme un état excité approché de la molécule.

Nous comparons également notre nouvelle définition aux autres approches et tentons d'expliquer les défauts qui sont parfois observés avec les méthodes antérieures.

Notre définition est variationnelle. Pour calculer le premier état excité, des chemins doivent être déformés sur une certaine variété, comme pour le calcul d'un état de transition entre réactants et produits sur une surface d'énergie potentielle. Nous proposons ici une méthode générale pour la déformation de chemins, pouvant être utile dans d'autres situations.

Enfin, nous présentons des résultats numériques pour le cas de systèmes à deux électrons : nous traçons l'énergie du premier état excité en fonction de la distance inter-atomique pour la molécule H_2, et donnons des résultats pour des atomes à deux électrons comme l'hélium.

Computing Electronic Structures: a new Multiconfiguration Approach for Excited States

Éric CANCÈS[a], Hervé GALICHER[a] and Mathieu LEWIN[b]

Abstract

We present a new method for the computation of electronic excited states of molecular systems. This method is based upon a recent theoretical definition (due to one of us, see M. Lewin, *Solutions of the Multiconfiguration Equations in Quantum Chemistry*, Arch. Rat. Mech. Anal. 171 (2004) 83-114) of the Multiconfiguration Self-Consistent-Field (MCSCF) excited states. Contrarily to previously used methods, our algorithm always converges to a stationnary te of the model, which can be interpreted as an approximate excited state of the molecule.

We also compare our definition to other approaches and give some explanation of the unsatisfactory behaviours which are sometimes observed when using the latter.

Our definition is variational. To compute the first excited state, one has to deform paths on a manifold, like this is usually done in the search for transition states between reactants and products on potential energy surfaces. We propose here a general method for the deformation of paths which could also be useful in other settings.

Numerical results for the special case of two-electrons systems are provided. We compute the first singlet excited state energy potential energy surface of the H_2 molecule, and give some numerical results concerning Helium-like atoms.

Electronic excited states play an essential role in various phenomena of high interest, such as photo-induced chemical reactions, femtosecond spectroscopy, or laser control of molecular processes. Whereas most of the currently used electronic structure models, notably the Hartree-Fock and the Kohn-Sham models, are rigorously founded and quite successful in the description of ground states, their approach to excited states is questionable [HG96]. The method which seems to be best-adapted to this issue is to date the multiconfiguration self-consitent field (denoted by MCSCF in the following) method [Löw55c, Löw59, She87]; loosely speaking, this approach leads to variational models which fill the gap between the mean-field Hartree-Fock and the N-body Schrödinger models [Löw59]. However the definition of what actually is an excited state for a nonlinear theory such as MCSCF is still unclear; it is indeed observed that nonlinear electronic structure models have a lot of spurious critical points that cannot be interpreted as approximations

of excited states. In other words, solving the equations of the model is clearly not sufficient to obtain a state which really approximates some excited state. In addition, even if we leave aside the above mentioned difficulty, the practical calculation of MCSCF critical points is problematic: numerical algorithms available to date do not always converge. More precisely, they sometimes lead to oscillations between two states of different energies, none of them being a critical point. For all these reasons, the computation of electronic excited states remains one of the main challenges of modern Quantum Chemistry.

In [Lew04], it is emphasized that those difficulties are likely to stem from the currently used definitions of MCSCF excited states that are not correct, for they do not fully take into account the nonlinearity of the model. The purpose of [Lew04] was to provide a more rigorous definition of MCSCF excited states. Our goal in this paper is to show that this theoretical definition can actually be used in practice, at least for the computation of the *first* excited state.

The paper is organized as follows. In Section 1, we introduce the MCSCF description of electronic structures. In Section 2, we present the new definition of MCSCF excited states and compare it to other definitions currently used in Computational Chemistry. Finally, in Section 3, we describe in details our new algorithm and present numerical results for the case of two-electrons systems.

1 MCSCF approximation of the time-independent Schrödinger equation

In this section we recall some classical properties of the N-body time-independent Schrödinger equation, and briefly present the MCSCF approximation. We refer the reader to [AF97, CDK$^+$03, Lew04, She87] for more details.

Let us consider a molecular system consisting of N electrons, and of M nuclei of positive charges $z_1, \cdots, z_M$. The nuclei are supposed to be correctly described by a classical model and are represented by pointwise charges clamped at positions $\bar{x}_1, \cdots, \bar{x}_M$ ($\bar{x}_m \in \mathbb{R}^3$ for $1 \leq m \leq M$). This is the so-called Born-Oppenheimer approximation. The electrons are described by the N-body quantum hamiltonian (written in atomic units, see e.g. [CDK$^+$03])

$$H_N = \sum_{i=1}^{N} \left(-\frac{1}{2}\Delta_{x_i} + V(x_i) \right) + \sum_{1 \leqslant i < j \leqslant N} \frac{1}{|x_i - x_j|}, \tag{II.1}$$

which acts on normalized electronic wavefunctions $\Psi(x_1, ..., x_N) \in L^2_a((\mathbb{R}^3)^N)$, $\|\Psi\|_{L^2} = 1$. The subscript a indicates that, due to the fermionic nature of the electrons, one solely considers wavefunctions Ψ which are antisymmetric under permutations of variables:

$$\forall \sigma \in S_N,\ \Psi(x_1, ..., x_N) = \varepsilon(\sigma)\Psi(x_{\sigma(1)}, ..., x_{\sigma(N)}) \quad \text{a.e.}$$

Here and below, S_N denotes the set of the permutations of the indices $\{1, ..., N\}$ and $\varepsilon(\sigma)$ the signature of the permutation σ. Finally, V is the electrostatic potential generated by the nuclei

$$V(x) = -\sum_{m=1}^{M} \frac{z_m}{|x - \bar{x}_m|}.$$

In what follows, we denote by $Z = \sum_{m=1}^{M} z_m$ the total nuclear charge which is an integer as we work in atomic units.

For the sake of clarity, we do not take the spin into account in the first two sections of the article, but the following arguments can be straightforwardly adapted to the case of spin-dependent wavefunctions. The spin will be reintroduced in Section 3, in which numerical examples on real molecular systems will be provided.

The operator H_N is self-adjoint in $L^2_a((\mathbb{R}^3)^N)$, with domain $H^2_a((\mathbb{R}^3)^N)$ and form domain $H^1_a((\mathbb{R}^3)^N)$. Some of its spectral properties are known: the spectrum $\sigma(H_N)$ of H_N is bounded from below, its essential spectrum is a range of the form $[\Sigma;+\infty)$ with $\Sigma \leqslant 0$, and its point spectrum is a subset of $(-\infty;0]$. The *N-body ground state energy* is the minimum of $\sigma(H_N)$ also defined by

$$E_N = \inf\{\langle \Psi, H_N \Psi\rangle,\ \Psi \in H^1_a(\mathbb{R}^{3N}),\ \|\Psi\|_{L^2_a(\mathbb{R}^{3N})} = 1\}. \tag{II.2}$$

The condition $E_N < \Sigma$ translates as the physical property of the nuclei being able to bind the N electrons in their vicinity. Zhislin showed in [Zhi60] that when $Z > N-1$ (an assumption that we will make throughout this article),

$$\sigma(H_N) = \{E_N = \lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n \leqslant \cdots\} \cup [\Sigma;+\infty)$$

where $(\lambda_i)_{i\geqslant 1}$ are eigenvalues strictly below and which converge to Σ.

The eigenfunctions corresponding to the $\lambda_i > E_N$ are called *excited states*. Both the ground states and the excited states obviously solve the *time-independent Schrödinger equation*

$$H_N \Psi = \lambda_i \Psi. \tag{II.3}$$

Recall that the excited state energies λ_d, $d \geq 1$, can be obtained by the Rayleigh-Ritz principle

$$\lambda_d = \min_{dim(W)=d} \max_{\substack{\Psi\in W,\\ \|\Psi\|_{L^2}=1}} \langle \Psi, H\Psi\rangle \tag{II.4}$$

where the first minimum is taken over all d-dimensional subspaces W of the domain of H_N.

The Schrödinger equation is a model of extremely high accuracy (except for heavy atoms for which core electrons are relativistic). For systems involving a few (say today six or seven) electrons, a direct Galerkin discretization of problem (II.3) is possible; such a technique is refered to as *Full CI* in Computational Chemistry. For larger systems, this approach is out of reach, due to the excessive dimension of the space $\mathbb{R}^{3N}$ on which the wavefunctions are defined, and problem (II.3) must then be approximated. To date, the most commonly used approximations are the Hartree-Fock model (see e.g. [Löw55b]) on the one hand, and the Kohn-Sham model (see e.g. [KS65, DG90]) on the other hand. Both of them have been designed for the calculation of ground states and are not really adapted to the calculation of excited states. On the contrary, the MCSCF approximation can be applied to both ground and excited state calculations.

The MCSCF method is based on the following remark:

$$L^2_a((\mathbb{R}^3)^N) = \bigwedge_{n=1}^{N} L^2(\mathbb{R}^3),$$

an equality which can be explicited in the following way. Consider an orthonormal basis $(\varphi_i)_{1\leq i<+\infty}$ of $L^2(\mathbb{R}^3)$. It is well-known that the sequence $(\varphi_{i_1}\otimes\cdots\otimes\varphi_{i_N})_{1\leq i_k<\infty}$ forms an orthonormal basis of $L^2((\mathbb{R}^3)^N)=\bigotimes_{n=1}^N L^2(\mathbb{R}^3)$, where by definition

$$(\varphi_{i_1}\otimes\cdots\otimes\varphi_{i_N})(x_1,...,x_N)=\varphi_{i_1}(x_1)\cdots\varphi_{i_N}(x_N).$$

An orthonormal basis of the subspace $L^2_a((\mathbb{R}^3)^N)$ of $L^2((\mathbb{R}^3)^N)$ can then be obtained by simply considering the antisymmetrized product $(\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N})_{1\leq i_1<\cdots<i_N<+\infty}$ where $(\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N})$ denotes the so-called Slater determinant of the φ_{i_k}:

$$\begin{aligned}(\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N})(x_1,...,x_N) &= \frac{1}{\sqrt{N!}}\sum_{\sigma\in S_N}\varepsilon(\sigma)\varphi_{i_1}(x_{\sigma(1)})\cdots\varphi_{i_N}(x_{\sigma(N)})\\ &= \frac{1}{\sqrt{N!}}\det(\varphi_{i_k}(x_l)).\end{aligned}$$

In other words *every antisymmetric wavefunction Ψ is an infinite linear combination of such Slater determinants:*

$$\Psi=\sum_{1\leq i_1<\cdots<i_N<+\infty}c_{i_1\ldots i_N}\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N},$$

the sum being convergent in $L^2_a((\mathbb{R}^3)^N)$. Remark that $\|\Psi\|_{L^2}=1$ is then equivalent to the condition $\sum_{i_1<\cdots<i_N}|c_{i_1\ldots i_N}|^2=1$.

An integer $K\geq N$ being fixed, we now consider the subset of $L^2_a((\mathbb{R}^3)^N)$ consisting of the wavefuntions Ψ which are *finite* linear combinations of the $\binom{K}{N}$ Slater determinants constructed from a set of K orthonormal functions $(\varphi_1,\cdots,\varphi_K)$ of $L^2(\mathbb{R}^3)$, i.e.

$$\Psi=\sum_{1\leq i_1<\cdots<i_N\leq K}c_{i_1\ldots i_N}\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N}. \tag{II.5}$$

The MCSCF approach is a variational method for approximating (II.3) in which both the coefficients $c_{i_1\ldots i_N}$ and the functions $(\varphi_1,\cdots,\varphi_K)$ are variational parameters. Let us mention incidently that the MCSCF method differs from the Configuration-Interaction (CI) method [Löw55a], for in the latter, only the coefficients $c_{i_1\ldots i_N}$ are variational parameters (in a CI calculation, the functions $(\varphi_1,\cdots,\varphi_K)$ are issued from a previous Hartree-Fock or Kohn-Sham calculation and are kept fixed).

Following our purpose to describe the MCSCF approach, we therefore introduce the manifold

$$\mathcal{M}_N^K=\left\{(c,\Phi)\in\mathbb{R}^{\binom{K}{N}}\times(H^1(\mathbb{R}^3))^K,\ \sum_{i_1<\cdots<i_N}|c_{i_1\ldots i_N}|^2=1,\ \int_{\mathbb{R}^3}\varphi_i\varphi_j=\delta_{ij}\right\}$$

where we have denoted by

$$c=(c_{i_1\cdots i_N})\in\mathbb{R}^{\binom{K}{N}},\qquad \Phi=(\varphi_1,...,\varphi_K)\in H^1(\mathbb{R}^3)^K$$

(we arrange the $c_{i_1\cdots i_N}$ in a column vector c using for instance the lexicographical order). Let us note that the functions $(\varphi_1,\cdots,\varphi_K)$ are now requested to have a H^1 regularity, in

order to ensure that the MCSCF energy (see formula (II.8) below) is well defined. Indeed, the MCSCF energy functional that we denoted here by $\mathcal{E}_N^K$, is defined by the formula

$$\mathcal{E}_N^K(c,\Phi) = \langle \Psi_{(c,\Phi)}, H_N \Psi_{(c,\Phi)} \rangle \tag{II.6}$$

$$\Psi_{(c,\Phi)} = \sum_{1\leq i_1<\cdots<i_N\leq K} c_{i_1\ldots i_N} \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}.$$

The MCSCF ground state energy is then

$$E_N^K = \inf_{\mathcal{M}_N^K} \mathcal{E}_N^K. \tag{II.7}$$

If we write $\Psi_{(c,\Phi)}$ as

$$\Psi_{(c,\Phi)} = \sum_{1\leq i_1,\ldots,i_N\leq K} \alpha_{i_1\ldots i_N} \varphi_{i_1} \otimes \cdots \otimes \varphi_{i_N},$$

with

$$\begin{aligned} \alpha_{i_1\ldots i_N} &= 0 \quad \text{if } \#\{i_1\ldots i_N\} < N \\ &= \frac{\varepsilon(\sigma)}{\sqrt{N!}} c_{i_{\sigma(1)}\ldots i_{\sigma(N)}} \text{ otherwise,} \end{aligned}$$

where σ is the permutation of the indices $\{1,2,\cdots,N\}$ such that $i_{\sigma(1)} < \ldots < i_{\sigma(N)}$), the MCSCF energy functional reads

$$\begin{aligned} \mathcal{E}_N^K(c,\Phi) = N \sum_{1\leqslant k_2,\ldots,k_N\leqslant K} \int_{\mathbb{R}^3} \frac{1}{2} \left| \sum_{i=1}^K \alpha_{i,k_2\ldots k_N} \nabla\varphi_i \right|^2 + V \left(\sum_{i=1}^K \alpha_{i,k_2\ldots k_N} \varphi_i \right)^2 \\ + \frac{N(N-1)}{2} \sum_{1\leqslant k_3,\ldots,k_N\leqslant K} \iint_{\mathbb{R}^6} \frac{\left(\sum_{1\leqslant i,j\leqslant K} \alpha_{i,j,k_3\ldots,k_N} \varphi_i(x)\varphi_j(y) \right)^2}{|x-y|} dxdy. \end{aligned} \tag{II.8}$$

Let us point out that, whereas the Schrödinger energy functional $\langle \Psi, H_N \Psi \rangle$ is quadratic, the MCSCF energy functional is not. Consequently, the MCSCF equations, namely the first order stationarity conditions for the critical points of $\mathcal{E}_N^K$ on the manifold $\mathcal{M}_N^K$, will be nonlinear.

Remark that when $K = N$, the set $\mathcal{M}_N^N$ reduces to the set of Slater determinants and one recovers the celebrated Hartree-Fock approximation [Löw55b, LS77, Lio87]. The difference between the Hartree-Fock and the exact (non-relativistic) ground state energy [Löw59]

$$E^{\text{corr}} = E_N^N - E_N$$

is called the *correlation energy*, for it originates from correlations between the positions of individual electrons, which are averaged out by the mean-field Hartree-Fock scheme. Estimating the correlation energy is essential for reliably calculating many of the fundamental properties of molecules [AF97, HG96], in particular in situations where the Hartree-Fock method fails. Since

$$\lim_{K\to+\infty} E_N^K = E_N,$$

the MCSCF method is a method of choice for computing the correlation energy.

Mathematically, it is known that a minimizer of (II.7) exists, and that the associated wavefunction converges to the ground state of H_N as K goes to infinity [LB94, Fri03, Lew04]. A minimizer of (II.7) can be numerically computed by a Newton-like method [Wer87, WM80, She87, ER81, MFS92] in which the orbitals φ_i and the coefficients c are optimized simultaneously. For the Hartree-Fock model, efficient numerical methods based on combinations of fixed-point and optimization strategies are available [CDK+03]. Unfortunately, such algorithms are specifically designed for solving the Hartree-Fock problem and seem to be difficult to adapt to the more general MCSCF setting.

Remark that in (II.5), all the Slater determinants that can be built with the functions φ_i are taken into account. Most often, this cannot be done in practice for $\binom{K}{N}$ is too large a number. It is then necessary to resort to an additional approximation consisting in dividing the electrons into two groups, the *inactive* electrons that are supposed to be correctly described by a Hartree-Fock type model, and the *active* electrons that mostly contribute to the correlation energy, and in using the MCSCF methodology for the active electrons only. This is the so-called CASSCF (Complete Active Space Self-Consistent Field) approach [Roo87]. All what we shall mention here can be straightforwardly adapted to the CASSCF setting. In particular, the first excited state of a CASSCF model can be computed using a slightly modified version of the numerical algorithm presented in Section 2.3.

2 On the definition of MCSCF excited states

Numerical investigations show that the MCSCF energy $\mathcal{E}_N^K$ possesses a lot of critical points on the manifold $\mathcal{M}_N^K$. This abundance of critical points is probably due to the nonlinearity of the energy functional. In [Lew04, Theorem 2 (i)] it is shown how an infinity of critical points can be exhibited, following a variational method previously applied for the Hartree-Fock theory in [Lio87]. It seems very difficult to decide in practice whether some critical point can be interpreted as an excited state of the molecular system (in particular, considering the Morse index[1] is not enough). To achieve this goal, we need to be able to "follow" the critical points as K grows to infinity, and see whether they actually converge to the true excited states of the linear N-body Schrödinger model (II.3).

2.1 Currently used definition of MCSCF excited states

Since there are two sets of parameters (the orbitals φ_i and the coefficients c), the *multi-configuration equations* take the following general form

$$\left\{ \begin{array}{l} \gamma_i \left(-\frac{\Delta}{2} + V \right) \varphi_i + \displaystyle\sum_{1 \leq j,k,l \leq K} b_{ijkl} \left((\varphi_j \varphi_k) * \frac{1}{|x|} \right) \varphi_l = \sum_{j=1}^{K} \lambda_{ij} \varphi_j, \ 1 \leq i \leq K \\ H_\Phi \cdot c = \beta c, \end{array} \right. \tag{II.9}$$

where the b_{ijkl} are real numbers which can be expressed in terms of c (see [Fri03] and [Lew04], where a compact form of the first equations is also given). The first line of (II.9) is in fact a system of K nonlinear coupled partial differential equations accounting for the stationarity conditions with respect to Φ; the symmetric matrix (λ_{ij}) is the Lagrange

[1] Recall that the Morse index of a critical point is the number of negative eigenvalues of the Hessian matrix.

multiplier associated with the orthonormality constraints on Φ. The numbers γ_i are called the *occupation numbers* and satisfy $0 \leq \gamma_i \leq 1$ (see [Lew04] for details); they all equal 1 in the Hartree-Fock method. The second equation conveys the stationarity condition with respect to c. As both the energy and the constraints are quadratic with respect to c, it has the form of an eigenvalue problem. It follows from (II.6) that the matrix H_Φ is defined by

$$(H_\Phi)_{IJ} = \langle \Phi_I, H_N \Phi_J \rangle,$$

where

$$\Phi_I = \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N}, \quad \text{when} \quad I = \{i_1 < \cdots < i_N\}.$$

In other words, H_Φ is the $\binom{K}{N} \times \binom{K}{N}$ matrix of the quadratic form associated with H_N when it is restricted to the $\binom{K}{N}$-dimensional space

$$V_\Phi = \mathrm{Span}(\varphi_{i_1} \wedge \cdots \wedge \varphi_{i_N},\ 1 \leq i_1 < \cdots < i_N \leq K).$$

Let us introduce $\lambda_d^K(\Phi)$, $d = 1, \cdots, \binom{K}{N}$, the eigenvalues of the matrix H_Φ. Obviously,

$$\lambda_d \leq \lambda_d^K(\Phi)$$

for all K and d such that $d \leq \binom{K}{N}$, and all $\Phi \in H^1(\mathbb{R}^3)^K$ such that $\int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}$. This inequality suggests the currently used definition of excited state energies [Löw59, She87, WM80, CER79, Wer87, MFS92]

$$\boxed{\mu_d^K = \inf_\Phi \lambda_d^K(\Phi),} \tag{II.10}$$

that is to say, quoting [She87], *"the MCSCF energy results from minimizing the appropriate eigenvalue of the hamiltonian matrix with respect to orbital variations"*. It can be shown that the expected condition

$$\lim_{K \to \infty} \mu_d^K = \lambda_d$$

is actually fulfilled. We nevertheless believe that this commonly admitted definition of MCSCF excited state energies is the source of various difficulties of both practical and theoretical nature. We now draw up a list of the latter.

i) Practical difficulties in eigenvalue optimization. Solving problem (II.10) amounts to minimizing the d^{th} eigenvalue of a matrix depending on a set of parameters Φ. This is known to be a challenging task. Indeed no completely satisfactory numerical method dedicated to solving such problems is available to date, except for very special cases (for instance when the matrix linearly depends on the parameters, see e.g. [LO96]). In fact, we shall see in the next paragraph that the algorithms which are currently implemented in the Quantum Chemistry simulation packages are not totally adapted to this issue [She87, WM80, CER79, Wer87].

ii) Problems of degeneracies and non-existence of stationary points. Serious difficulties can occur when optimizing $\lambda_d^K(\Phi)$, due to a possible loss of differentiability of this function in case of degeneracies. As an illustration, let us simply mention a celebrated example due to Rellich and reported in [RS78]: consider the family of 2×2 matrices $(A(x,y))_{(x,y)\in\mathbb{R}^2}$ defined by

$$A(x,y) = \begin{pmatrix} -\sin x & \sin y \\ \sin y & \sin x \end{pmatrix}; \tag{II.11}$$

the second eigenvalue $\sqrt{\sin^2(x)+\sin^2(y)}$ degenerates at its minimum $(x_0,y_0)=(0,0)$, and is not differentiable. At this point, there exists no critical point of the associated energy $(x,y,v)\mapsto\langle A(x,y)v,v\rangle$ on $\mathbb{R}^2\times S^1$. Coming back to our main context, if no stationary MCSCF state exists at the energy level μ_d^K, it is not clear whether definition (II.10) is relevant.

iii) The non-linearity of $\mathcal{E}_N^K$ is not fully accounted for. An explanation of the difficulties mentioned above can be given in the following way. According to the Rayleigh-Ritz principle, we may write

$$\mu_d^K=\inf_{\Phi}\inf_{\substack{V\subset\mathbb{R}^{\binom{K}{N}}\\ \dim V=d}}\sup_{\substack{c\in V\\ \|c\|=1}}\mathcal{E}_N^K(c,\Phi)=\inf_{A\in\mathcal{A}^{d-1}}\sup_{(c,\Phi)\in A}\mathcal{E}_N^K(c,\Phi), \tag{II.12}$$

where

$$\mathcal{A}^{d-1}=\left\{f(S^{d-1})\times\{\Phi\},\ f\in C^0(S^{d-1},S^{\binom{K}{N}-1})\text{ odd},\ \Phi\in H^1(\mathbb{R}^3)^K,\ \int\Phi\Phi^T=I_K\right\}.$$

Such an $\inf-\sup$ method cannot give *a priori* a critical point of $\mathcal{E}_N^K$, for the set $\mathcal{A}^{d-1}$ on which the outer infimum in (II.12) is clearly not invariant under *nonlinear* deformations [Gho93]. We propose below a new $\inf-\sup$ definition of MCSCF excited states which fully takes into account the nonlinearity of the model.

But before proceeding further in this direction, we would like to make some comments on the numerical methods used to solve (II.10). Following [She87, WM80, CER79, Wer87], the general form of the numerical algorithms currently used to calculate the $(d-1)^{th}$ excited state can be summarized as follows:

1. Start with some (c,Φ) obtained for instance from a previous Hartree-Fock or Configuration-Interaction calculation.
2. Compute the matrix H_Φ of the quadratic form associated to H_N on the subspace $\mathrm{Span}\,\{\varphi_{i_1}\wedge\cdots\wedge\varphi_{i_N},\ i_1<\cdots<i_N\}$.
3. Find c' as the d^{th} eigenvector of this matrix.
4. This c' being fixed, minimize the energy with respect to Φ to obtain a new Φ'.
5. Replace (c,Φ) by (c',Φ') and return to step 2.

The main difficulty with this method is that the energy is not necessarily decreasing during the computation; it can in fact oscillate, as this can be easily seen when this algorithm is applied to the following toy problem (due to Rellich and reported in [RS78]): find the first excited state for the energy functional

$$\tilde{\mathcal{E}}(c,\Phi)=c^T\begin{pmatrix}-\sin\Phi & 0\\ 0 & \sin\Phi\end{pmatrix}c$$

with $c\in S^1$ and $\Phi\in]-\pi,\pi[$ (an oscillation between $\Phi=-\pi/2$ and $\Phi=\pi/2$ is obtained). This phenomenon is observed in practice in MCSCF calculations (see, e.g., [CER79, page 1092]). This is a severe limitation of the above numerical method.

2.2 A new definition of MCSCF excited states

In this section, we present our new definition of MCSCF excited states and explain how this definition can be used in practice to compute the first excited state.

We introduce the following group action of $\mathbb{Z}_2$ on $\mathcal{M}_N^K$:

$$(-)\cdot(c,\Phi) = (-c,\Phi).$$

A continuous function $f : S^{d-1} \longrightarrow \mathcal{M}_N^K$ is said to be $\mathbb{Z}_2$-*equivariant* [Gho93] if it satisfies $f(-x) = (-)\cdot f(x)$ for all $x \in S^{d-1}$, or equivalently

$$f(x) = (c,\Phi) \Longrightarrow f(-x) = (-c,\Phi).$$

We denote by $\mathcal{C}^{d-1}$ the set of all such continuous $\mathbb{Z}_2$-equivariant functions from S^{d-1} into $\mathcal{M}_N^K$, and introduce

$$\mathcal{B}^{d-1} = \left\{ f(S^{d-1}),\ f \in \mathcal{C}^{d-1} \right\}.$$

Our definition of an excited state energy is now the following

$$\boxed{\lambda_d^K = \inf_{B\in\mathcal{B}^{d-1}} \sup_{(c,\Phi)\in B} \mathcal{E}_N^K(c,\Phi).} \tag{II.13}$$

Notice the important property

$$\mathcal{A}^{d-1} \subset \mathcal{B}^{d-1} \tag{II.14}$$

A direct consequence of (II.14) is the following inequality

$$\lambda_d^K \leq \mu_d^K.$$

Let us now state the following result, established by one of us in [Lew04].

Theorem II.1 (Existence of MCSCF excited states). *Assume $Z > N-1$ and $1 \leq d \leq \binom{K}{N}$. Then there exists a critical point (c_0,Φ_0) of the energy $\mathcal{E}_N^K$ on $\mathcal{M}_N^K$, with a Morse index lower than or equal to $d-1$, and which satisfies $\mathcal{E}_N^K(c_0,\Phi_0) = \lambda_d^K$. Moreover, λ_d^K satisfies*

$$\lambda_d \leq \lambda_d^K \leq \mu_d^K \tag{II.15}$$

and therefore

$$\lim_{K\to\infty} \lambda_d^K = \lambda_d.$$

This result shows that contrarily to what occurs with the definition (II.10), one always obtains with (II.13) a critical point which is solution to the MCSCF equations (II.9), together with a possibly better estimate of the true excited state energy as shown by (II.15). We have no general criterion to decide whether the strict inequality $\lambda_d^K < \mu_d^K$ holds or not. In our simulation (see section 3.3.1 below) we have observed cases where it was true and other cases for which it was not. One can guess that it holds in practice when, due to a problem of degeneracy, no critical point exists at the level μ_d^K (λ_d^K is always a critical value by Theorem II.1). For the Rellich example defined above (II.11), a simple calculation indeed shows that, with obvious notations, $-1 = \lambda_2 < \mu_2 = 0$.

Remark. In a non-interacting system, i.e. when the interaction term

$$\sum_{1\leqslant i<j\leqslant N} \frac{1}{|x_i - x_j|}$$

in the expression (II.1) of H_N is turned off, one can see that for any $d \leq \binom{K}{N}$, $\lambda_d^K = \lambda_d$ (i.e. the MCSCF and the Schrödinger excited state energies coincide). Moreover, the critical point which is found in Theorem II.1 is precisely in this case the d^{th} eigenfunction of the N-body hamiltonian.

Remark. Notice that when $K = N$, $\binom{K}{N} = 1$, which implies that Theorem II.1 *does not provide any definition of excited states for the Hartree-Fock method*. This strengthens the argument that the Hartree-Fock method is not adapted to excited state calculations [HG96]. Let us point out that the infinitely many solutions to the Hartree-Fock equations constructed by P.-L. Lions in [Lio87] cannot be interpreted as excited states of the molecule. Indeed, the sequence $(\nu_d)_{d\geq 1}$ of the critical values of [Lio87] satisfies $\nu_d \to 0$ as $d \to \infty$, whereas $\lambda_d \to \Sigma < 0$ when $N \geq 2$ (recall that Σ is the bottom of the essential spectrum of H_N). Moreover, using the fact that

$$(\cos\theta\varphi_1 + \sin\theta\varphi_2) \wedge (-\sin\theta\varphi_1 + \cos\theta\varphi_2) = \varphi_1 \wedge \varphi_2$$

for all $(\varphi_1, \varphi_2) \in H^1(\mathbb{R}^3)^2$ with $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$ and $\theta \in [0; 2\pi]$, it can be easily seen that $\nu_1 = \nu_2$ for two-electrons systems. Finally, in a non-interacting system, the ν_d do not equal the λ_d, and they satisfy in particular $\nu_{2d-1} \geq 2\lambda_d$ when $N = 2$. The same remarks apply to the sequence constructed in [Lew04, Theorem 2 (i)]. Let us also mention that a definition of Hartree-Fock excited states has been proposed by J.F. Léon in [Léo88]: the first excited state is defined as a minimizer of the Hartree-Fock energy, restricted to the set of Slater determinants which are orthogonal to the Hartree-Fock ground state, and so on for the higher excited states. Note that the so-obtained excited states are not solutions to the Hartree-Fock equations and we therefore do not know whether this definition can be useful in practice.

2.3 A new method for the computation of the first excited state

We can deduce from (II.13) a new computational approach for the calculation of excited states. In this section, we explain the main lines of this algorithm for the case of the first excited state.

Let us emphasize that our definition (II.13) is valid for all the excited states for which $1 \leq d \leq \binom{K}{N}$. However, the exploitation of (II.13) for numerical purposes when $d > 2$ is not obvious (one has to deform surfaces of dimension $d-1$), and will be the subject of a future work. For the first excited state (i.e. for $d = 2$), our definition can be easily transformed into a computational method, that we present here.

Let us first clarify the structure of the set $\mathcal{B}^1$. A $\mathbb{Z}_2$-equivariant function of $\mathcal{C}^1$ can be seen as a function $t \in [0; 2] \mapsto (c(t), \Phi(t)) \in \mathcal{M}_N^K$ which satisfies $c(1+t) = -c(t)$ and $\Phi(1+t) = \Phi(t)$. Since $\mathcal{E}_N^K$ is $\mathbb{Z}_2$-invariant which means

$$\mathcal{E}_N^K(-c, \Phi) = \mathcal{E}_N^K(c, \Phi),$$

we see that

$$\sup_{t\in[0;2]} \mathcal{E}_N^K(c(t), \Phi(t)) = \sup_{t\in[0;1]} \mathcal{E}_N^K(c(t), \Phi(t)).$$

Therefore, we can rewrite (II.13) as

$$\lambda_2^K = \inf_{(c,\Phi)\in\mathcal{M}_N^K} \left\{ \inf_{\gamma\in\Gamma_{(c,\Phi)}} \sup_{t\in[0;1]} \mathcal{E}_N^K(\gamma(t)) \right\} \tag{II.16}$$

where

$$\Gamma_{(c,\Phi)} = \left\{ \gamma \in C^0\left([0;1], \mathcal{M}_N^K\right),\ \gamma(0) = (c,\Phi),\ \gamma(1) = (-c,\Phi) \right\}.$$

Notice that the $\inf - \sup$ problem which is in brackets in (II.16) is a mountain-pass problem (between (c,φ) and $(-c,\varphi)$), similar to those encountered in molecular simulation in the search for transition states between reactants and products on potential energy surfaces [Sch87, QH84]. To compute the term in brackets, one thus has to deform paths, as this is usually done in the latter setting.

We propose to simplify the resolution of problem (II.16): we simply assume that a global minimizer of the MCSCF energy $(\bar{c},\bar{\Phi})$ is is also a minimizer of the outer minimization in (II.16). Therefore, we clamp both ends of the trial paths at $(\bar{c},\bar{\Phi})$ and $(-\bar{c},\bar{\Phi})$ respectively, and solve the mountain pass problem

$$\inf_{\substack{\gamma\in C^0\left([0;1],\mathcal{M}_N^K\right),\\ \gamma(0)=(\bar{c},\bar{\Phi}),\ \gamma(1)=(-\bar{c},\bar{\Phi})}} \sup_{t\in[0;1]} \mathcal{E}_N^K(\gamma(t)). \tag{II.17}$$

This simplication is justified provided there exists a path linking $(\bar{c},\bar{\Phi})$ and an actual minimizer of the outer minimization of (II.16), along which the energy does not exceed λ_2^K. It is indeed likely to be the case, at least for K large enough.

Thus, our algorithm to compute the first excited state can be summarized as follows (we shall give details for each step in the next section):

1. Use a Newton-like method to compute a ground state $(\bar{c},\bar{\Phi})$ of the MCSCF energy.
2. Find the path γ on the manifold $\mathcal{M}_N^K$ such that $\gamma(0) = (\bar{c},\bar{\Phi})$ and $\gamma(1) = (-\bar{c},\bar{\Phi})$, which minimizes
$$\max \mathcal{E}_N^K(\gamma([0;1])).$$
3. The first MCSCF excited state is the state which possesses the highest energy on this optimal path.

In practice, solving the mountain-pass problem (Step 2) is rather demanding in terms of CPU time. Therefore, we shall always choose a not too tight convergence criteria to stop the path optimization step. The state of highest energy on the final path is then used as initial guess in a Newton-like procedure to solve (II.9).

We have found many algorithms in the literature for the optimization of paths (often applied to the simulation of chemical reactions on potential energy surfaces) [Sch87, HJJ00, HJ99, HJU00, HJ00, CMH94, EK87, JS88, QH84, QHIH98, CE90, WWVE02], some of them being quite peculiar in our opinion. The method that we propose below for the deformation of paths, and which seems to give good results on our problem, is of general concern and could therefore also be useful for some other problems.

2.4 Solving the mountain pass problem: a method of deformation of paths

Let us first point out that solving a mountain pass problem is by no means equivalent to finding a saddle point somewhere "between" two minima. The example of the search of the first excited state of the Helium atom (Section 3.3.1) is an illustration of this statement. In this example indeed, the optimal path obtained with our algorithm contains two saddle points of different energies; an algorithm of saddle point localization could converge toward the one of lower energy, and thus underestimate the mountain pass energy.

The best way for properly solving a mountain pass problem is in fact to deform paths. A mathematical study of an algorithm of this type can be found in [CM93, CMR93]. Our method has been inspired by the one described in these references, but it is not identical. In this section, we present it in the following abstract setting: solve the mountain pass problem on the energy surface defined by the functional $\mathcal{E}$ on the Riemann manifold $\mathcal{M}$ between the two points M_0 and M_0' of $\mathcal{M}$, or in other words, find a minimizer of

$$\inf_{\substack{\gamma \in C^0([0;1],\mathcal{M}) \\ \gamma(0)=M_0,\ \gamma(1)=M_0'}} \max_{t\in[0;1]} \mathcal{E}(\gamma(t)).$$

Like in [HJJ00, HJU00, HJ00, CE90, WWVE02], the main idea is to sample a given path linking M_0 and M_0' with a sequence of points $M_0, M_1, \cdots, M_{N+1}$ of $\mathcal{M}$, such that $M_{N+1} = M_0'$. During the optimization process, the number N of points used to represent the current path is not necessarily fixed. In our method, we associate with each sequence $(t_k, M_k)_{0\leq k\leq N+1}$ where $0 = t_0 < t_1 < \cdots < t_N < t_{N+1} = 1$ are real numbers and where $M_0, M_1, \cdots, M_{N+1}$ are points on $\mathcal{M}$, a uniquely defined continuous path $\gamma : [0;1] \to \mathcal{M}$ which satisfies $\gamma(t_k) = M_k$. This is done by selecting once and for all, a convenient interpolation scheme. A possible choice is to take for $\gamma(t)$ some piecewise geodesic curve on the manifold $\mathcal{M}$. Simplest interpolation schemes can also be chosen, for in practice, M_k and M_{k+1} will be close together. In some cases, spline-type interpolation functions can also be used.

A sequence $(t_k, M_k)_{0\leq k\leq N+1}$ being given, one can use the gradient field of the functional $\mathcal{E}$ to deform the associated continuous path. A naive approach consists in simply moving each M_k in the direction opposite to the gradient with a step-length α_k. Remark that since the new point M_k' has to lay on the (curved) manifold $\mathcal{M}$, one has to make precise the statement "in the direction opposite to the gradient". The most intrinsic rule is to move M_k on the geodesic curve which spurts out from M_k in the direction opposite to the gradient. A simpler alternative is first to move M_k in the tangent space, then to project the so-obtained point on the manifold (we shall use this method in our problem).

When this naive procedure is iterated, each point M_k falls off in one of the valleys of the function. In [HJJ00, HJU00, HJ00], it is suggested to circumvent this problem by linking the points $(M_k)_{0\leq k\leq N+1}$ with strings or elastic bands. Our main contribution is to propose a simpler but apparently more efficient solution: it consists in first computing the path γ' associated to (t_k, M_k'), and then finding new points (t_k', N_k) which are better distributed in some sense on the (uniquely defined) continuous path γ'. We have observed that for stability reasons, the points need to be redistributed after each minimization step. We use in addition the following rule: the larger the difference between the maximum of the energy on γ and the $\mathcal{E}(M_k)$, the smallest the step-length α_k in the direction opposite to the gradient. This simple trick helps in preventing the points M_k from falling off in the valleys.

We have applied the above method to several test cases (notably to the ones described in [QHIH98]) and we have observed a convergence to the saddle point in all the cases, when the number N of points is large enough. We have also checked on these test cases that switching to a Newton-like method once the mountain-pass algorithm has found out a state close enough to the saddle point, is an efficient strategy.

In the following section, we apply this method to the calculation of MCSCF excited states.

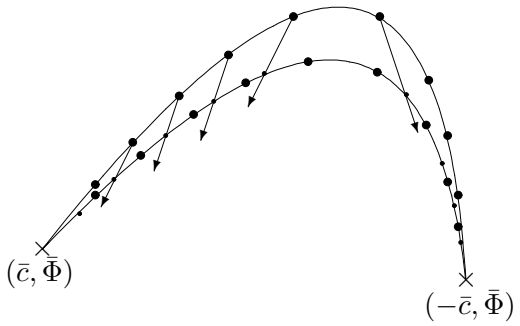


Figure II.1: The deformation method

3 Computation of the first excited state of two-electrons systems

3.1 Singlet and triplet states

In order to be able to simulate real molecular systems, we now need to reintroduce the spin variables. As the N-body hamiltonian H_N and the spin operators S^2 and S_z (see e.g. [LL77]) commute, it is convenient to search for eigenfunctions of H_N that also are eigenfunctions of S^2 and S_z.

For two-electron systems, the situation is particularly simple. There are only two types of wavefunctions which are eigenfunctions of both S^2 and S_z, namely the so-called singlet and triplet states.

A singlet state is a wavefunction of the form

$$\Psi^s(x, \sigma; y, \sigma') = \psi(x, y)|\alpha\beta\rangle(\sigma, \sigma')$$

where $\psi(x, y)$ is *symmetric* in $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$, i.e. such that $\psi(y, x) = \psi(x, y)$. The antisymmetry is carried by the spin function $|\alpha\beta\rangle(\sigma, \sigma')$ which is defined for $(\sigma, \sigma') \in \{|\uparrow\rangle, |\downarrow\rangle\} \times \{|\uparrow\rangle, |\downarrow\rangle\}$ by

$$|\alpha\beta\rangle(\sigma, \sigma')\frac{1}{\sqrt{2}}\left(\alpha(\sigma)\beta(\sigma') - \beta(\sigma)\alpha(\sigma')\right)$$

where

$$\alpha(|\uparrow\rangle) = 1, \qquad \alpha(|\downarrow\rangle) = 0, \qquad \beta(|\uparrow\rangle) = 0, \qquad \beta(|\downarrow\rangle) = 1.$$

A triplet state takes the form

$$\Psi^t(x, \sigma; y, \sigma') = \psi(x, y)\alpha(\sigma)\alpha(\sigma')$$

where $\psi(x, y)$ is *antisymmetric* in $L^2(\mathbb{R}^3 \times \mathbb{R}^3)$, i.e. $\psi(y, x) = -\psi(x, y)$ (the spin function $\alpha(\sigma)\alpha(\sigma')$ is symmetric and the antisymmetry is carried by the function of the space variables).

For two electron systems, the MCSCF wavefunctions thus read

$$\psi = \sum_{1 \le i,j \le K} c_{ij}\varphi_i \otimes \varphi_j, \tag{II.18}$$

where the $K \times K$ matrix $C = (c_{ij})$ is symmetric for singlet states and antisymmetric for triplet states. The condition $\|\psi\|_{L^2} = 1$ also reads $\|C\| = 1$ where $\|C\| = \mathrm{tr}(CC^T)^{1/2}$.

For numerical simulations, one most often resorts to a Galerkin approximation. One expands each φ_i on a finite basis $(\chi_\mu)_{1\leq\mu\leq N_b}$ of $H^1(\mathbb{R}^3)$ functions specially designed for electronic structure calculations, the so-called atomic orbitals. This approximation is refered to as the Linear Combination of Atomic Orbitals (LCAO) approximation in the Computational Chemistry literature (see e.g. [HRSP86]). Let S be the matrix defined by

$$S_{\mu\nu} = \int_{\mathbb{R}^3} \chi_\mu \chi_\nu,$$

and $\varphi = (\varphi_{\mu i})$ the $N_b \times K$ coordinate matrix of the functions $(\varphi_i)_{1\leq i\leq K}$ in the basis $(\chi_\mu)_{1\leq\mu\leq N_b}$. The condition $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$ also reads $\varphi^T S\varphi = I_K$ (of course, N_b must be chosen greater or equal to K) and the energy of a state Ψ^s or Ψ^t, as a function of C and φ, has the following expression:

$$\mathcal{E}(C,\varphi) = 2\ \mathrm{tr}(C^T\varphi^T h\varphi C) + \mathrm{tr}\left((\varphi C\varphi^T)\mathcal{W}(\varphi C\varphi^T)\right), \tag{II.19}$$

where h is the $N_b \times N_b$ matrix defined by

$$h_{\mu\nu} = \frac{1}{2}\int_{\mathbb{R}^3} \nabla\chi_\mu \cdot \nabla\chi_\nu + \int_{\mathbb{R}^3} V\chi_\mu\chi_\nu$$

and where $\mathcal{W}$ is the linear map associated with the tensor W defined by

$$W_{\mu\nu\kappa\lambda} = \int_{\mathbb{R}^3\times\mathbb{R}^3} \frac{\chi_\mu(x)\chi_\nu(y)\chi_\kappa(x)\chi_\lambda(y)}{|x-y|}\,dx\,dy,$$

i.e. for any $N_b \times N_b$ matrix X

$$[\mathcal{W}(X)]_{\mu\nu} = \sum_{\kappa,\lambda=1}^{N_b} W_{\mu\nu\kappa\lambda}X_{\kappa\lambda}. \tag{II.20}$$

Remark that expression (II.19) is valid for both singlet and triplet states, but that the matrix C appearing in this formula is symmetric for singlet states and antisymmetric for triplet states.

Let us first deal with the singlet state case. The manifold of admissible singlet states is

$$\mathcal{M} = \left\{(C,\varphi) \in M(K,K)\times M(N_b,K),\ C^T = C,\ \mathrm{tr}(C^TC) = 1,\ \varphi^T S\varphi = I_K\right\},$$

where $M(K,K')$ denotes the set of $K \times K'$ matrices. The MCSCF equations (i.e. the stationarity conditions of the MCSCF energy (II.19) on the manifold $\mathcal{M}$) take the form

$$\begin{cases} \left(\varphi^T h\varphi C + C^T\varphi^T h\varphi\right) + \varphi^T\mathcal{W}(\varphi C\varphi^T)\varphi = \beta C \\ h\varphi CC^T + \mathcal{W}(\varphi C\varphi^T)\,\varphi C = S\varphi\Lambda \end{cases} \tag{II.21}$$

where $\beta \in \mathbb{R}$ and where Λ is a $K\times K$ symmetric matrix. The problem can be dramatically simplified by introducing the following reduced model (see [Fri03, And63, Col63, CY00] and [Lew04, Appendix]). Let us first remark that (II.18) can be recast as

$$\psi(x,y) = \Phi(y)^T C\Phi(x)$$

where $\Phi = (\varphi_1, ..., \varphi_K)$. Since C is symmetric and real, there exists a rotation matrix $U \in M(K,K)$ such that $C = U^T \mathrm{diag}(c_1, ..., c_K) U$. This means that if we introduce $\Phi' = U\Phi$, we obtain

$$\psi(x,y) = \sum_{i=1}^{K} c_i \varphi'_i(x) \varphi'_i(y).$$

Therefore, there is no restriction in assuming that the matrix C is diagonal. When using this reduced model (which is *not* an approximation) in a finite basis of atomic orbitals, the manifold of admissible singlet states reads

$$\mathcal{M}^{\mathrm{red}} = S^{K-1} \times \mathcal{W}_K^{N_b}$$

with

$$\mathcal{W}_K^{N_b} = \left\{ \varphi \in M(N_b, K),\ \varphi^T S \varphi = I_K \right\},$$

and the energy functional is given by

$$\mathcal{E}^{\mathrm{red}}(c, \varphi) = \mathcal{E}(C^s(c), \varphi)$$

where $C^s(c) = \mathrm{diag}(c_1, ..., c_K)$. Lastly, the MCSCF equations become

$$\left\{ \begin{array}{l} H(\varphi) \cdot c = \beta \cdot c \\ h\varphi (C^s(c))^2 + \mathcal{W}(\varphi C^s(c) \varphi^T)\, \varphi C^s(c) = S\varphi\Lambda \end{array} \right. \tag{II.22}$$

with

$$H(\varphi)_{ij} = 2\varphi_i^T h \varphi_i \delta_{ij} + \varphi_i^T \mathcal{W}(\varphi_j \varphi_j^T) \varphi_i.$$

The same type of reduction can be done for triplet states, since for any antisymmetric matrix C, there exists a rotation matrix U such that $C = U^T \mathrm{diag}(C_1, ..., C_p) U$ if $K = 2p$, and $C = U^T \mathrm{diag}(C_1, ..., C_p, 0) U$ if $K = 2p+1$, where

$$C_i = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & c_i \\ -c_i & 0 \end{pmatrix}.$$

This means that the wavefunction of a generic triplet state reads [Col63, Fri03, Lew04]

$$\psi(x,y) = \sum_{i=1}^{p} c_i |\varphi_{2i-1}\varphi_{2i}\rangle(x,y) = \frac{1}{\sqrt{2}} \sum_{i=1}^{p} c_i \left(\varphi_{2i-1}(x)\varphi_{2i}(y) - \varphi_{2i}(x)\varphi_{2i-1}(y) \right).$$

3.2 Description of the algorithm

For the sake of brevity, we only deal with the reduced model of the singlet state.

We shall make use of the following interpolation rule: a discrete path on $\mathcal{M}^{\mathrm{red}} = S^{K-1} \times \mathcal{W}_K^{N_b}$ being given as a finite sequence $\left(t_k, (c^k, \varphi^k) \right)_{0 \le k \le N+1}$ where

- $t_0 = 0 < t_1 < \cdots < t_N < t_{N+1} = 1$ are real numbers ;
- $(c^k, \varphi^k)_{0 \le k \le N+1}$ are points of $S^{K-1} \times \mathcal{W}_K^{N_b}$,

we define the associated continuous path $\gamma \in C^0([0,1], S^{K-1} \times \mathcal{W}_K^{N_b})$ according to

$$\forall t \in [0,1], \quad \gamma(t) = (c(t), \varphi(t))$$

where

$$\forall 0 \leq k \leq N, \qquad \forall t \in [t_k, t_{k+1}], \qquad c(t) = \cos\left(\theta_k(t)\right) c^k + \sin\left(\theta_k(t)\right) \widetilde{c}^{k+1}$$

with

$$\begin{cases} \widetilde{c}^{k+1} = \dfrac{c^{k+1} - \left(c^{k+1}, c^k\right)\, c^k}{\| c^{k+1} - \left(c^{k+1}, c^k\right)\, c^k \|} \\ \\ \theta_k(t) = \dfrac{t - t_k}{t_{k+1} - t_k} \arccos\left(c^{k+1}, c^k\right). \end{cases}$$

and

$$\forall 0 \leq k \leq N, \qquad \forall t \in [t_k, t_{k+1}], \qquad \varphi(t) = \widetilde{\varphi}(t)\, \left[\widetilde{\varphi}(t)^T S \widetilde{\varphi}(t)\right]^{-1/2}$$

with

$$\widetilde{\varphi}(t) = \varphi^k + \frac{t - t_k}{t_{k+1} - t_k} \left(\varphi^{k+1} - \varphi^k\right).$$

We can now describe our algorithm for computing the first excited state of two-electron systems.

Step A: search for a MCSCF ground state $(\bar{c}, \bar{\varphi})$, i.e. solve

$$\inf \left\{ \mathcal{E}^{\mathrm{red}}(c, \varphi), \quad (c, \varphi) \in S^{K-1} \times \mathcal{W}_K^{N_b} \right\}.$$

with a Newton-like algorithm; a convenient initial guess is the Hartree-Fock ground state, which can itself be obtained by a self-consistent field algorithm [CDK+03].

Step B : construction of an initial trial path.

As already mentioned in Section 2.3, we get rid of the outer minimization in (II.16) and concentrate on solving

$$\lambda_2^{s,r} = \min_{\substack{\gamma \in \mathcal{C}^0([0,1], S^{K-1} \times \mathcal{W}_K^{N_b}) \\ \gamma(0) = (\bar{c}, \bar{\varphi}) \\ \gamma(1) = (-\bar{c}, \bar{\varphi})}} \max_{t \in [0;1]} \mathcal{E}^{\mathrm{red}}(\gamma(t)).$$

Let $\bar{c}_1$ be the second eigenvector of the hamiltonian matrix

$$[H(\bar{\varphi})]_{ij} = 2\, \bar{\varphi}_i^T h \bar{\varphi}_i\, \delta_{ij} + \bar{\varphi}_i^T \mathcal{W}(\bar{\varphi}_j \bar{\varphi}_j^T) \bar{\varphi}_i \tag{II.23}$$

(note that $\bar{c}$ is the ground state of $H(\bar{\varphi})$). A possible initial trial path is one of the minimizers of

$$\min_{\substack{\gamma(t) = (c(t), \bar{\varphi}) \\ c(t) \in \mathcal{C}^0([0,1], S^{K-1}) \\ \gamma(0) = (c, \bar{\varphi}) \\ \gamma(1) = (-c, \bar{\varphi})}} \max_{t \in [0;1]} \mathcal{E}^{\mathrm{red}}(\gamma(t)),$$

for instance

$$\gamma^0(t) = (c(t), \bar{\varphi}) \tag{II.24}$$

with $c(t) = \cos(t\pi)\,\bar{c} + \sin(t\pi)\,\bar{c}'$. A better initial guess can however be obtained by random perturbations of that reference path. In practice, we randomly choose a collection of $N_{\rm sto}$ states $(\bar{c}'_j, \bar{\varphi}'_j) \in \left(\mathrm{vect}(\bar{c})^\perp \cap S^{K-1}\right) \times \mathcal{W}_K^{N_b}$ such that for all j

$$\|\bar{c}'_j - \bar{c}_1\| \leq \varepsilon \|\bar{c}_1\| \quad \text{and} \quad \|\bar{\varphi}'_j - \bar{\varphi}\| \leq \varepsilon \|\bar{\varphi}\|$$

for a small ε, and we consider the $N_{\rm sto}$ continuous paths $\gamma^j(t)$ associated with the three-point discrete paths

$$\gamma^j(0) = (\bar{c}, \bar{\varphi}),\ \gamma^j(1/2) = (\bar{c}'_j, \bar{\varphi}'_j),\ \gamma^j(1) = (-\bar{c}, \bar{\varphi}).$$

One then selects, among the $N_{\rm sto}$ paths γ^j, the path $\gamma_0(t)$ for which the maximal energy $\max \mathcal{E}^{\rm red}(\gamma^j([0;1]))$ is minimum. The above method can obviously be generalized to discrete paths containing more than three points and can also be used to improve the following step C (path optimization) when necessary.

One then sets $m = 0$, $t_k = k/(N+1)$, $\gamma_0^k = (c^{k,0}, \varphi^{k,0}) = \gamma_0(t_k)$ for $1 \leq k \leq N$, and

$$E_0^{\min} = \mathcal{E}^{\rm red}(\bar{c}, \bar{\varphi}), \qquad E_0^{\max} = \max_{1 \leq k \leq N} \mathcal{E}^{\rm red}\left(\gamma_0^k\right),$$

$$i_0^{\max} = \operatorname*{argmax}_{1 \leq k \leq N} \mathcal{E}^{\rm red}\left(\gamma_0^k\right).$$

Step C : path optimization.

For the sake of simplicity, we displace the nodes in the direction opposite to the gradient; for this purpose,

1. we compute for each $1 \leq k \leq N$, the MCSCF energy at the point $\gamma_m^k = (c^{k,m}, \varphi^{k,m})$

$$E_m^k = \mathcal{E}^{\rm red}(\gamma_m^k) = \left(H(\varphi^{k,m})c^{k,m}, c^{k,m}\right)\ ;$$

2. we project the components $\nabla_c \mathcal{E}^{\rm red}(\gamma_m^k)$ and $\nabla_\varphi^S \mathcal{E}^{\rm red}(\gamma_m^k)$ of the gradients of the energy at the points γ_m^k on the tangent spaces of the underlying manifolds:

$$\begin{aligned} g_m^k &= \nabla_c \mathcal{E}^{\rm red}(c^{k,m}, \varphi^{k,m}) - \left(\nabla_c \mathcal{E}^{\rm red}(c^{k,m}, \varphi^{k,m}), c^{k,m}\right)\, c^{k,m} \\ &= 2\,\left(H(\varphi^{k,m}) - E_m^k\right) c^{k,m}\ ; \end{aligned}$$

$$G_m^k = 4\, P^S_{\varphi^{k,m}} \left[S^{-1}\left(h\varphi^{k,m} C^s(c^{k,m})^2 + \mathcal{W}\left(\varphi^{k,m} C^s(c^{k,m})(\varphi^{k,m})^T\right)\varphi^{k,m} C^s(c^{k,m})\right)\right],$$

where $P^S_{\varphi^{k,m}}$ is the orthogonal projector of $M(N_b, K)$ on the tangent space $T_{\varphi^{k,m}}\mathcal{W}_K^{N_b}$ for the scalar product $\langle \cdot, \cdot \rangle_S$ (defined on $M(N_b, K)$ by $\langle A, B \rangle_S = \mathrm{tr}(A^T S B)$). One has

$$\forall \varphi \in \mathcal{W}_K^{N_b}, \quad \forall Z \in M(N_b, K), \quad P^S_\varphi Z = Z - \frac{1}{2}\varphi\left(\varphi^T S Z + Z^T S \varphi\right); \qquad \text{(II.25)}$$

3. we set

$$E_m^{\max} = \max_{1 \leq k \leq N} \mathcal{E}^{\rm red}\left(\gamma_m^k\right), \qquad i_m^{\max} = \operatorname*{argmax}_{1 \leq k \leq N} \mathcal{E}^{\rm red}\left(\gamma_m^k\right)\ ;$$

4. if $\|E_m^{\max} - E_{m-1}^{\max}\| < \eta$, M times consecutively, set $(\tilde{c}, \tilde{\varphi}) = \gamma_m^{i_m^{\max}}$ and go to Step D (i.e. switch to a Newton-like algorithm);

5. for all $1 \leq k \leq N$, search for an optimal step $0 \leq \alpha_m^k \leq 1$ and set

$$d_m^k = -\alpha_m^k g_m^k, \qquad D_m^k = -\alpha_m^k G_m^k \ ;$$

6. displace the nodes γ_m^k along the descent directions:

$$c^{k,m+1} = \frac{1}{\|c^{k,m} + d_m^k\|}(c^{k,m} + d_m^k)$$

$$\varphi^{k,m+1} = \widetilde{\varphi}\left[\widetilde{\varphi}^T S \widetilde{\varphi}\right]^{-1/2} \qquad \text{with} \qquad \widetilde{\varphi} = \varphi^{k,m+1} + D_m^k \ ;$$

7. reparametrize the path. For this purpose, define the length of the discrete path by

$$L = \sum_{k=1}^{N+1} \|c^{k,m+1} - c^{k-1,m+1}\| + \|\varphi^{k,m+1} - \varphi^{k-1,m+1}\|$$

and search for a sequence of points $\gamma_{m+1}(t_k)$ of the continuous path $\gamma^{k+1}(t)$ satisfying

$$\|\gamma_{m+1}(t_{k+1}) - \gamma_{m+1}(t_k)\| \in [L/N, L/N + \varepsilon']$$

where ε' is a small enough threshold. The points $\gamma_{m+1}(t_k)$ define a reparametrized path $\gamma_{m+1}^{\text{rep}}(t)$; we finally set

$$\gamma_{m+1}^k = \gamma_{m+1}^{\text{rep}}(k/N);$$

8. set $m = m + 1$ and return to step C.1.

Step D : apply a Newton-like algorithm with $(\tilde{c}, \tilde{\varphi})$ as initial guess.

3.3 Numerical results

In this section, we present numerical results concerning the calculation of the first singlet excited state of some two-electrons systems, namely the H_2 molecule and Helium-like atoms.

These results have been obtained with a Scilab [GBC+99] program, interfaced with a few C routines aiming in particular at speeding up the tensor-matrix products (II.20). Let us mention that the overlap matrices S, the core hamiltonians h, and the bielectronic integral tensors W have been extracted from Gaussian 98 calculations [FTS+98].

3.3.1 The H_2 molecule

We have computed the first singlet excited state of the H_2 molecule, for various interatomic distances R. We have used for these calculations the double zeta Dunning's correlation consistent atomic basis set (CC-PVDZ), for which $N_b = 10$. The number of iterations in Step C (path optimization) necessary to reach a given convergence criterion strongly depends on the choice of the initial guess. In that respect, the randomly perturbed initial paths constucted in Step B of our algorithm are of better quality than the one given by

formula (II.24). Let us mention that one iteration of the path optimization procedure for $N_b = 10$ and $N = 100$ is about five seconds long on a Pentium IV 2.4 GHz.

Let us first analyze the results for a fixed interatomic distance equal to $R = 0.5\,\mathring{\mathrm{A}}$. The energy profiles of the successive paths generated by the path optimization procedure (Step C) have been reported on the same graph (Figure II.2). One can see that the energy profile of the initial trial path is a single hump and that the energy profiles of the iterates progressively turn into a double hump shape. The energy profiles of the earlier iterates have a rough shape for the initial trial path results from a stochastic local deformation of a reference path (Step B of the algorithm). The optimization process rapidly smoothes the trial path. Notice that due to the reparametrization procedure, the graph of $\mathcal{E}^{\mathrm{red}}(\gamma_k)$ is not necessarily below the graph of $\mathcal{E}^{\mathrm{red}}(\gamma_{k-1})$.

The optimal path γ obtained with our algorithm exhibits a double hump energy profile (see Figure II.3) with two local maxima. Let us point out that we have run on this case many tests with different stochastic initial trial paths; we have always obtained a double hump profile at convergence.

Our method thus provides two saddle points of Morse index equal to one. According to our definition of MCSCF excited states, the one of higher energy (denoted by M on Figure II.3) corresponds to the first singlet excited state. On the other hand, the saddle point M' cannot be interpreted as a MCSCF excited state; it is rather a spurious solution of the MCSCF equations (II.22). Notice that the energy of M' corresponds to the *first* eigenvalue of the matrix $H(\varphi)$ appearing in equation (II.22), whereas the energy of M corresponds to the *second* eigenvalue of $H(\varphi)$. We therefore obtain $\lambda_2^K = \mu_2^K$ in this case.

Let us now display the vectors c for the three points I, M and M' introduced in Figure II.3 (of course $c_F = -c_I$)

$$c_I = \begin{bmatrix} 0.9953010 \\ -0.0369945 \\ -0.0087985 \\ -0.0369945 \\ -0.0441540 \\ -0.0441540 \\ -0.0498874 \\ -0.0088056 \\ -0.0088057 \\ -0.0042793 \end{bmatrix} \qquad c_M = \begin{bmatrix} 0.7009818 \\ 0.7009818 \\ -0.0139024 \\ -0.0536869 \\ -0.0458169 \\ -0.0536868 \\ -0.0942362 \\ -0.0141266 \\ -0.0074823 \\ -0.0074823 \end{bmatrix} \qquad c_{M'} = \begin{bmatrix} -0.7069785 \\ 0.7072304 \\ -1.029\ 10^{-08} \\ -0.0021659 \\ -1.159\ 10^{-08} \\ -1.321\ 10^{-08} \\ 0.0009807 \\ -8.252\ 10^{-09} \\ 0.0009807 \\ -8.252\ 10^{-09} \end{bmatrix}$$

It appears that c_M possesses two dominant coefficients. This shows the usefulness of the MCSCF method for the calculation of excited states: the Hartree-Fock method is not able to correctly describe such a two-configuration state (recall that the square of the coefficients of c are the weights of the different configurations of the multiconfiguration wavefunction). The structure of $c_{M'}$ is very similar to the one of c_M, apart from a change of sign.

We can have an insight into the shape of the optimal path γ by displaying the projection of γ on several 3D representative subspaces selected by principal components analysis. Figure II.4 shows the projection of the component $c(t)$ of $\gamma(t)$ as well as that of three different columns of $\varphi(t)$ (each of them representing a partially occupied one-electron wavefunction). The approximate location of the points I, M, M' and F introduced in Figure II.3 are reported on the graphs.

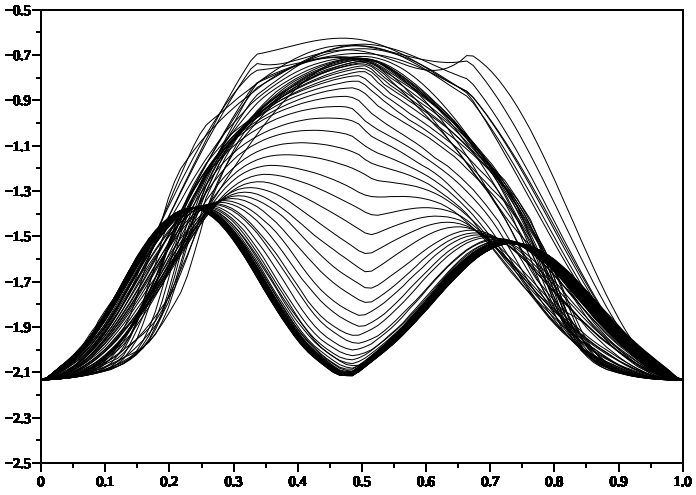


Figure II.2: Energy profiles of the successive paths generated by the path optimization procedure (H_2 molecule, interatomic distance equal to 0.5 $\mathring{\text{A}}$).

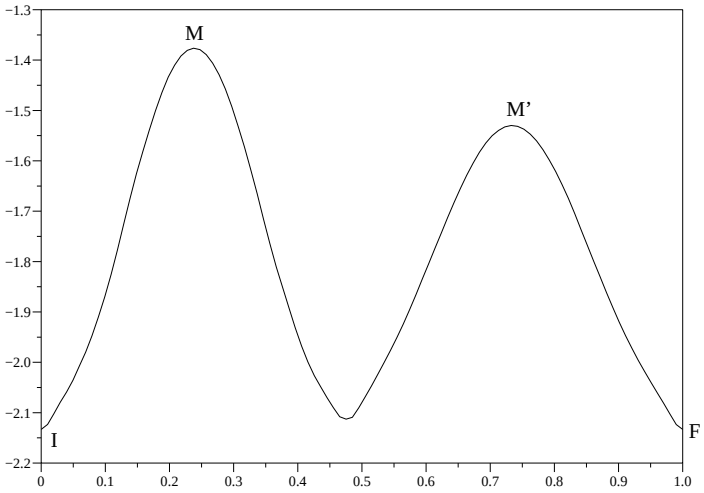


Figure II.3: Path at convergence for H_2 with $R = 0.5$ $\mathring{\text{A}}$.

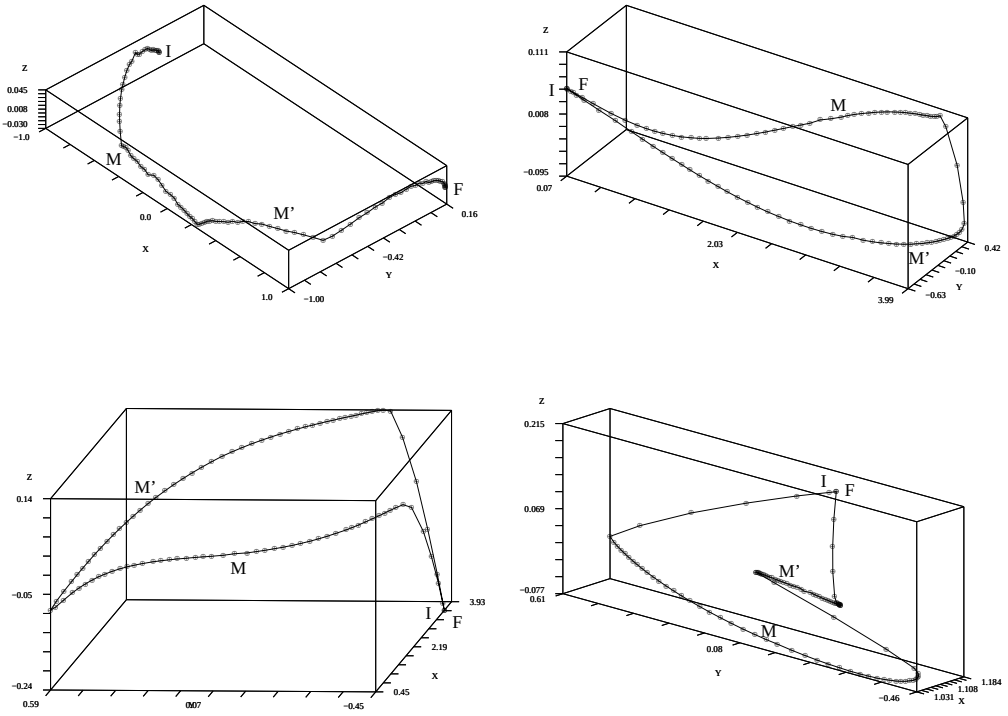


Figure II.4: Projections of $c(t)$ (top left), φ_1 (top right), φ_2 (bottom left) and φ_7 (bottom right) on 3D representative subspaces selected by principal components analysis.

Let us now present our results for different values of the interatomic distance R. Figure II.5 puts together three potential energy surfaces (PES) of H_2: the Hartree-Fock and MCSCF ground state PES and the first MCSCF singlet excited state PES.

For values of R in the range $[1.5\ \mathring{\text{A}}; 2.5\ \mathring{\text{A}}]$, the optimal path exhibits the same characteristics as for the case $R = 0.5\ \mathring{\text{A}}$ reported above. However, the optimal path is more difficult to obtain than for smaller values of R. We have actually observed that in this range of values of R, the choice of the convergence criteria plays a crucial role in the quality of the results. Indeed, the difference $\| \max(\mathcal{E}^{\text{red}}(\gamma_k)) - \max(\mathcal{E}^{\text{red}}(\gamma_{k+1})\|$ can be very small during many consecutive iterations, just as if convergence was reached. But if we run many additional iterations, the algorithm finally escapes this trap and converges toward the (supposed) optimal path. We have observed that such a sequence of small changes in $\max(\mathcal{E}^{\text{red}}(\gamma_k))$ occurs when the energy profile of the trial path turns from a single hump shape into a double hump shape (see Figure II.6).

Many solutions can be proposed to avoid this drawback. One of them consists in using a continuation method. Actually, there are two ways for "propagating" an optimal path by a continuation method. The first way is to choose for the continuation parameter, the interatomic distance R (i.e. inject the optimal path obtained for $R = R_0$ as initial guess in the calculation for $R = R_0 + dR$). The second way is to introduce an additional parameter s, varying in the range $[0; 1]$, allowing to progressively switch on the electronic interaction term. For instance, one can simply multiply by s the second term of the right-hand side of (II.19). When $s = 0$, the two electrons are independent and an optimal path is easy to construct. We have observed that an efficient numerical strategy consists in making the continuation parameter s increase logarithmically from 0 to 0.1, then linearly from 0.1 to 1. Numerical simulations also show that it is more efficient to progressively turn on the electronic interaction than to use the interatomic distance as a continuation parameter. The former method has thus been used in the calculation of the excited state PES of the H_2 molecule for values of R in the range $[1.5; 2.5]$.

For $R = 1.5$ and $R = 2$, the first singlet excited state that we obtain with our method corresponds to the *first eigenvalue* of the matrix H_Φ (see formula (II.22). This means that, using the notations of Sections 2.1 and 2.2, we have in these cases

$$\lambda_2^K < \mu_2^K.$$

Therefore, our definition (II.13) of MCSCF excited states provides better results than the usual definition (II.10).

We notice that when R grows, the values of the local maxima M and M' get closer and closer, until some critical value of R (between 2 and 2.5) for which both values are equal. Beyond this critical value, we obtain, depending on the initial guess, optimal paths that have either one or two humps and give the same excited state (see Figure II.7).

Our numerical experiments thus show that there is a bifurcation phenomenon between single and double hump paths, in the vicinity of the minimum of the first excited state PES. It apparently corresponds to situations for which our method gives better results than the other methods currently used in Computational Chemistry.

3.3.2 Helium-like atoms

We have also tested our algorithm on the calculation of the first singlet excited state of Helium-like atoms (one nucleus of charge $Z \geq 2$ and two electrons). For these calculations,

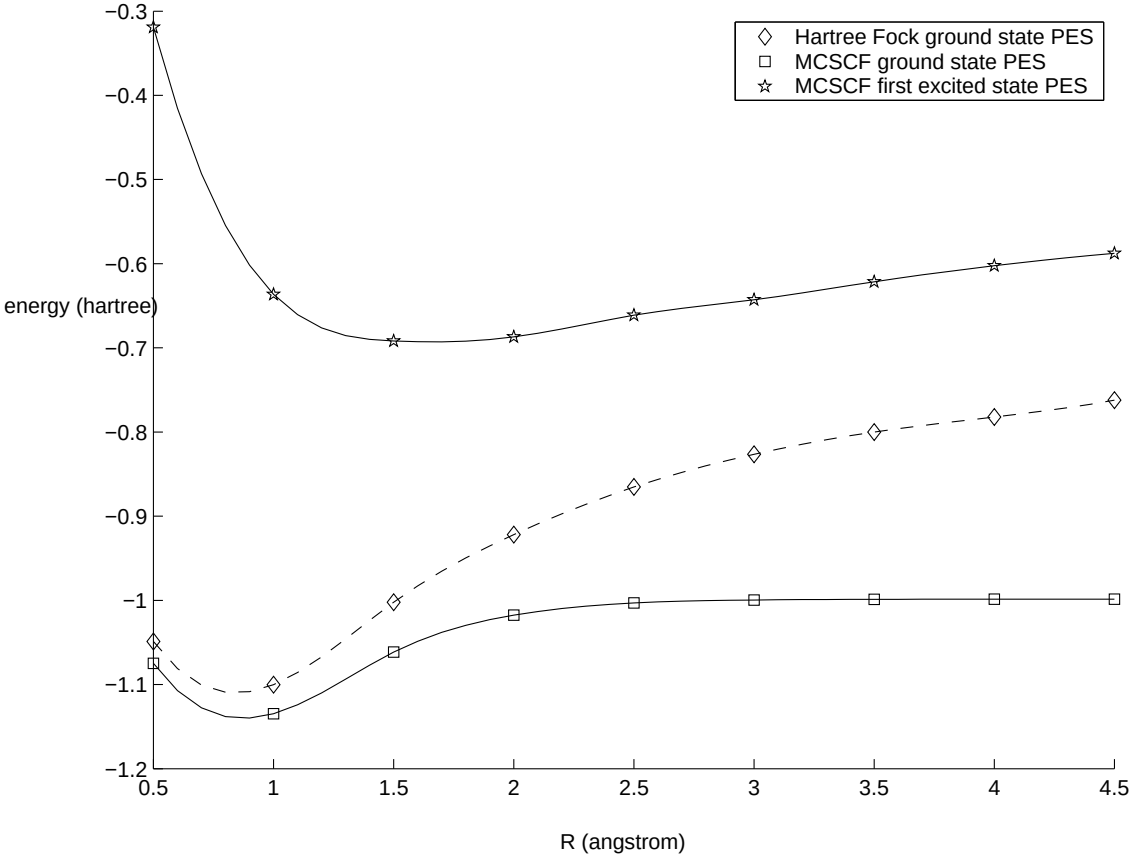


Figure II.5: Potential energy surfaces (PES) of the H_2 molecule.

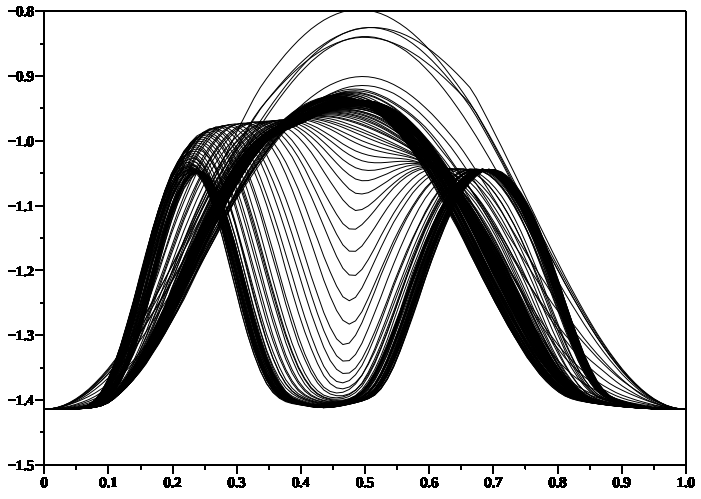


Figure II.6: Iterates of γ_k during the optimization process until convergence, for $R = 1.5$ $\mathring{\text{A}}$.

we have used a simple basis set made of the first six radial eigenstates of the corresponding Hydrogen-like atom (one nucleus of the same charge Z and one electron), that are known analytically (see e.g. [LL77]). The results are shown on Table II.1 for the values $Z = 2, 3, 4$. It is to be noticed that the energy profiles of the optimal paths present double humps. This is consistent with the results obtained with the H_2 molecule for small interatomic distances. The energy profiles of the sequence of paths obtained during the optimization process are put together on Figure II.8.

	$Z = 2$	$Z = 3$	$Z = 4$
HF ground state energy	-2.835374	-7.200180	-13.570185
MCSCF ground state energy	-2.844340	-7.204910	-13.573562
MCSCF first singlet excited state	-1.741472	-4.4432095	-8.4031421

Table II.1: Results for Helium-like atoms (one nucleus of charge Z and two electrons).

Acknowledgment: M.L. would like to thank Éric Séré for useful discussions and for suggesting him the works [CM93, CMR93].

Bibliography

[AF97] P.W. Atkins and R.S. Friedman. *Molecular Quantum Mechanics*. Oxford University Press, third edition, 1997.

[And63] T. Ando. Properties of fermions density matrices. *Rev. Mod. Phys.*, 35(3):690–702, 1963.

[CDK+03] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, and Y. Maday. *Computational quantum chemistry: a primer*, volume X of *Handbook of Numerical Analysis*, pages 3–270. Elsevier, 2003.

[CE90] R. Czerminski and R. Elber. Self-avoiding walk between two fixed points as a tool to calculate reaction paths in large molecular systems. *Int. J. Quantum Chemistry*, 24:167–185, 1990.

[CER79] L.M. Cheung, S.T. Elbert, and K. Ruedenberg. MCSCF optimization through combined use of natural orbital and the Brillouin-Levy-Berthier theorem. *Int. J. Quantum Chem.*, 16:1069–1101, 1979.

[CM93] Y.S. Choi and P.J. McKenna. A mountain pass method for the numerical solution of semilinear elliptic problems. *Nonlinear Anal.*, 20(4):417–437, 1993.

[CMH94] S.S-L. Chiu, J.W. McDouall, and I.H. Hiller. Prediction of whole reaction paths for large molecular systems. *J. Chem. Soc. Faraday Trans.*, 90(12):1575–1579, 1994.

[CMR93] Y.S. Choi, P.J. McKenna, and M. Romano. A mountain pass method for the numerical solution of semilinear wave equations. *Numer. Math.*, 64(4):487–509, 1993.

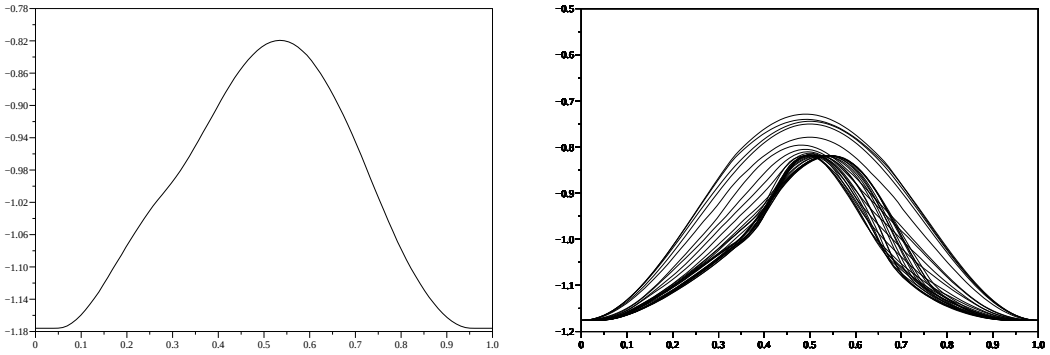


Figure II.7: Path at convergence and iterates of the path during the optimization process for H_2 with $R = 3$ Å.

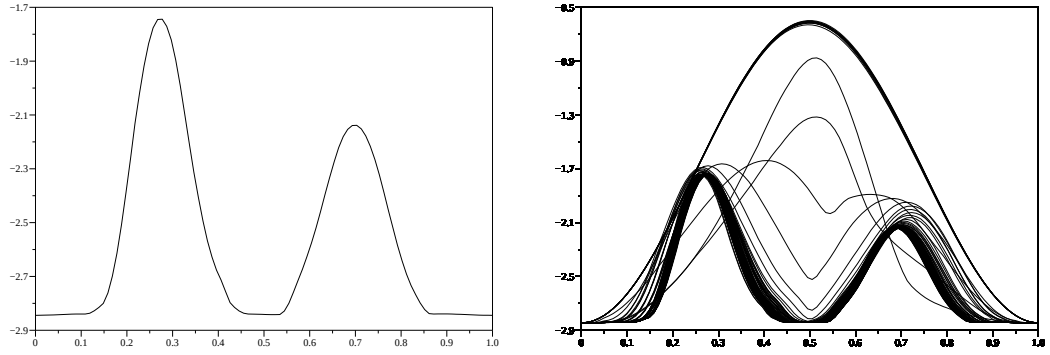


Figure II.8: Optimal path and iterates of the path during the optimization process for the Helium atom.

[Col63] A.J. Coleman. Structure of fermions density matrices. *Rev. Mod. Phys.*, 35(3):668–689, 1963.

[CY00] A.J. Coleman and V.I. Yukalov. *Reduced Density Matrices: Coulson's Challenge*. Springer Verlag, 2000.

[DG90] R.M. Dreizler and E.K.U. Gross. *Density functional theory*. Springer Verlag, 1990.

[EK87] R. Elber and M. Karplus. A method for determining reaction paths in large molecules: application to myoglobin. *Chem. Phys. Letters*, 139(5):375–380, 1987.

[ER81] R.H.A. Eade and M.A. Robb. Direct minimization in MCSCF theory. The quasi-Newton method. *Chem. Phys. Letters*, 83(2):362–368, 1981.

[Fri03] G. Friesecke. The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions. *Arch. Rat. Mech. Analysis*, 169:35–71, 2003.

[FTS+98] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. liu, A. Liashenko, P. Piskorz, I. Kpmaromi, G. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, and J.A. Pople. *Gaussian 98 (Revision A.7)*. Gaussian Inc., Pittsburgh PA, 1998.

[GBC+99] C. Gomez, C. Bunks, J.P. Chancelier, F. Delebecque, M. Goursat, R. Nikoukhah, and S. Steer. *Engineering and Scientific Computing with Scilab*. Birkauser, 1999.

[Gho93] N. Ghoussoub. *Duality and perturbation methods in critical point theory*. Cambridge University Press, 1993.

[HG96] M. Head-Gordon. Quantum chemistry and molecular processes. *J. Phys. Chem.*, 100:13213–13225, 1996.

[HJ99] G. Henkelman and H. Jónsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Physics*, 111(15):7010–7022, 1999.

[HJ00] G. Henkelman and H. Jónsson. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Physics*, 113(22):9978–9985, 2000.

[HJJ00] G. Henkelman, G. Jóhannesson, and H. Jónsson. *Methods for Finding Saddle Points and Minimum Energy Paths*. Ed. S. D. Schwartz (Kluwer Academic Publishers), 2000.

[HJU00] G. Henkelman, H. Jónsson, and B.P. Uberuaga. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Physics*, 113(22):9901–9904, 2000.

[HRSP86] W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople. *Ab initio molecular orbital theory*. Wiley, 1986.

[JS88] P.G. Jasien and R. Shepard. A general polyatomic potential energy surface fitting method. *Int. J. Quantum Chemistry*, 22:183–198, 1988.

[KS65] W. Kohn and L.J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[LB94] C. Le Bris. A general approach for multiconfiguration methods in quantum molecular chemistry. *Ann. Inst. H. Poincaré Anal. Non linéaire*, 11(6):441–484, 1994.

[Léo88] J.F. Léon. Excited states for Coulomb systems in the Hartree-Fock approximation. *Comm. Math. Phys.*, 120(2):261–268, 1988.

[Lew04] M. Lewin. Solutions of the multiconfiguration equations in quantum chemistry. *Arch. Rat. Mech. Anal.*, 171(1):83–114, 2004.

[Lio87] P.L. Lions. Solutions of Hartree-Fock equations for Coulomb systems. *Commun. Math. Phys.*, 109:33–87, 1987.

[LL77] L.D. Landau and E.H. Lifchitz. *Quantum mechanics*. Pergamon Press, 1977.

[LO96] A.S. Lewis and M.L. Overton. Eigenvalue optimization. *Acta Num.*, 5:149–190, 1996.

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

[Löw55a] 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):1474–1489, 1955.

[Löw55b] P.O. Löwdin. Quantum theory of many-particle systems. II. Study of the ordinary Hartree-Fock approximation. *Phys. Rev.*, 97(6):1490–1508, 1955.

[Löw55c] P.O. Löwdin. Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects. *Phys. Rev.*, 97(6):1509–1520, 1955.

[Löw59] P.O. Löwdin. Correlation problem in many-electron quantum mechanics. I. Review of different approaches and discussion of some current ideas. *Adv. Chem. Phys.*, 2:207–322, 1959.

[MFS92] M.A. Robb M. Frisch, I.N. Ragazos and H.B. Schlegel. An evaluation of three direct mcscf procedures. *Chem. Phys. Letters*, 189(6):524–528, 1992.

[QH84] W. Quapp and D. Heidrich. Analysis of the concept of minimum energy path on the potential energy surface of chemically reactiong systems. *Theoret. Chim. Acta*, 66:245–260, 1984.

[QHIH98] W. Quapp, H. Hirsch, O. Imig, and D. Heidrich. Searching for saddle points of potential energy surfaces by following a reduced gradient. *J. Comput. Chem.*, 19(9):1087–1100, 1998.

[Roo87] B. O. Roos. The Complete Active Space Self-Consistent Field method and its applications in electronic structure calculation. *Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys.*, 69:399–446, 1987.

[RS78] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, volume IV – Analysis of Operators. Academic Press, 1978.

[Sch87] H.B. Schlegel. *Adv. Chem. Phys.*, 67:249, 1987.

[She87] R. Shepard. The Multiconfiguration Self-Consistent Field method. *Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys.*, 69:63–200, 1987.

[Wer87] H.-J. Werner. Matrix-formulated direct Multiconfiguration Self-Consistent Field and Multiconfiguration Reference Configuration-Interaction methods. *Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys.*, 69:1–62, 1987.

[WM80] H.-J. Werner and W. Meyer. A quadratically convergent Multiconfiguration-Self-Consistent Field method with simultaneous optimization of orbitals and CI coefficients. *J. Chem. Phys.*, 73(5):2342–2356, 1980.

[WWVE02] E. Weinan, R. Weiqing, and E. Vanden-Eijnden. String method for the study of rare events. *Phys. Rev. B*, 66:052301, 2002.

[Zhi60] G. M. Zhislin. Discussion of the spectrum of Schrödinger operators for systems of many particles. (In Russian). *Trudy Moskovskogo matematiceskogo obscestva*, 9:81–120, 1960.

[a] CERMICS, École Nationale des Ponts et Chaussées,
6 & 8 avenue Blaise Pascal, Cité Descartes,
77455 Marne-La-Vallée Cedex 2, France.
cances, galicher@cermics.enpc.fr

[b] CEREMADE, CNRS UMR 7534,
Université Paris IX Dauphine,
Place du Maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France.
lewin@ceremade.dauphine.fr

Appendix: the Newton algorithm

We describe here the Newton-like algorithm which is used in our method. Define

$$P_c = 1 - cc^T$$

the orthogonal projector on $\{c\}^\perp$,

$$Q_\varphi = P^S_\varphi$$

where P^S_φ is defined in formula (II.25), and

$$Q^*_\varphi Z = Z - \frac{1}{2} S\varphi \left(\varphi^T Z + Z^T \varphi\right)$$

the adjoint of Q_φ for the euclidian scalar product. Note that Q_φ is an orthogonal projector, only for the scalar product defined by the matrix S (see page 110).

The Newton equations take the form, $(c, \varphi, \mu, \Lambda) \in \mathbb{R}^K \times M(N_b, N) \times \mathbb{R} \times M_s(K)$ being given,

$$\left\{ \begin{array}{l} \text{Find } (d, \zeta, \nu, M) \in \mathbb{R}^K \times M(N_b, K) \times \mathbb{R} \times M_s(K) \text{ such that} \\ \left[H(\varphi) - \mu\right] d + \left[H'(\varphi) \cdot \zeta\right] c = \nu c - \left(H(\varphi) \cdot c - \mu c\right) \\ \mathcal{L}_{(c,\varphi,\mu,\Lambda)} \cdot (d, \zeta) = S\varphi M - \left(h\varphi C^s(c)^2 + \mathcal{W}(\varphi C^s(c)\varphi^T)\, \varphi C^s(c) - S\varphi\Lambda\right) \\ c^T d = -\left(c^T c - 1\right) \\ \varphi^T S\zeta + \zeta^T S\varphi = -\left(\varphi^T S\varphi - I_K\right) \end{array} \right. \qquad \text{(II.26)}$$

with

$$\begin{aligned} \mathcal{L}_{(c,\varphi,\mu,\Lambda)} \cdot (d, \zeta) = {} & h\zeta C^s(c)^2 + 2h\varphi C^s(c) C^s(d) + \\ & + \mathcal{W}(\varphi C^s(c)\zeta^T + \zeta C^s(c)\varphi^T)\, \varphi C^s(c) + \mathcal{W}(\varphi C^s(c)\varphi^T)\, \zeta C^s(c) + \\ & + \mathcal{W}(\varphi C^s(d)\varphi^T)\, \varphi C^s(c) + \mathcal{W}(\varphi C^s(c)\varphi^T)\, \varphi C^s(d) - S\zeta\Lambda. \end{aligned} \qquad \text{(II.27)}$$

Note that

$$\begin{aligned} \left[\left[H'(\varphi) \cdot \zeta\right] c\right]_i &= \sum_{j=1}^K \left[H'(\varphi) \cdot \zeta\right]_{ij} c_j \\ &= 2\left(\zeta_i^T h \varphi_i + \varphi_i^T h \zeta_i\right) c_i + \Big(\zeta_i^T \mathcal{W}\left(\varphi C^s(c)\varphi^T\right) \varphi_i \\ &\quad + \varphi_i^T \mathcal{W}\left(\varphi C^s(c)\varphi^T\right) \zeta_i \Big) + \sum_{j=1}^K \left(\zeta_j^T \mathcal{W}(\varphi_i \varphi_i^T)\varphi_j + \varphi_j^T \mathcal{W}(\varphi_i \varphi_i^T)\zeta_j\right) c_j \\ &= 2\left(\zeta_i^T h \varphi_i + \varphi_i^T h \zeta_i\right) c_i + \Big(\zeta_i^T \mathcal{W}\left(\varphi C^s(c)\varphi^T\right) \varphi_i + \\ &\quad \varphi_i^T \mathcal{W}\left(\varphi C^s(c)\varphi^T\right) \zeta_i \Big) + \mathrm{tr}\left(\mathcal{W}(\varphi_i \varphi_i^T)\left(\varphi C^s(c)\zeta^T + \zeta C^s(c)\varphi^T\right) \right). \end{aligned}$$

Let us consider a MCSCF state $(c, \varphi) \in S^{K-1} \times \mathcal{W}_K^{N_b}$ satisfying the constraints, and let

$$\mu = c^T H(\varphi) c, \qquad \Lambda = \varphi^T \left(h\varphi C^s(c)^2 + \mathcal{W}(\varphi C^s(c)\varphi^T)\varphi C^s(c)\right).$$

We define the linear map $\delta(\zeta) : M(N_b, K) \to \mathbb{R}^K$ by

$$\delta(\zeta) = -\left[P_c\left(H(\varphi) - \mu\right) P_c\right]^{-1} \cdot \left(P_c\left(\left[H'(\varphi)\cdot\zeta\right] c\right)\right).$$

Notice that $\delta(\zeta)$ is well defined as soon as $H(\varphi) - \mu$ is invertible on $\{c\}^\perp$. Let be

$$\alpha = \left[P_c\left(H(\varphi) - \mu\right) P_c\right]^{-1} \cdot \left(P_c\left(\left[H(\varphi) - \mu\right] c\right)\right).$$

We then introduce

$$L_{(c,\varphi,\mu,\Lambda)} \cdot \zeta = Q_\varphi^* \mathcal{L}_{(c,\varphi,\mu,\Lambda)}(\delta(Q_\varphi\zeta), Q_\varphi\zeta).$$

and

$$\begin{aligned} b_{(c,\varphi,\mu,\Lambda)} \;=\; & -Q_\varphi^*\Big(h\varphi C^s(c)^2 + \mathcal{W}(\varphi C^s(c)\varphi^T)\,\varphi C^s(c) - S\varphi\Lambda \\ & -2h\varphi C^s(c)C^s(\alpha) - \mathcal{W}(\varphi C^s(\alpha)\varphi^T)\varphi C^s(c) - \mathcal{W}(\varphi C^s(c)\varphi^T)\varphi C^s(\alpha)\Big). \end{aligned}$$

Newton algorithm

Initial guess. Some $(c_0, \varphi_0) \in S^{K-1} \times \mathcal{W}_K^{N_b}$. Set $k = 0$.

Iterates.

1. Set

$$\mu_k = c_k^T H(\varphi_k) c_k, \qquad \Lambda_k' = \varphi_k^T\left\{h\varphi_k C^s(c_k)^2 + \mathcal{W}(\varphi_k C^s(c_k)\varphi_k^T)\varphi_k C^s(c_k)\right\}$$

and (the symmetry of Λ_k is imposed)

$$\Lambda_k = \frac{1}{2}\left(\Lambda_k' + (\Lambda_k')^T\right).$$

2. Compute

$$r_k = H(\varphi_k)\cdot c_k - \mu_k c_k$$
$$R_k = h\varphi_k C^s(c_k)^2 + \mathcal{W}(\varphi_k C^s(c_k)\varphi_k^T)\,\varphi_k C^s(c_k) - S\varphi_k\Lambda_k.$$

If $\|r_k\| \le \varepsilon_c$ and if $\|R_k\| \le \varepsilon_\varphi$, STOP.

3. Use an orthonormalization method (Gramm-Schmidt for instance) to build a matrix $B_k \in M(K, K-1)$ such that $[c_k|B_k] \in U(K)$. Then, let

$$T_k = \left(B_k^T\left[H(\varphi_k) - \mu_k\right] B_k\right)^{-1},$$

$$\alpha_k = B_k T_k B_k^T r_k$$

and

$$\begin{aligned} b_k = -Q_k^*\Big(& R_k - 2h\varphi_k C^s(c_k)C^s(\alpha_k)) \\ & - \mathcal{W}(\varphi_k C^s(\alpha_k)\varphi_k^T)\varphi_k C^s(c_k) - \mathcal{W}(\varphi_k C^s(c_k)\varphi_k^T)\varphi_k C^s(\alpha_k)\Big), \end{aligned}$$

with $Q_k^* Z = Z - \frac{1}{2} S\varphi_k\left(\varphi_k^T Z + Z^T\varphi_k\right)$.

4. Solve the linear symmetric system

$$L_k \zeta_k = b_k,$$

where $L_k = L_{(c_k,\varphi_k,\mu_k,\Lambda_k)}$. To this end, one can use for instance the GMRes method. To apply the operator L_k to some vector ζ, one follows :

   (a) Compute $Q_k\zeta$ where $Q_k = Q_{\varphi_k}$.

   (b) Compute

$$\delta(Q_k\zeta) = -B_k T_k B_k^T \left(\left[H'(\varphi_k) \cdot (Q_k\zeta) \right] c_k \right).$$

   (c) Compute $L_k\zeta = Q_k^* \mathcal{L}_{(c_k,\varphi_k,\mu_k,\Lambda_k)} \left(\delta(Q_k\zeta), Q_k\zeta \right)$ using formula (II.27).

5. Set

$$\zeta_k = Q_k \zeta_k,$$
$$d_k = -B_k T_k B_k^T \left(\left[H'(\varphi_k) \cdot \zeta_k \right] c_k \right) - \alpha_k.$$

6. Set

$$c_{k+1} = \|c_k + d_k\|^{-1} \left(c_k + d_k \right),$$
$$\widetilde{\varphi}_{k+1} = \varphi_k + \zeta_k,$$
$$\varphi_{k+1} = \widetilde{\varphi}_{k+1} \left(\widetilde{\varphi}_{k+1}^T S \widetilde{\varphi}_{k+1} \right)^{-1/2}.$$

7. Set $k = k + 1$ and return to Step 1.

This Newton-like method has been used on the one hand to find the MCSCF ground state, and on the other hand to achieve the convergence of the mountain pass point.

For the minimization of $\mathcal{E}^{\mathrm{red}}$, the step 6 has to be improved by using a linear search method to minimize the energy on the curve $(c(t), \varphi(t))_{t\in[0,1]}$ with

$$c(t) = \|c_k + t d_k\|^{-1} \left(c_k + t d_k \right),$$
$$\widetilde{\varphi}(t) = \varphi_k + t \zeta_k,$$
$$\varphi(t) = \widetilde{\varphi}(t) \left(\widetilde{\varphi}(t)^T S \widetilde{\varphi}(t) \right)^{-1/2}.$$

Deuxième partie

Optimisation de géométrie : la modélisation de réactions chimiques adiabatiques

B.I

A Mountain Pass for Reacting Molecules

Ce chapitre reprend le texte intégral d'un article à paraître dans *Annales Henri Poincaré*.

Résumé

Dans cet article, nous considérons une molécule neutre dont les noyaux possèdent deux positions stables distinctes, et nous étudions un lemme du col entre ces deux minima, dans le cadre de la théorie linéaire de Schrödinger.

Tout d'abord, nous établissons certaines propriétés concernant le spectre et les fonctions propres d'une molécule qui se sépare en plusieurs morceaux, c'est-à-dire dont les noyaux s'écartent. Ce comportement est observé lorsqu'une suite de Palais-Smale obtenue par le procédé de min-max est non compacte. Cela nous permet d'identifier précisément les valeurs possibles de l'énergie et les *points critiques à l'infini* (voir [Bah89]), en cas de perte de compacité.

Nous restreignons ensuite notre étude à un model simplifié, mais toujours pertinent : une molécule comprenant deux sous-systèmes libres mais dont les géométries internes sont fixées. Nous démontrons alors que la perte de compacité n'a pas lieu, sous certaines hypothèses très naturelles portant sur les configurations "à l'infini". Ceci démontre l'existence d'un point selle entre les deux minima. Plus précisément, nous supposons soit que les molécules à l'infini sont chargées, soit qu'elles sont neutres, mais possèdent un moment dipolaire.

A Mountain Pass for Reacting Molecules

Mathieu Lewin

Communicated by R. Benguria

Abstract

In this paper, we consider a neutral molecule that possesses two distinct stable positions for its nuclei, and look for a mountain pass point between the two minima in the non-relativistic Schrödinger framework.

We first prove some properties concerning the spectrum and the eigenstates of a molecule that splits into pieces, a behaviour which is observed when the Palais-Smale sequences obtained by the mountain pass method are not compact. This enables us to identify precisely the possible values of the mountain pass energy and the associated "critical points at infinity" (a concept introduced by Bahri [Bah89]) in this non-compact case.

We then restrict our study to a simplified (but still relevant) model: a molecule made of two interacting parts, the geometry of each part being frozen. We show that this lack of compactness is impossible under some natural assumptions about the configurations "at infinity", proving the existence of the mountain pass in these cases. More precisely, we suppose either that the molecules at infinity are charged, or that they are neutral but with dipoles at their ground state.

AMS Subject Classification: 35B38, 35Q40, 35J10, 49J35, 81V70, 81V55, 81Q05, 49S05.

Keywords: mountain pass, critical points, critical points at infinity, Morse index, quantum mechanics, ground state, excited state, quantum chemistry, reaction path, transition state.

Introduction

In this paper, we study in the non-relativistic quantum Schrödinger framework the case of a molecule that possesses two distinct stable positions for its nuclei, as this is for instance the case for HCN and CNH. Our purpose is somewhat simple: can we obtain a critical point of the energy by using the classical mountain pass method between the two minima ? Experiment suggests that this is the case (at least for the HCN↔CNH reaction). Indeed, such mountain pass points are frequently computed by chemists who need to understand the possible behaviour of the molecule: it corresponds to a "transition state" during an infinitely slow reaction leading from one minimum to the other. But

as far as we know, this problem has never been tackled from the mathematical point of view for the N-body quantum problem, or even in the context of the classical Hartree or Thomas-Fermi type models which are approximations of the exact theory.

For a neutral molecule, the proof that there is a *minimum* with regards to the position of the nuclei can be found in the fundamental work of E.H. Lieb and W. E. Thirring [LT86] for the Schrödinger model, and in a series of papers by I. Catto and P.-L. Lions [CL92, CL93a, CL93b, CL93c] for approximate models (Hartree or Thomas-Fermi type), the latter being really more complicated due to the non linearity of these models. In these two works, the authors had to prove that minimizing sequences are compact, the non-compactness behaviour being related to the fact that the molecule can split into parts, each moving away from the others. Remark that binding does not occur for the Thomas-Fermi model (see the works of E. Teller [Tel62], E.H. Lieb and B. Simon [LS77], and the references in [CL92]) and that the result is not known for the Hartree-Fock model, except in very special cases [CL93c].

Let us make some comment on a tool used in these proofs that cannot be simply adapted to our setting. A common idea in these two works is to average over all the possible orientations of each piece in order to simplify the computation of the interaction energy between them, by suppressing the multipoles. To show that the energy can be lower than the energy at infinity, a new term using the correlation between the electrons is then created in [LT86] to obtain a Van Der Waals term of the form $-C/R^6$ (R is the distance between the molecules), while a very detailed computation of the (exponentially small) combined energy is done in [CL92, CL93a, CL93b, CL93c] to conclude that the system can bind. Because of the preliminary averaging, the conclusion is that there exists some orientation of the molecules for which this is true, but this position is unknown *a priori*.

In the case of the mountain pass method that we propose here, the non-compactness is obviously also due to a possible splitting of the molecule. However, we want to insist on the fact that we cannot use in this setting the same idea of averaging over all the rotations of the molecules, because we have to pull down the energy *along a path*. In other words, a comparison between the energies is not sufficient to conclude, and a precise information on the directions on which the energy decreases is needed.

This is why we failed to treat the problem in its full generality and we had to add some hypothesis about the configurations "at infinity". Nevertheless, we wish to ameliorate this first work in the future, and hope that it will stimulate further results.

The results proved in this paper are the following.

First, we study the spectrum and the eigenfunctions of the Hamiltonian when a molecule splits into parts. We obtain some bounds on the eigenvalues and the bottom of the essential spectrum which allow to show that the "electrons remain in the vicinity of the nuclei" when a fixed excited state is studied. In other words, no electron is lost during the process. This is obtained by a non-isotropic exponential decay of the electronic density, which is shown to be uniform when the distance between the molecules grows. We also specify the behaviour of the associated wavefunctions and define the "critical points at infinity", a concept introduced by A. Bahri [Bah89]. Some parts of this first result are necessary for our min-max problem.

Then, we prove a result that enables to identify the possible behaviour of the non-compact min-maxing paths. As it is suggested by the intuition, it is shown that the optimal

energy of the mountain pass corresponds in this case to a system where the molecule is split into independent parts (the electrons are shared among them), each being at its ground state. This Morse information on the critical points at infinity is rather intuitive.

As announced, we were unable to treat the general case and we end this article by showing that this non-compactness behaviour is impossible in the special case of two interacting molecules with fixed nuclei. This is done under the hypothesis that we are in the easy case of two charged molecules at infinity, or in the more difficult case of two neutral molecules at infinity, but with dipoles at their ground state. This enables to obtain the required result for many practical situations. As explained before, the crucial step is to evaluate the interaction energy between the molecules and we use here a multipole expansion, even for wavefunctions that are not a simple tensor product of two ground states as in [LT86]. Finally, the expected result is deduced from the fact that the critical points of the dipole/dipole interaction energy which have a nonnegative energy have a Morse index which is at least 2.

From a practical point af view, the study of this mountain pass method is really important. As mentioned above, the main idea is that a path leading from one minimum to the other represents an infinitely slow chemical reaction. The mountain pass energy is then interpreted as the lowest energy threshold for the reaction to happen. The numerical computation of this energy and of the optimum (even the whole path) is then a prime necessity for chemists, who have to understand the possible behaviours of the molecule (see for instance [Sch87, FF96] for chemical and numerical aspects).

However, chemists only consider paths on which the molecule is at its ground state all along it, which leads to obvious problems of smoothness in the case of degeneracy of the first eigenvalue of the Hamiltonian, and can obstruct convergence. For mathematical reasons, we were thus forced to abandon this hypothesis and relax the problem by considering that the wavefunction can vary independently of the nuclear geometry, in order to obtain a critical point with respect to nuclei's variations. Since we shall show that our min-max energy is in fact the same that the one used in practice, this approach could also be interesting for numerical computations.

We conclude with a few words on the mathematical tools used in this paper. As in [CL92, CL93a, CL93b, CL93c], the proof is guided by P.-L. Lions' Concentration-Compactness ideas [Lio84], although the localization of the electrons is given by the uniform exponential decay of Theorem I.2, and not by this theory. Let us remark that the physical intuition is somewhat often related to the behaviour of the electronic density. For instance, when the molecule splits into parts, the latter becomes a sum of functions localized near the nuclei. But this point of view is not sufficient to understand the problem since the main object is not the density, but the wavefunction. The latter will not split into sums, but into sums of tensor products of wavefunctions in lower dimensions (see the work of G. Friesecke [Fri03] for a very clear explanation of this phenomenon). Therefore, we use a variant of N-body geometric methods for Schrödinger operators [Sim77, Sig82, Hun66] that enables to relate the behaviour of the wavefunction to those of the associated electronic density. This method is used in [Fri03] and enabled G. Friesecke to notice an interesting link between the celebrated HVZ Theorem [Hun66, Van64, Zhi60] and the Concentration-Compactness method [Lio84].

Moreover, the HVZ Theorem (which enables to identify the bottom of the spectrum as the ground state energy of the same system but with an electron removed), and Zhislin's

Theorem (which states the existence of excited states for positive or neutral molecules) are abundantly used in this article.

Finally, we use the results and methods developed by G. Fang and N. Ghoussoub [FG92, Gho93] which enable to obtain Morse information on the Palais-Smale sequences, related to the fact that the deformed object are paths (i.e. deformations of $[0;1]$). We also use the duality theory developed in [Gho93] which permits to locate critical points.

The paper is organized as follows. In the next section, we describe the model in detail and recall known results on the Hamiltonian and its eigenfunctions. Then, in section 2, we present our results without proof: for the sake of clarity, we have brought all the proofs together in the last section.

Acknowledgments: I would like to thank Éric Séré for his constant attention and helpful remarks.

1 The model

1.1 Framework

We consider here a positive or neutral molecule with N non relativistic electrons, and M nuclei of charges $Z_1+\cdots+Z_M \geq N$. The nuclei are supposed to be correctly described by a classical model (Born-Oppenheimer approximation) and are thus represented as pointwise charges at $R_1, ..., R_M \in \mathbb{R}^3$. In what follows, we let

$$R = (R_1, ..., R_M) \in \Omega := (\mathbb{R}^3)^M \setminus (\cup_{i\neq j}\{R_i = R_j\})$$

and

$$Z = (Z_1, ..., Z_M) \in (\mathbb{N}^*)^M, \quad |Z| = Z_1 + \cdots + Z_M \geq N.$$

The system is described by the purely coulombic N-body Hamiltonian

$$H^N(R,Z) = \sum_{i=1}^{N}\left(-\frac{1}{2}\Delta_{x_i} + V_R(x_i)\right) + \sum_{1\leq i<j\leq N}\frac{1}{|x_i - x_j|} + \sum_{1\leq i<j\leq M}\frac{Z_iZ_j}{|R_i - R_j|},$$

$$V_R(u) = -\sum_{j=1}^{M}\frac{Z_j}{|u - R_j|}.$$

Its operator domain is the Sobolev space $H^2_a(\mathbb{R}^{3N}, \mathbb{C})$, and its quadratic form domain is $H^1_a(\mathbb{R}^{3N}, \mathbb{C})$. Throughout this paper, the subscript a indicates that we consider wavefunctions Ψ which are antisymmetric under interchanges of variables (expression of the Pauli exclusion principle):

$$\forall \sigma \in S_N,\ \Psi(x_1, ..., x_N) = \varepsilon(\sigma)\Psi(x_{\sigma(1)}, ..., x_{\sigma(N)}).$$

The quantum energy of the system in a state $\Psi \in H^1_a(\mathbb{R}^{3N}, \mathbb{C})$ is the associated quadratic form

$$\mathcal{E}^N(R,\Psi) = \left\langle \Psi, H^N(R,Z)\Psi \right\rangle.$$

We refer the reader to [Lie90, C$^+$03] for a description of this model and a detailed explanation of the Born-Oppenheimer approximation. The properties of $H^N(R,Z)$ and

its eigenfunctions are recalled below. For the sake of simplicity, we have neglected the spin and the dynamic of the nuclei, as in [LT86]. We would like also to mention that all the results in this paper can be adapted to the case of smeared nuclei, that is to say when the Coulomb potential $\frac{1}{|x-R_i|}$ is replaced by $\int_{\mathbb{R}^3} \frac{1}{|x-y-R_i|} d\mu_i(y)$ where μ_i is a probability measure on $\mathbb{R}^3$. Of course, $\frac{1}{|R_i-R_j|}$ has to be replaced by $\iint_{\mathbb{R}^6} \frac{1}{|R_i+y-(R_j+z)|} d\mu_i(y) d\mu_j(z)$. Remark that, in contrast to many other papers dealing with minimization, we work here with complex-valued wavefunctions, a hypothesis that plays a role in our results (see for instance Theorem I.4 and the associated remarks).

Z and N being fixed such that $N \leq |Z|$, for each $R \in \Omega$, the problem

$$E^N(R,Z) = \min\{\mathcal{E}^N(R,\Psi),\ \|\Psi\|_{L^2} = 1\}$$

has a solution Ψ, which is the ground state of the N electrons interacting with the M nuclei localized at the R_i.

For neutral molecules ($N = |Z|$), it is also known that the problem

$$E^N = \min_{R\in\Omega} E^N(R,Z)$$

admits a solution [LT86], proving the stability of neutral molecules.

We shall assume that (R,Ψ) and (R',Ψ') are two local minima of $\mathcal{E}^N$. We then consider the classical mountain pass method

$$c = \inf_{\gamma\in\Gamma} \max_{t\in[0;1]} \mathcal{E}^N(\gamma(t)) \tag{I.1}$$

where

$$\Gamma = \{\gamma \in \mathcal{C}^0([0;1], \Omega \times SH^1_a(\mathbb{R}^{3N})),\ \gamma(0) = (R,\Psi),\ \gamma(1) = (R',\Psi')\}$$

$$SH^1_a(\mathbb{R}^{3N}) = \{\Psi \in H^1_a(\mathbb{R}^{3N}),\ \|\Psi\|_{L^2} = 1\}$$

and want to show that c is a critical value of $\mathcal{E}^N$.

As mentioned in the introduction, the physical interpretation of this min-max method is that paths $\gamma \in \Gamma$ represent an infinitely slow reaction leading from one minimum to the other. c is thus interpreted as the lowest energy threshold for passing from (R,Ψ) to (R',Ψ').

In practice, the following definition is used

$$c' = \inf_{r\in\mathcal{R}} \max_{t\in[0;1]} E^N(r(t),Z)$$

where

$$\mathcal{R} = \{r \in \mathcal{C}^0([0;1],\Omega),\ r(0) = R,\ r(1) = R'\}.$$

As explained in the introduction, the function $R \mapsto E^N(R,Z)$ is continuous but not necessary differentiable and this is why we shall study the min-max method $(I.1)$. However, it will be shown that in fact $c = c'$.

1.2 Properties of $H^N(R,Z)$

Let us now recall some well-known facts about the spectrum of $H^N(R,Z)$. We introduce

$$\lambda_d^N(R,Z) = \inf_{dim(V)=d} \sup_{\substack{\Psi \in V, \\ \|\Psi\|_{L^2}=1}} \langle \mathcal{H}(R,Z)\Psi, \Psi\rangle. \tag{I.2}$$

In the sequel, we shall denote for all $d \geq 1$

$$E^0(R,Z) = \lambda_d^0(R,Z) := \sum_{1\leq i<j\leq M} \frac{Z_i Z_j}{|R_i - R_j|}, \quad \Sigma^0(R,Z) = +\infty.$$

For a wavefunction $\Psi \in H_a^1(\mathbb{R}^3, \mathbb{C})$, the electronic density and the electronic kinetic energy density are respectively defined by

$$\rho_\Psi(x) = N \int_{R^{3(N-1)}} |\Psi(x, x_2, ..., x_N)|^2 dx_2 ... dx_N$$

$$t_\Psi(x) = N \int_{R^{3(N-1)}} |\nabla\Psi(x, x_2, ..., x_N)|^2 dx_2 ... dx_N.$$

We have brought together the main known results in the following

Theorem I.1. *We assume $N \geq 1$. The following results are known :*

1. **(Self-adjointness [Kat51])** *$H^N(R,Z)$ is self-adjoint on $L_a^2(\mathbb{R}^{3N})$ with operator domain $H_a^2(\mathbb{R}^{3N})$ and quadratic form domain $H_a^1(\mathbb{R}^{3N})$.*

2. *We have [HS00]*
$$\sigma_{ess}\left(H^N(R,Z)\right) = [\Sigma^N(R,Z); +\infty)$$
$$\forall d \geq 1,\ \lambda_d^N(R,Z) \leq \Sigma^N(R,Z) \leq \sum_{1\leq i<j\leq M} \frac{Z_i Z_j}{|R_i - R_j|}$$

3. **(HVZ Theorem [Hun66, Van64, Zhi60, HS00])** *We have*
$$\Sigma^N(R,Z) = E^{N-1}(R,Z).$$

4. **(Compactness below the essential spectrum)** *If $\lambda_d^N(R,Z) < \Sigma^N(R,Z)$, then $\lambda_d^N(R,Z)$ is an eigenvalue of finite multiplicity and in particular, there exists a $\Psi_d \in H_a^2(\mathbb{R}^{3N})$ such that*
$$H^N(R,Z)\Psi_d = \lambda_d^N(R,Z)\Psi_d.$$
It is locally lipschitz [Kat57], i.e.
$$\Psi_d \in \mathcal{C}^0(\mathbb{R}^{3N}) \quad \text{and} \quad |\nabla\Psi_d| \in L^\infty_{loc}(\mathbb{R}^{3N}),$$
and real analytic [FHØ02a] on $U^N \setminus \{x_i = x_j\}$, where $U = \mathbb{R}^3 \setminus \{R_i\}_{i=1}^M$. If ρ_d is the associated electronic density, then [FHØ02a]
$$\rho_d \in \mathcal{C}^\omega(U) \cap \mathcal{C}^{0,1}(\mathbb{R}^3).$$

5. **(Zhislin Theorem [Zhi60])** *For positive and neutral molecules $N \leq |Z|$, then*

$$\lambda_d^N(R,Z) < \Sigma^N(R,Z)$$

for all $d \geq 1$, so that

$$\sigma\left(H^N(R,Z)\right) = \{\lambda_1^N(R,Z) \leq \cdots \leq \lambda_d^N(R,Z) \leq \cdots\} \cup [\Sigma^N(R,Z); +\infty).$$

6. **(Negative molecules [HS00, Lie84])** *For negative molecules $N > |Z|$, there exists a δ such that $\lambda_\delta^N(R,Z) = \Sigma^N(R,Z)$, and $\delta = 1$ when $N \geq 2|Z| + M$.*

Note that the functions $\Sigma^N(R,Z)$ and $\lambda_d^N(R,Z)$ $(d \geq 1)$ are continuous with respect to R.

2 The results

In this section, we present the results that we have obtained concerning the mountain pass method defined above. As mentioned, all the proofs are postponed to the next section.

2.1 The spectrum of a molecule that splits into pieces

We begin the study by some general results about the spectrum and the behaviour of the eigenstates when the molecule splits into pieces, that is to say when $|R_i - R_j| \to +\infty$ for some i and j. As mentioned above, this splitting of the molecule will be shown to be the main reason for the possible lack of compactness of Palais-Smale sequences.

Although only the case of ground states will be necessary for the sequel, we tackle here arbitrary excited states. In this section, we consider a positive or neutral molecule $(N \leq |Z|)$.

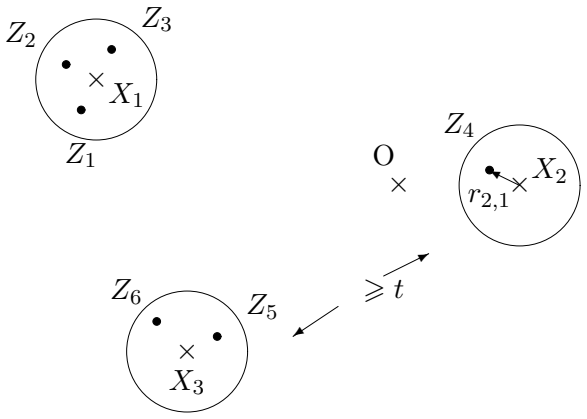


Figure I.1: An example with $M = 6$, $p = 3$, $m_1 = 3$, $m_2 = 1$ and $m_3 = 2$.

We fix a $2 \leq p \leq M$ (number of pieces).

Let $X_1, ..., X_p : \mathbb{R}^+ \mapsto \mathbb{R}^3$ be p functions that satisfy $|X_i(t) - X_j(t)| \geq t$ for all $i \neq j$ and t large enough. Let be $m = (m_1, ..., m_p) \in (\mathbb{N}^*)^p$ such that $\sum_{j=1}^p m_j = M$. We fix

a positive constant R_0 and some $r_{j,k} \in \mathbb{R}^3$ and $z_{j,k} \in \mathbb{N}$ for $j = 1, ..., p$ and $k = 1, ..., m_j$ such that $|Z| := \sum_{j=1}^{p} \sum_{k=1}^{m_j} z_{j,k} \geq N$, $|r_{j,k}| \leq R_0$ and $r_{j,k} \neq r_{j,l}$ when $k \neq l$. We then let

$$z_j = (z_{j,1}, ..., z_{j,m_j}), \quad Z = (z_1, ..., z_p), \quad r_j = (r_{j,1}, ..., r_{j,m_j}),$$

$$\tilde{r}_j(t) = (X_j(t) + r_{j,1}, ..., X_j(t) + r_{j,m_j}), \quad R(t) = (\tilde{r}_1(t), ..., \tilde{r}_p(t)).$$

We also introduce

$$\omega_j = \{(r_k) \in B(0, R_0)^{m_j},\ r_{k_1} \neq r_{k_2} \text{ if } k_1 = k_2\}.$$

and

$$\mathcal{U}(R) = \mathbb{R}^3 \setminus \left(\bigcup_{j=1}^{p} \overline{B(X_j, R)} \right).$$

2.1.1 Spectrum and uniform exponential decay

We have the following result:

Theorem I.2 (Spectrum and uniform exponential decay). *For all $1 \leq N \leq |Z|$ and all $d \geq 1$, we have*

1. $\lim_{t \to +\infty} E^N(R(t), Z) = \min \left\{ \sum_{j=1}^{p} E^{N_j}(r_j, z_j),\ N_1 + ... + N_p = N \right\}.$

2. $\limsup_{t \to +\infty} \lambda_d^N(R(t), Z) \leq \min \left\{ \sum_{j=1}^{p} \lambda_{\delta_j}^{N_j}(r_j, z_j),\ N_1 + ... + N_p = N,\ \prod_{j=1}^{p} \delta_j = d \right\}.$

3. $\inf_{r_j \in \omega_j} \liminf_{t \to +\infty} \left(\Sigma^N(R(t), Z) - \lambda_d^N(R(t), Z) \right) > 0.$

4. *Let Ψ_R be an eigenfunction associated to the eigenvalue $\lambda_d^N(R, Z)$, with associated densities ρ_R and t_R. Then there exist positive constants R_1, C and α, depending only on N, d, R_0, p such that*

$$\rho_R(x) \leq C \exp(-\alpha \delta(x)) \quad \text{and} \quad t_R(x) \leq C \exp(-\alpha \delta(x)) \quad \text{on} \quad \mathcal{U}(R_1)$$

where $\delta(x) = \min\{|x - X_j|,\ j = 1, ..., d\}$.

The first part 1) identifies the limit of the ground state energy. This type of result is rather intuitive and classical. However, since we do not know a reference in this precise setting, a proof will be given in the next section.

The interpretation of the last part 4) is that if a neutral or positively charged molecule splits into parts, then for a fixed excited state, the electrons remain in the vicinity of the nuclei.

For the sake of simplicity, let us denote, for $r = (r_1, ..., r_p)$ and $z = (z_1, ..., z_p)$

$$\Lambda_d^N(r, z) := \min \left\{ \sum_{j=1}^{p} \lambda_{\delta_j}^{N_j}(r_j, z_j),\ N_1 + ... + N_p = N,\ \prod_{j=1}^{p} \delta_j = d \right\}.$$

2.1.2 Behaviour of the wavefunctions, critical points at infinity

Now that we have some bounds on the eigenvalues and the bottom of the essential spectrum, we want to prove a result describing the behaviour of the eigenfunctions. This will enable us to define the "critical points at infinity" of the model, a concept that was introduced by A. Bahri [Bah89].

The right hand side of Theorem I.2 - 1), or more generally an equality like

$$c = \sum_{j=1}^{p} \lambda_{\delta_j}^{N_j}(r_j, z_j), \tag{I.3}$$

is rather standard from P.-L. Lions' Concentration-Compactness point of view: when the molecule splits into pieces, then the energy of an electronic excited states becomes a sum of excited states energies of the pieces. In other words, a "critical point at infinity" would be a system constituted by p molecules in some excited state, each being infinitely far from the others, so that the interactions between them vanish.

Let us consider a sequence $t_n \to +\infty$, some $r = (r_1, ..., r_p) \in \bigotimes_{j=1}^{p} \omega_j$, and denote by $R^n = R(t_n) := (X_1(t_n) + r_1, ..., X_p(t_n) + r_p)$, $X_j^n = X_j(t_n)$. For the electronic density, such a configuration is then clearly obtained in the case of dichotomy, that is to say $\rho^n \simeq \sum_{j=1}^{p} \rho_j^n$ where ρ_j^n is essentially supported in the vicinity of X_j^n (see the exponential decay of Theorem I.2). But the behaviour of the wavefunction Ψ^n is less simple since these functions will not split into sum of functions, but into sums of antisymmetric tensor products of wavefunctions in lower dimensions. In other words, a simple way to represent these non interacting molecules in terms of the wavefunction is to take

$$\Psi^n = \tau_{X_1^n} \cdot \psi_1 \wedge \cdots \wedge \tau_{X_p^n} \cdot \psi_p \tag{I.4}$$

where each ψ_j is an eigenfunction of $H^{N_j}(r_j, z_j)$ associated to the eigenvalue $\lambda_{\delta_j}^{N_j}(r_j, z_j)$. We have used here the notation

$$\tau_v \cdot \Psi(x_1, ..., x_N) := \Psi(x_1 - v, ..., x_N - v)$$

and we recall that the tensor product is defined for $\psi \in L_a^2(\mathbb{R}^{3N_1})$ and $\psi' \in L_a^2(\mathbb{R}^{3N_2})$ by

$$\psi \wedge \psi'(x_1, ..., x_{N_1+N_2}) = \frac{1}{\sqrt{N!N_1!N_2!}} \sum_{\sigma \in S_N} \varepsilon(\sigma) \psi(x_\sigma^1) \psi'(x_\sigma^2).$$

where $x_\sigma^1 := (x_{\sigma(1)}, ..., x_{\sigma(N_1)})$ and $x_\sigma^2 = (x_{\sigma(N_1+1)}, ..., x_{\sigma(N)})$.

With $(I.4)$, one easily sees that $\rho^n = \sum_{j=1}^{p} \tau_{X_j^n} \cdot \rho_j$ with obvious notations, and that

$$H^N(R^n, Z)\Psi^n - c\Psi^n \to 0$$

in $L^2(\mathbb{R}^{3N})$ as $n \to +\infty$.

When a $\lambda_{\delta_j}^{N_j}(r_j, z_j)$ is degenerated, we can obtain the same behaviour by taking a wavefunction which is a sum of such antisymmetric tensor products

$$\Psi^n \in \bigwedge_{j=1}^{p} \tau_{X_j^n} \cdot ker\left(H^{N_j}(r_j, z_j) - \lambda_{\delta_j}^{N_j}(r_j, z_j)\right).$$

To simplify notations, we shall denote $\tau_n \cdot (\psi_1 \wedge \cdots \wedge \psi_p) := \tau_{X_1^n} \cdot \psi_1 \wedge \cdots \wedge \tau_{X_p^n} \cdot \psi_p$, so that $\Psi^n = \tau_n \cdot \Psi$ where $\Psi \in \bigwedge_{j=1}^{p} ker\left(H^{N_j}(r_j, z_j) - \lambda_{\delta_j}^{N_j}(r_j, z_j)\right)$.

Suppose now that a molecule splits into two identical pieces: $r_1 = r_2$ and $z_1 = z_2$, and that $N_1 \neq N_2$. At infinity, we shall obtain two molecules with the same configurations of the nuclei, but not the same number of electrons. Since there is no reason to distinguish the two states obtained by inverting the electrons between the two molecules, a wavefunction can be *a sum* of these two states with the same energies. We are thus led to introduce the following definition.

Definition I.1. *Let* $r^n = (r_1^n, ..., r_p^n) \in \bigotimes_{j=1}^p \omega_j$ *be such that* $r_j^n \to r_j \in \omega_j$, *and* $R^n = R(t_n) := (X_1(t_n) + r_1, ..., X_p(t_n) + r_p)$, $X_j^n = X_j(t_n)$, *for some* $t_n \to +\infty$.
Let c *be such that the set*

$$\mathcal{A}_c^N(r,z) = \left\{ (N_j, \delta_j) \in (\mathbb{N}^p)^2,\ N_1 + \cdots + N_p = N,\ \sum_{j=1}^p \lambda_{\delta_j}^{N_j}(r_j, z_j) = c \right\} \tag{I.5}$$

is not empty. The sequence (R^n, Ψ^n) *in* $\Omega \times SH_a^1(\mathbb{R}^{3N})$ converges to a critical point at infinity of energy c *if there exists some*

$$\Psi \in \sum_{(N_j,\delta_j)\in\mathcal{A}_c^N(r,z)} \left(\bigwedge_{j=1}^p ker\left(H^{N_j}(r_j, z_j) - \lambda_{\delta_j}^{N_j}(r_j, z_j) \right) \right) \tag{I.6}$$

such that

$$\|\Psi^n - \tau_n \cdot \Psi\|_{H_a^1(\mathbb{R}^{3N})} \to 0. \tag{I.7}$$

To justify the term *critical point*, we remark that one can prove

Lemma I.1. *Let be* c *and* Ψ *that satisfy* $(I.5)$ *and* $(I.6)$. *Then*

$$(H^N(R^n, Z) - c)(\tau_n \cdot \Psi) \to 0$$

in $L^2(\mathbb{R}^{3N})$.

Details will be given later on.

Saying differently, a *critical point at infinity* is a class

$$\tau_K(r, \Psi) = \left\{ \tau_{X_1,...,X_p}(r, \Psi),\ X_j \in \mathbb{R}^3,\ |X_i - X_j| \geq K \right\}$$

where $\Psi \in \sum_{(N_j,\delta_j)\in\mathcal{A}_c^N(r,z)} \left(\bigwedge_{j=1}^p ker\left(H^{N_j}(r_j, z_j) - \lambda_{\delta_j}^{N_j}(r_j, z_j) \right) \right)$, $r_j \in \omega_j$ and $\tau_{X_1,...,X_p}$ is defined on each $\Omega \times \bigwedge_{j=1}^p H_a^1(\mathbb{R}^{3N_j})$ by

$$\tau_{X_1,...,X_p} \cdot r = (X_1 + r_1, \cdots, X_p + r_p)$$

$$\tau_{X_1,...,X_p} \cdot (\psi_1 \wedge \cdots \wedge \psi_p) = (\tau_{X_1}\psi_1) \wedge \cdots \wedge (\tau_{X_p}\psi_p).$$

A sequence (R^n, Ψ^n) converges to this critical point at infinity when there exists a $K_n \to +\infty$ such that

$$\lim_{n\to+\infty} d[(R^n, \Psi^n); \tau_{K_n}(r, \Psi)] = 0.$$

Let us now fix a sequence $t_n \to +\infty$ and some $r^n = (r_1^n, ..., r_p^n) \in \bigotimes_{j=1}^p \omega_j$ such that $r_j^n \to r_j \in \omega_j$, and denote $R^n = R(t_n)$, $X_j^n = X_j(t_n)$. We then have the following result concerning the behaviour of the eigenfunctions

Theorem I.3. *We assume $1 \leq N \leq |Z|$ and $d \geq 1$. Let (Ψ^n) be a sequence of wavefunctions such that*

$$H^N(R^n, Z) \cdot \Psi^n = \lambda_d^N(R^n, Z) \cdot \Psi^n.$$

Then, up to a subsequence, we have $\lim_{n\to+\infty} \lambda_d^N(R^n, Z) := c$ with

$$\Lambda_1^N(r, z) \leq c = \sum_{j=1}^{p} \lambda_{\delta_j'}^{N_j'}(r_j, z_j) \leq \Lambda_d^N(r, z)$$

for some $(N_j', \delta_j') \in \mathcal{A}_c^N(r, z)$, and (R^n, Ψ^n) converges to a critical point at infinity of energy c.

2.2 The mountain pass method: a general result

Let us now come back to our mountain pass method, and consider again a neutral molecule ($N = |Z|$). Recall that (R, Ψ) and (R', Ψ') are two local minima of $(R, \Psi) \mapsto \mathcal{E}^N(R, \Psi)$, and that c and c' are defined by

$$c = \inf_{\gamma\in\Gamma} \max_{t\in[0;1]} \mathcal{E}^N(\gamma(t)) \tag{I.8}$$

$$\Gamma = \{\gamma \in \mathcal{C}^0([0;1], \Omega \times SH_a^1(\mathbb{R}^{3N})),\ \gamma(0) = (R, \Psi),\ \gamma(1) = (R', \Psi')\}.$$

$$c' = \inf_{r\in\mathcal{R}} \max_{t\in[0;1]} E^N(r(t), Z)$$

$$\mathcal{R} = \{r \in \mathcal{C}^0([0;1], \Omega),\ r(0) = R,\ r(1) = R'\}.$$

The following result enables to identify c in the case of lack of compactness:

Theorem I.4. *We assume $N = |Z|$. We have $c = c'$. There exists a min-maxing sequence $(R^n, \Psi^n) \in \Omega \times H_a^1(\mathbb{R}^{3N})$ such that:*

If $\sum_{i\neq j} |R_i^n - R_j^n|$ is bounded then, up to a translation, (R^n, Ψ^n) converges strongly in $\Omega \times H_a^1(\mathbb{R}^{3N})$ to some critical point (R, Ψ) of $\mathcal{E}^N$ such that

$$H^N(R, Z) \cdot \Psi = c \cdot \Psi, \quad c = \lambda_1^N(R, Z).$$

If $\sum_{i\neq j} |R_i^n - R_j^n|$ is not bounded, there exists a $2 \leq p \leq M$, some $X_j^n \in \mathbb{R}^3$ with $j = 1, ..., p$ and a $R_0 > 0$ such that, changing the indices if necessary, $R^n = (X_1^n + r_1^n, ..., X_p^n + r_p^n)$, $\|r_j^n\| \leq R_0$, $Z = (z_1, ..., z_p)$, and $\lim_{n\to+\infty} |X_i^n - X_j^n| = +\infty$, $r_j^n \to r_j$. Then

$$c = \Lambda_1^N(r, z) = \min\left\{\sum_{j=1}^{p} E^{N_j}(r_j, z_j),\ N_1 + \cdots + N_p = N\right\}$$

and (R^n, Ψ^n) converges up to a subsequence to a critical point at infinity of energy $c = \Lambda_1^N(r, z)$.

As a consequence, in the non-compact case, the molecule splits into pieces, the electrons being shared among them and at their ground state. We also believe that the r_j correspond to positions of the nuclei with a Morse index equal to 0, but this is not necessary for the sequel.

This result should be seen as the first step towards concluding the existence of a critical point of energy c, by proving that the second case in Theorem I.4 does not

happen. Unfortunately, we met with serious difficulties when trying to solve this general problem. This is why the compactness will be shown in the next section for the special case of two interacting molecules with fixed nuclei.

Remark – Throughout this paper, we work with complex-valued wavefunctions Ψ. Although in other situations (minimization for instance) one often works with real-valued functions without any change, this is not the case here. In particular, the equality $c = c'$ is very easily obtained in this setting, while one can prove that this is also true for real-valued wavefunctions, but for well-chosen ground states Ψ and Ψ' only. See the proof for more details.

2.3 Compactness in the case of two interacting molecules

Now that we have identified the critical points at infinity for the mountain pass method, the next step is to show that min-maxing paths cannot approach these critical points. We study here the case of two interacting molecules with fixed nuclei. The parameters are then

- the distance between the two molecules (denoted by α in the sequel),
- the orientation of each molecule (represented by two rotations u and u'),
- the electronic wavefunction.

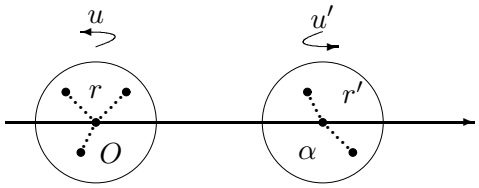


Figure I.2: Two molecules with fixed nuclei.

So we consider $r = (r_1, ..., r_m) \in B(0, R_0)^m$ and $r' = (r'_1, ..., r'_{m'}) \in B(0, R_0)^{m'}$ such that $r_1 = r'_1 = 0$, $r_i \neq r_j$ and $r'_i \neq r'_j$ for $i \neq j$, and some $z = (z_1, ..., z_m)$, $z' = (z'_1, ..., z'_{m'})$. We denote by $Z = (z, z')$, and introduce

$$R(\alpha, u, u') = (u \cdot r, \alpha\vec{v} + u' \cdot r'),$$

where $\vec{v}$ is a fixed vector of norm 1, $\alpha \in \mathbb{R}$, and u, u' are rotations in $\mathbb{R}^3$. We have used the notation $u \cdot r = (u \cdot r_1, ..., u \cdot r_m)$. We suppose now that $N = |Z|$ and define

$$\mathcal{E}^N(\alpha, u, u', \Psi) := \mathcal{E}^N(R(\alpha, u, u'), \Psi).$$

In [LT86], it is proved that $\mathcal{E}^N$ admits a minimum on $\mathbb{R} \times (S\mathcal{O}_3(\mathbb{R}))^2 \times SH^1_a(\mathbb{R}^{3N})$. As in the previous sections, we shall assume that $\mathcal{E}^N$ possesses two local minima M and M'. Up to a rotation of each molecule, we may suppose that $\alpha(M) > 0$ and $\alpha(M') > 0$. We then consider

$$c = \inf_{\gamma \in \Gamma} \max_{t \in [0;1]} \mathcal{E}^N(\gamma(t))$$

where Γ is the set of all the continuous functions $\gamma : [0;1] \to X := (0;+\infty) \times (S\mathcal{O}_3(\mathbb{R}))^2 \times SH^1_a(\mathbb{R}^{3N})$ such that $\gamma(0) = M$ and $\gamma(1) = M'$.

2.3.1 The mountain pass method

We begin this section by stating a result which is the analogue of Theorem I.4 in this special setting.

Theorem I.5. *We have*

- *either there exists a critical point* (α, u, u', Ψ) *of* $\mathcal{E}^N$ *on* X*, such that*

$$H^N(R(\alpha,u,u'),Z)\cdot\Psi = c\cdot\Psi, \quad c = \lambda_1^N(R(\alpha,u,u'),Z),$$

- *or*

$$c = \min\left\{E^{N_1}(r,z) + E^{N_2}(r',z'),\ N_1+N_2 = N\right\}.$$

Roughly speaking, the non compactness of min-maxing sequences is related to the existence of two gradient lines going from a local minimum to some critical point at infinity of index 0. The idea is that an "optimal path" has to follow these lines, and then to connect the two critical points at infinity. Since the molecule is split here into two independent parts, the problem is now to find two mountain pass paths connecting each configuration of the two molecules – two similar problems of lower dimension. When the position of the nuclei in each molecule is fixed, these paths can be obtained by only applying some rotations. In other words, the minima always belong to the same connected component and this is why the situation will be simpler in this setting. When the position is not fixed, even if we may assume the existence of such paths (by induction), the situation is much more complicated and we hope to come back to this more general issue in the future.

The proof of Theorem I.5 is very similar to the one of Theorem I.4, and will be ommitted.

In order to prove that the second case in Theorem I.5 does not happen, we need some information on the directions on which the energy decreases near the critical points at infinity. We shall thus need an expansion of the interaction energy between the two molecules when α grows. The terms involving in these expansion are classical. Let us first recall the definitions of the first multipoles.

Definition I.2. *Let be* $R = (R_1, ..., R_M) \in \Omega$*,* $Z = (Z_1, ..., Z_M)$*, and* $\rho \in L^1(\mathbb{R}^3) \cap \mathcal{S}(\mathbb{R}^3 \setminus \{R_j\})$ *a non negative function such that* $\int_{\mathbb{R}^3}\rho = N > 0$*. Then*

1. *the* total density of charge *is the measure* $\tilde{\rho} := \rho - \sum_{j=1}^M Z_j\delta_{R_j}$*. The* total charge *is* $q := \int_{\mathbb{R}^3}\tilde{\rho} = N - |Z|$*,*
2. *the* dipole moment *is the vector* $P := \int_{\mathbb{R}^3} x\tilde{\rho}(x)\,dx = \int_{\mathbb{R}^3} x\rho(x)\,dx - \sum_{j=1}^M Z_jR_j$*,*
3. *the* quadrupole moment *is the matrix* $Q := \int_{\mathbb{R}^3}\left(xx^T - \frac{1}{3}|x|^2 I\right)\tilde{\rho}(x)\,dx$*.*

When ρ is the electronic density associated to some eigenstate Ψ, we shall use the notations ρ_Ψ, $\tilde{\rho}_\Psi$, P_Ψ and Q_Ψ.

This multipoles will be used in the expansion of the interaction energy. To illustrate this point, we give here the following

Lemma I.2. *We assume that N_1 and N_2 are such that $N_1+N_2 = N$, $E^{N_1}(r,z) < \Sigma^{N_1}(r,z)$ and $E^{N_2}(r',z') < \Sigma^{N_2}(r',z')$. Let ψ_1 and ψ_2 be two ground states of respectively $H^{N_1}(r,z)$ and $H^{N_2}(r',z')$. Denoting*

$$\Psi(\alpha,u,u') = (u\cdot\psi_1)\wedge(\tau_{\alpha\vec{v}}\cdot u'\cdot\psi_2),$$

we have

$$\begin{aligned}\mathcal{E}^N\left(R(\alpha,u,u'),\Psi(\alpha,u,u')\right) &= E^{N_1}(r,z)+E^{N_2}(r',z')\\ &+\frac{q_1q_2}{\alpha}+q_2\frac{uP_1\cdot\vec{v}}{\alpha^2}-q_1\frac{u'P_2\cdot\vec{v}}{\alpha^2}+\frac{(uP_1)\cdot(u'P_2)-3(uP_1\cdot\vec{v})(u'P_2\cdot\vec{v})}{\alpha^3}\\ &+\frac{3(q_2uQ_1u^T+q_1u'Q_2u'^T)v\cdot v}{2\alpha^3}+O\left(\frac{1}{\alpha^4}\right)\end{aligned}$$

for all $u,u'\in S\mathcal{O}_3(\mathbb{R})$ and when α goes to $+\infty$.

In this result, q_k, P_k and Q_k are respectively the total charge, the dipole and the quadrupole moment associated to the electronic densities ρ_k of the states ψ_k.

The terms of this expansion can be interpreted respectively as the energies of the molecules, and the interaction energy between them, which decomposes into the charge/charge ($1/\alpha$), dipole/charge ($1/\alpha^2$), dipole/dipole and charge/quadrupole ($1/\alpha^3$) terms.

We are now able to state our main compactness results. As mentioned above, we had to add some hypothesis about the molecules "at infinity", concerning their multipoles in their ground state.

2.3.2 The case of charged molecules at infinity

Our first result will concern the case of monopoles at infinity, that is to say when the molecules are charged.

Theorem I.6 (Charged molecules at infinity). *Let us assume that*

$$E^{N_1}(r,z)+E^{N_2}(r',z') = \min\left\{E^{n_1}(r,z)+E^{n_2}(r',z'),\ n_1+n_2=N\right\} \tag{I.9}$$

for some N_1 and N_2 with $(N_1-|z|)(N_2-|z'|)\neq 0$.

Then the case 2) in Theorem I.5 does not happen. Therefore c is a critical value of $\mathcal{E}^N$ on X.

Remark – By $(I.9)$, we have for instance

$$\mu := E^{N_1}(r,z)-E^{|z|}(r,z) < E^{|z'|}(r',z')-E^{N_2}(r',z') := \mu'$$

for some N_1, N_2 such that $N_1+N_2=N$ and $N_1<|z|$. This can be viewed as a comparison between oxydo-reduction potentials. So $(I.9)$ will be probably true if one molecule is a oxydant and the other is a reductor.

2.3.3 The case of neutral molecules with dipole moments at infinity

If the two molecules at infinity are neutral, the first term involving in the expansion of the interaction energy is the dipole/dipole term. This is why we shall now consider the case of molecules that possess some dipole moment in their ground state (experiment suggests that this is the case for every non symmetric molecule).

Let us introduce the following definition

Definition I.3. *Let be $R = (R_1, ..., R_M) \in \Omega$, $Z = (Z_1, ..., Z_M)$ and $N > 0$ such that $\lambda_1^N(R, Z) < \Sigma^N(R, Z)$. We shall say that the molecule (R, Z, N) possesses a dipole moment at its ground state if $P_\Psi \neq 0$ for all ground state Ψ.*

Since $V := ker\left(H^N(R, Z) - E^N(R, Z)\right)$ is finite dimensional, let us notice that this implies $\min\{|P_\Psi|,\ \Psi \in V,\ \|\Psi\|_{L^2} = 1\} > 0$.

We then have the following result:

Theorem I.7 (Neutral molecules with dipole moments at infinity).
Let us assume that

(H1) $E^{|z|}(r, z)+E^{|z'|}(r', z') < E^{N_1}(r, z)+E^{N_2}(r', z')$ *for all* N_1, N_2 *such that* $N_1+N_2 = N$ *and* $(N_1 - |z|)(N_2 - |z'|) \neq 0$,

(H2) the two molecules $(r, z, |z|)$ *and* $(r', z', |z'|)$ *possess a dipole moment at their ground state,*

(H3) $E^{|z|}(r, z)$ *or* $E^{|z'|}(r', z')$ *is non-degenerated.*

Then the case 2) *in Theorem I.5 does not happen. Therefore c is a critical value of* $\mathcal{E}^N$ *on* X.

Remark – *(H3)* is a purely mathematical restriction that simplifies the proof.

Let us explain the general idea of the proof. Recall that the dipole/dipole interaction energy can be written $F(P, P')/\alpha^3$ (see Lemma I.2). It is shown in Appendix 2 that the critical points of F which have a non-negative energy have a Morse index which is at least one. If a path approaches a critical point at infinity then, to pull down the energy along the path, one may use either the rotations of the molecules if the dipole/dipole interaction energy is positive (thanks to this Morse index information on F), or the distance between them if it is negative (because $\alpha \mapsto F(P, P')/\alpha^3$ is then increasing). This is why min-maxing paths do not approach the critical points at infinity, and give thus a compact Palais-Smale sequence. Obviously, this general idea does not suffice to lead the proof and there are some other difficulties (essentially due to the complexity of the model) that are explicited in the next section.

Remark – This general information on the Morse index is probably true for the others multipoles interaction energies, a fact that could be used to treat the general case.

3 Proofs

3.1 Proof of Theorems I.2 and I.3

3.1.1 Preliminaries

We shall use the following lemma, which is an adaptation of results in [HO77, HØ01, FHØ02b, FHØ02c, FHØ02a], and which is proved in Appendix 1.

Lemma I.3. *Let Ψ_R be an eigenfunction associated to the eigenvalue $\lambda_d^N(R,Z)$ and ρ_R be the electronic density. We introduce $\varepsilon_R = \Sigma^N(R,Z) - \lambda_d^N(R,Z)$. Then*

1. *ρ_R satisfies the inequation*

$$-\frac{1}{2}\Delta\rho_R + V_R\rho_R + \varepsilon_R\rho_R \leq 0. \tag{I.10}$$

2. *With $R_1(\varepsilon) := max\left(R_0+1, R_0+\frac{2Np}{\varepsilon}\right)$ and $C(\varepsilon) := \cup_{j=1}^p \{x,\ |x-X_j| = R_1(\varepsilon)\}$, and if $r > 2R_1(\varepsilon_R)$, then we have*

$$\begin{aligned}
\rho_R(x) &\leq \|\rho_R\|_{L^\infty(C(\varepsilon_R))} \sum_{j=1}^p e^{-\sqrt{\varepsilon_R/p}(|X_j-x|-R_1(\varepsilon_R))} \\
&\leq p\|\rho_R\|_{L^\infty(C(\varepsilon_R))} e^{-\sqrt{\varepsilon_R/p}(\delta(x)-R_1(\varepsilon_R))} \\
&\leq M e^{-\sqrt{\varepsilon_R/p}(\delta(x)-R_1(\varepsilon_R))},
\end{aligned} \tag{I.11}$$

on $\mathcal{U}(R_1(\varepsilon_R))$, where $\delta(x) = \min\{|x-X_j|,\ j=1,...,d\}$, and $M = M(p,N,R_1(\varepsilon_R))$.

The explicit bound (I.11) has been written in order to show the dependence of all the constants with regard to ε_R. It is clearly not optimal. It shows a non-isotropic exponential decay of the electronic density, which will be uniform if $\varepsilon_R \nrightarrow 0$. This type of bounds is studied in the work of Agmon [Agm82] and we do not know if one can use his formalism to obtain the same result. Isotropic exponential bounds for N-body eigenfunctions are frequently seen in the literature, but surprising is the fact that such non-isotropic bounds has not yet been noticed.

The next two lemmas will be useful to prove the exponential decay of Theorem I.2.

Lemma I.4. *For all $\alpha > 0$, there exists a constant $M = M(\alpha, N, R_0)$ such that*

$$t_R(x) \leq M \int_{B(x,\alpha)} \rho_R(y)\, dy$$

on $\mathcal{U}(R_0 + 1/2 + \alpha)$.

Proof of Lemma I.4 – see [HØ01]. □

Lemma I.5. *For all $j = 1,...,p$, $d \geq 1$ and $n \leq |z_j|$, we have*

$$\inf_{r_j \in \omega_j} (\Sigma^n(r_j,z_j) - \lambda_d^n(r_j,z_j)) > 0.$$

Proof – We have

$$\Sigma^n(r_j,z_j) - \lambda_d^n(r_j,z_j) = \tilde{\Sigma}^n(r_j,z_j) - \tilde{\lambda}_d^n(r_j,z_j)$$

where $\tilde{\lambda}_d^n(r_j,z_j)$ and $\tilde{\Sigma}^n(r_j,z_j)$ are the d^{th} eigenvalue and the bottom of the essential spectrum of the Hamiltonian with the nuclei interaction removed

$$\tilde{H}^n(r_j,z_j) = \sum_{i=1}^n \left(-\frac{1}{2}\Delta_{x_i} + V_{r_j}(x_i)\right) + \sum_{1\leq i<j\leq n} \frac{1}{|x_i-x_j|}.$$

By Zhislin's Theorem, it is known that

$$\tilde{\Sigma}^n(r_j, z_j) - \tilde{\lambda}_d^n(r_j, z_j) > 0$$

for all $r_j \in \overline{\omega_j}$ and, since this function is continuous with regard to r_j,

$$\inf_{r_j \in \omega_j} (\Sigma^n(r_j, z_j) - \lambda_d^n(r_j, z_j)) > 0. \quad \square$$

In the next result, we use both HVZ and Zhislin's Theorems. This lemma will be useful in the proof of Theorem I.2 to construct test functions.

Lemma I.6. *If the minimum*

$$\min \left\{ \sum_{j=1}^{p} \lambda_{\delta_j}^{N_j}(r_j, z_j),\ N_1 + ... + N_p = N,\ \prod_{j=1}^{p} \delta_j = d \right\}.$$

is attained for $N_1, ..., N_j$ and $\delta_1, ..., \delta_p$, then necessarily

$$\lambda_{\delta_j}^{N_j}(r_j, z_j) < \Sigma_{\delta_j}^{N_j}(r_j, z_j)$$

for all $j = 1, ..., p$.

Proof of Lemma I.6 – Remark that by definition $\lambda_{\delta_j}^0(r_j, z_j) < \Sigma^0(r_j, z_j) = +\infty$ for all δ_j. We argue by contradiction and suppose that there exists a k such that $N_k > 0$ and $\lambda_{\delta_k}^{N_k}(r_k, z_k) = \Sigma^{N_k}(r_k, z_k) = E^{N_k - 1}(r_k, z_k)$. Theorem I.1 implies $N_k \geq |z_k| + 1$. Since $\sum_{j=1}^{p}(N_k - |z_k|) = N - |Z| \leq 0$, there exists a $l \neq k$ such that $N_l < |z_l|$. We then let $\delta_j' = \delta_j$, $N_j' = N_j$ for $j \notin \{k, l\}$, $\delta_k' = 1$, $N_k' = N_k - 1$, $\delta_l' = \delta_k \delta_l$, and $N_l' = N_l + 1$. We obtain

$$\begin{aligned} \sum_{j=1}^{p} \lambda_{\delta_j}^{N_j}(r_j, z_j) - \sum_{j=1}^{p} \lambda_{\delta_j'}^{N_j'}(r_j, z_j) &= \lambda_{\delta_l}^{N_l}(r_l, z_l) - \lambda_{\delta_l'}^{N_l+1}(r_l, z_l) \\ &\geq E^{N_l}(r_l, z_l) - \lambda_{\delta_l'}^{N_l+1}(r_l, z_l) \\ &= \Sigma^{N_l+1}(r_l, z_l) - \lambda_{\delta_l'}^{N_l+1}(r_l, z_l) \\ &> 0 \end{aligned}$$

since $N_l + 1 \leq |z_l|$ (Zhislin Theorem), which is a contradiction. $\square$

3.1.2 Proof of Theorem I.2

We are now able to prove Theorem I.2.

We first prove 2). Suppose that the right hand side is attained for some $N_1, ..., N_j$ and $\delta_1, ..., \delta_p$ such that $N_1 + ... + N_p = N$ and $\prod_{j=1}^{p} \delta_j = p$. For the sake of simplicity, we may assume that $N_j > 0$ for all $j = 1, ..., p$. By Lemma I.6 and Theorem I.1, there exist eigenfunctions $\Psi_j^k \in L_a^2(\mathbb{R}^{3N_j})$ satisfying

$$H^{N_j}(r_j, z_j)\Psi_j^k = \lambda_k^{N_j}(r_j, z_j)\Psi_j^k, \quad \int_{\mathbb{R}^{3N_j}} \Psi_j^k \Psi_j^l = \delta_{kl}$$

for all $j = 1, ..., p$ and $k = 1, ..., \delta_j$. If

$$V_j = span(\Psi_j^k,\ k = 1, ..., \delta_j) \subset L_a^2(\mathbb{R}^{3N_j})$$

then we have

$$\max_{\Psi\in V_j,\ \|\Psi\|_L^2=1} \langle H^{N_j}(r_j,z_j)\Psi,\Psi\rangle = \lambda_{\delta_j}^{N_j}(r_j,z_j).$$

We now consider a sequence $t_n \to +\infty$ such that $\lim_{n\to+\infty}\lambda_d^N(R(t_n),Z) = \limsup_{t\to+\infty}\lambda_d^N(R(t),Z)$. If $\Psi\in L^2(\mathbb{R}^{3N})$, we introduce

$$\Psi_j^{k,n} = \tau_{X_j(t_n)}\cdot\Psi_j^k$$

$$\tilde V_j^n = span(\Psi_j^{k,n},\ k=1,...,\delta_j)\subset L_a^2(\mathbb{R}^{3N_j}).$$

(we recall that τ_v is the translation by v).

Now, let be

$$W_n = \tilde V_1^n\wedge\cdots\wedge\tilde V_p^n = span(\Psi_1^{k_1,n}\wedge\cdots\wedge\Psi_p^{k_p,n},\ 1\le k_j\le\delta_j)$$

which is a space of dimension $\prod_{j=1}^d\delta_j = d$. If

$$\Psi = \sum_{1\le k_j\le\delta_j} c_{k_1,...,k_p}\Psi_1^{k_1,n}\wedge\cdots\wedge\Psi_p^{k_p,n}\in W_n,$$

and $\sum|c_{k_1,...,k_p}|^2=1$, we have

$$\langle H^N(R,Z)\Psi,\Psi\rangle = \sum_{j=1}^p\sum_{1\le k_j\le\delta_j}|c_{k_1,...,k_p}|^2\langle H^{N_j}(r_j,z_j)\Psi_j^{k_j,n},\Psi_j^{k_j,n}\rangle + e_n$$

where e_n is the interaction energy between the p molecules. It is the sum of three terms

$$e_n = e_n^1 + e_n^2 + e_n^3.$$

e_n^1 is the interaction between electrons in different molecules, and contains terms like

$$\iint\frac{(\Psi_{j_1}^{k_{j_1}}\Psi_{j_1}^{k'_{j_1}})(x,...)\Psi_{j_2}^{k_{j_2}}\Psi_{j_2}^{k'_{j_2}})(y,...)}{|x-y+X_{j_2}(t_n)-X_{j_1}(t_n)|}\,dx\,dy.$$

with $j_1\neq j_2$. e_n^2 is the interaction between electrons and nuclei of different molecules, and contains terms like

$$\iint\frac{z_{j_2,i}(\Psi_{j_1}^{k_{j_1}}\Psi_{j_1}^{k'_{j_1}})(x,...)}{|x-r_{j_2,i}+X_{j_2}(t_n)-X_{j_1}(t_n)|}\,dx\,dy$$

with $j_1\neq j_2$. Finally, e_n^3 is the interaction between nuclei of different molecules

$$e_n^3 = \sum_{j_1<j_2}\sum_{1\le k_j\le m_j}\frac{z_{j_1,k_{j_1}}z_{j_2,k_{j_1}}}{|r_{j_1,k_{j_1}}-r_{j_2,k_{j_2}}+X_{j_2}(t_n)-X_{j_1}(t_n)|}.$$

It is now easy to see that each of this term tends to 0 as $n\to+\infty$.

By definition, we have

$$\begin{aligned}\lambda_d^N(R,Z) &\le \max_{\sum|c_{k_1,...,k_p}|^2=1}\langle H^N(R,Z)\Psi,\Psi\rangle\\ &\le \sum_{j=1}^p\max_{\sum|c_{k_1,...,k_p}|^2=1}\sum_{1\le k_j\le\delta_j}|c_{k_1,...,k_p}|^2\lambda_{k_j}^{N_j}(r_j,z_j)+\max e_n\\ &\le \sum_{j=1}^p\lambda_{\delta_j}^{N_j}(r_j,z_j)+\max e_n.\end{aligned}$$

We may now pass to the limit as $n \to +\infty$ in this inequality and obtain the bound

$$\limsup_{t\to+\infty} \lambda_d^N(R(t),Z) \leq \min\left\{ \sum_{j=1}^p \lambda_{\delta_j}^{N_j}(r_j,z_j),\ N_1+...+N_p=N,\ \prod_{j=1}^p \delta_j = d \right\}.$$

We then prove simultaneously 1) 3) 4) by induction on $N = 1, ..., |Z|$.
For $N = 1$, it is known that

$$\Sigma^1(R,Z) = E^0(R,Z) = \sum_{1\leq i<j\leq M} \frac{Z_i Z_j}{|R_i - R_j|},$$

and so

$$\lim_{t\to+\infty} \Sigma^1(R(t),Z) = \sum_{j=1}^p \sum_{1\leq k<l\leq m_j} \frac{z_{j,k}z_{j,l}}{|r_{j,k}-r_{j,l}|} = \sum_{j=1}^p E^0(r_j,z_j).$$

As a consequence

$$\begin{aligned}\liminf_{t\to+\infty} &\left(\Sigma^1(R(t),Z) - \lambda_d^1(R(t),Z)\right) \\ &\geq \sum_{j=1}^p E^0(r_j,z_j) - \left(\lambda_d^1(r_1,z_1) + \sum_{j=2}^p E^0(r_j,z_j)\right) \\ &= E^0(r_1,z_1) - \lambda_d^1(r_1,z_1) = \Sigma^1(r_1,z_1) - \lambda_d^1(r_1,z_1)\end{aligned}$$

and

$$\inf_{r_i\in\omega_i} \liminf_{t\to+\infty} \left(\Sigma^1(R(t),Z) - \lambda_d^1(R(t),Z)\right) \geq \inf_{r_1\in\omega_1} \left(\Sigma^1(r_1,z_1) - \lambda_d^1(r_1,z_1)\right) > 0$$

by lemma I.5. The uniform exponential decay is then a consequence of lemmas I.3 and I.4.

Let $t_n \to +\infty$ be such that $\lim_{n\to+\infty} E^1(R(t_n),Z) = \liminf_{t\to+\infty} E^1(R(t),Z)$, and $\varphi^n \in L^2(\mathbb{R}^3)$ be such that

$$H^1(R(t_n),Z)\varphi^n = E^1(R(t_n),Z)\varphi^n.$$

By the uniform exponential decay, we may write $\varphi^n = \sum_{j=1}^p \varphi_j^n + \alpha_n$ where $supp(\varphi_j^n) \subset B(X_j, r_n/3)$, and $\|\alpha_n\|_{H^1} \to 0$. Then

$$\langle H^1(R(t_n),Z)\varphi^n, \varphi^n\rangle = \sum_{j=1}^p \langle \tilde{H}^1(r_j,z_j)\varphi_j^n, \varphi_j^n\rangle + \sum_{j=1}^p \sum_{1\leq k<l\leq m_j} \frac{z_{j,k}z_{j,l}}{|r_{j,k}-r_{j,l}|} + e_n$$

where $\tilde{H}$ is the Hamiltonian with the nuclei interaction removed, and $e_n \to 0$. We have

$$\begin{aligned}\sum_{j=1}^p \langle \tilde{H}^1(r_j,z_j)\varphi_j^n, \varphi_j^n\rangle &\geq \sum_{j=1,...,p} \tilde{E}^1(r_j,z_j)\|\varphi_j^n\|_{L^2}^2 \\ &\geq \min_{j=1,...,p} \tilde{E}^1(r_j,z_j) \left(\sum_{j=1,...,p} \|\varphi_j^n\|_{L^2}^2\right)\end{aligned}$$

so

$$\begin{aligned}\langle H^1(R(t_n),Z)\varphi^n,\varphi^n\rangle &\geq \min_{j=1,\dots,p}\tilde{E}^1(r_j,z_j)+\sum_{j=1}^{p}\sum_{1\leq k<l\leq m_j}\frac{z_{j,k}z_{j,l}}{|r_{j,k}-r_{j,l}|}\\ &= \min\left\{\sum_{j=1}^{p}E^{N_j}(r_j,z_j),\ N_1+\dots+N_p=1\right\}\end{aligned}$$

and finally

$$\lim_{t\to+\infty}E^1(R(t),Z)=\min\left\{\sum_{j=1}^{p}E^{N_j}(r_j,z_j),\ N_1+\dots+N_p=1\right\}.$$

Let us now assume that 1) 3) 4) have been proved for $N-1<|Z|$. We have $\Sigma^N(R,Z)=\lambda_1^{N-1}(R,Z)$ so

$$\begin{aligned}\lim_{t\to+\infty}\Sigma^N(R(t),Z) &= \min\left\{\sum_{j=1}^{p}E^{N_j}(r_j,z_j),\ N_1+\dots+N_p=N-1\right\}\\ &= \sum_{j=1}^{p}E^{N_j}(r_j,z_j)\end{aligned}$$

for some $N_1,\dots,N_p$. But $\sum_{j=1}^{p}(|z_j|-N_j)=|Z|-(N-1)>0$ so there exists a k such that $N_k<|z_k|$. We then have, for all $d\geq 1$,

$$\begin{aligned}&\liminf_{t\to+\infty}\left(\Sigma^N(R(t),Z)-\lambda_d^N(R(t),Z)\right)\\ &\qquad\geq\sum_{j=1}^{p}E^{N_j}(r_j,z_j)-\left(\lambda_d^{N_k+1}(r_k,z_k)+\sum_{j=1,\ j\neq k}^{p}E^{N_j}(r_j,z_j)\right)\\ &\qquad=E^{N_k}(r_k,z_k)-\lambda_d^{N_k+1}(r_k,z_k)=\Sigma^{N_k+1}(r_k,z_k)-\lambda_d^{N_k+1}(r_k,z_k)\end{aligned}$$

and

$$\begin{aligned}&\inf_{r_i\in\omega_i}\liminf_{t\to+\infty}\left(\Sigma^N(R(t),Z)-\lambda_d^N(R(t),Z)\right)\\ &\qquad\geq\inf_{\substack{j=1,\dots,p\\ n\leq|z_j|}}\ \inf_{r_j\in\omega_j}\left(\Sigma^n(r_j,z_j)-\lambda_d^n(r_j,z_j)\right)>0\end{aligned}$$

by lemma I.5. The uniform exponential decay 4) is then a consequence of lemmas I.3 and I.4.

We now prove the inequality

$$\liminf_{t\to+\infty}E^N(R(t),Z)\geq\min\left\{\sum_{j=1}^{p}E^{N_j}(r_j,z_j),\ N_1+\dots+N_p=N\right\}$$

by using a variant of classical N-body geometric methods for Schrödinger operators [Sim77, Sig82, HS00], which is used in [Fri03].

Let $t_n \to +\infty$ be such that $\lim_n E^N(R(t_n), Z) = \liminf_t E^N(R(t), Z)$, and Ψ_n an associated sequence of ground states, with densities ρ_{Ψ_n} and t_{Ψ_n}. We denote $R^n = R(t_n)$ and $X_j^n = X_j(t_n)$. Due to the uniform exponential decay, one has

$$\lim_{n\to+\infty} \int_{\mathcal{U}(t_n/3)} \rho_{\Psi_n} = \lim_{n\to+\infty} \int_{\mathcal{U}(t_n/3)} t_{\Psi_n} = 0. \tag{I.12}$$

Let $\xi_n \in \mathcal{C}^\infty(\mathbb{R}^3, [0;1])$ be a cutoff function such that $\xi_n \equiv 0$ on $\mathcal{U}(t_n/3)$, $\xi_n \equiv 1$ on $\mathbb{R}^3 \setminus \mathcal{U}(t_n/3-1)$, and $\|\nabla\xi_n\|_\infty \leq 1$, $\|\Delta\xi_n\|_\infty \leq 2$. We then introduce $\chi_n(x_1, ..., x_N) = \prod_{i=1}^N \xi_n(x_i)$ and $\tilde{\Psi}_n = \chi_n \Psi_n$. Using (I.12), it is then easy to see that $\|\Psi_n - \tilde{\Psi}_n\|_{H^1} \to 0$, and

$$\begin{aligned} H^N(R^n, Z)\cdot\tilde{\Psi}_n - \lambda_1^N(R^n, Z)\cdot\tilde{\Psi}_n &= -\sum_{i=1}^N (2\nabla_{x_i}\chi_n \cdot \nabla_{x_i}\Psi_n + \Psi_n \Delta_{x_i}\chi_n) \\ &\to 0 \quad \text{in } L^2(\mathbb{R}^{3N}), \end{aligned}$$

$$\mathcal{E}^N(R^n, \tilde{\Psi}_n) - \lambda_1^N(R^n, \Psi_n)\|\tilde{\Psi}_n\|_{L^2(\mathbb{R}^{3N})} = \sum_{i=1}^N \int |\Psi_n|^2 |\nabla_{x_i}\chi_n|^2 \to 0.$$

Now, we may write $\xi_n = \sum_{j=1}^p \xi_n^j$ where $Supp(\xi_n^j) \subset B(X_j, t_n/3)$, and

$$\tilde{\Psi}_n = \sum_{1\leq k_j\leq p} \xi_n^{k_1}(x_1)\cdots\xi_n^{k_N}(x_N)\Psi_n := \sum_{1\leq k_j\leq p} \Psi_n^{k_1,...,k_N}.$$

Since the $\Psi_n^{k_1,...,k_N}$ have disjoint supports,

$$|\tilde{\Psi}_n|^2 = \sum_{1\leq k_j\leq p} |\Psi_n^{k_1,...,k_N}|^2, \quad \mathcal{E}^N(R^n, \tilde{\Psi}_n) = \sum_{1\leq k_j\leq p} \mathcal{E}^N(R_n, \Psi_n^{k_1,...,k_N}),$$

$$\left(H^N(R^n, Z) - \lambda_1^N(R^n, Z)\right)\cdot\Psi_n^{k_1,...,k_N} \to 0$$

in $L^2(\mathbb{R}^{3N})$ for all $k_1, ...k_N$. To end the proof of Theorem I.2, it suffices to bound $\mathcal{E}^N(R^n, \tilde{\Psi}_n)$ from below by the appropriate constant.. We now fix $k_1, ..., k_N$ and introduce $C_j = \{i,\ k_i = j\}$, $N_j = |C_j|$. Remark that $\Psi_n^{k_1,...,k_N}$ is antisymmetric in $(x_i)_{i\in C_j}$ for all $j = 1, ..., p$. Then

$$\begin{aligned} \mathcal{E}^N(R^n, \Psi_n^{k_1,...,k_N}) &= \sum_{j=1}^p \left(\sum_{i\in C_j} \int \frac{1}{2}|\nabla_{x_i}\Psi_n^{k_1,...,k_N}|^2 - \sum_{i\in C_j}\int V_{\tilde{r}_j}(x_i)|\Psi_n^{k_1,...,k_N}|^2 \right. \\ &\quad \left. + \sum_{k,l\in C_j} \int \frac{|\Psi_n^{k_1,...,k_N}|^2}{|x_k - x_l|} + E^0(r_j, z_j) \right) + e_n \end{aligned}$$

where

$$e_n = \sum_{1\leq j\neq j'\leq p} \left(\sum_{i\in C_j} \int V_{\tilde{r}_{j'}}(x_i)|\Psi_n^{k_1,...,k_N}|^2 + \sum_{i\in C_j}\sum_{i'\in C_{j'}} \int \frac{|\Psi_n^{k_1,...,k_N}|^2}{|x_i - x_{i'}|} \right) + e_n',$$

e'_n being the interaction energy between nuclei in different molecules, which easily tends to 0 as $n \to +\infty$. Now

$$|e_n| \leq \sum_{1\leq j\neq j'\leq p} \left(\sum_{i\in C_j} \int \frac{3|z_j|}{t_n} |\Psi_n^{k_1,\dots,k_N}|^2 + 3 \sum_{i\in C_j} \sum_{i'\in C_{j'}} \int \frac{|\Psi_n^{k_1,\dots,k_N}|^2}{t_n} \right) + e'_n \to 0$$

as $n \to +\infty$. Finally, since $\Psi_n^{k_1,\dots,k_N}$ is antisymmetric in $(x_i)_{i\in C_j}$ for all $j = 1, ..., p$ and thanks to the translation invariance of the Hamiltonian,

$$\mathcal{E}^N(R^n, \Psi_n^{k_1,\dots,k_N}) \geq \left(\sum_{j=1}^p E^{N_j}(r_j, z_j) \right) \|\Psi_n^{k_1,\dots,k_N}\|_{L^2}^2 + e_n.$$

Passing to the limit, we obtain

$$\lim_{n\to+\infty} \mathcal{E}^N(R^n, \tilde{\Psi}_n) \geq \min \left\{ \sum_{j=1}^p E^{N_j}(r_j, z_j),\ N_1 + \cdots N_p = N \right\}$$

which ends the proof of Theorem I.2. □

3.1.3 Proof of Theorem I.3

The proof uses exactly the same N-body geometric method as the end of the proof of Theorem I.2, but with $\lambda_1^N(R^n, Z)$ replaced by $\lambda_d^N(R^n, Z)$. If we suppose that $\lim_{n\to+\infty} \lambda_d^N(R^n, Z) = c$, then passing to the limit and using Theorem I.2

$$c \leq \limsup_{n\to+\infty} \lambda_d^N(R^n, Z) \leq \Lambda_d^N(r, z).$$

We have

$$\left(H^N(R^n, Z) - c\right) \cdot \Psi_n^{k_1,\dots,k_N} \to 0$$

in $L^2(\mathbb{R}^{3N})$ for all $k_1, ...k_N$. Since all the interaction terms tend to 0 (see the proof of Theorem I.2), we obtain

$$\left(\sum_{j=1}^p H^{N_j}(X_j^n + r_j^n, z_j)_{C_j} - c \right) \cdot \Psi_n^{k_1,\dots,k_N} \to 0$$

where the Hamiltonian $H^{N_j}(X_j^n + r_j^n, z_j)_{C_j}$ acts on the variables $(x_i)_{i\in C_j}$. Due to the translation invariance, we obtain

$$\left(\sum_{j=1}^p H^{N_j}(r_j^n, z_j)_{C_j} - c \right) \cdot \tilde{\Psi}_n^{k_1,\dots,k_N} \to 0$$

where $\tilde{\Psi}_n^{k_1,\dots,k_N}(x_1, ..., x_N) = \Psi_n^{k_1,\dots,k_N}(X_{k_i}^n + x_i)$. But due to the exponential decay of Ψ_n, $\tilde{\Psi}_n^{k_1,\dots,k_N}$ is precompact in $H^1(\mathbb{R}^{3N})$ and converges up to a subsequence to some $\tilde{\Psi}^{k_1,\dots,k_N}$ such that

$$\left(\sum_{j=1}^p H^{N_j}(r_j, z_j)_{C_j} \right) \tilde{\Psi}^{k_1,\dots,k_N} = c \cdot \tilde{\Psi}^{k_1,\dots,k_N}.$$

We thus have either c is an eigenvalue of $\left(\sum_{j=1}^p H^{N_j}(r_j,z_j)_{C_j}\right)$ on the tensor product $\bigotimes_{j=1}^p L^2_a(\mathbb{R}^{3N_j})_{C_j}$ (with an obvious notation), or $\tilde{\Psi}^{k_1,\dots,k_N}=0$.

Lemma I.7. *We have* $\sigma\left(\sum_{j=1}^p H^{N_j}(r_j,z_j)_{C_j}\right) = \sum_{j=1}^p \sigma\left(H^{N_j}(r_j,z_j)_{C_j}\right)$ *so that* $\sigma_{ess}\left(\sum_{j=1}^p H^{N_j}(r_j,z_j)_{C_j}\right)=[\Sigma;+\infty)$ *with*

$$\begin{aligned}\Sigma &= \min\left\{\sum_{j\neq j_0} E^{N_j}(r_j,z_j)+\Sigma^{N_{j_0}}(r_{j_0},z_{j_0}),\ 1\leq j_0\leq p\right\}\\ &> \Lambda_d^N(r,z)\end{aligned}$$

for all $d\geq 1$.

Proof of Lemma I.7 – The fact that the spectrum of $\sum_{j=1}^p H^{N_j}(r_j,z_j)_{C_j}$ is the sum $\sum_{j=1}^p \sigma\left(H^{N_j}(r_j,z_j)_{C_j}\right)$ is standard (see for instance [RS72], Theorem VIII-33). Suppose now that $\Sigma=\sum_{j\neq j_0}E^{N_j}(r_j,z_j)+\Sigma^{N_{j_0}}(r_{j_0},z_{j_0})$ for some $1\leq j_0\leq p$. If $\lambda_d^{N_{j_0}}(r_{j_0},z_{j_0})<\Sigma^{N_{j_0}}(r_{j_0},z_{j_0})$ then obviously $\Sigma>\Lambda_d^N(r,z)$. If $\lambda_d^{N_{j_0}}(r_{j_0},z_{j_0})=\Sigma^{N_{j_0}}(r_{j_0},z_{j_0})$ then $\Sigma>\Lambda_d^N(r,z)$ by lemma I.6. □

As a consequence, if c is an eigenvalue of $\left(\sum_{j=1}^p H^{N_j}(r_j,z_j)_{C_j}\right)$, it is necessary below its essential spectrum. It is then easy to see that this implies

$$\tilde{\Psi}^{k_1,\dots,k_N}\in\bigotimes_{j=1}^p ker\left(H^{N_j}(r_j,z_j)-\lambda_{\delta_j}^{N_j}(r_j,z_j)\right)_{C_j}$$

for some $\lambda_{\delta_j}^{N_j}(r_j,z_j)<\Sigma^{N_j}(r_j,z_j)$.

Now, we have $\|\Psi^n-\tau_n\cdot\Psi\|_{H^1_a(\mathbb{R}^{3N})}\to 0$ where $\Psi=\sum_{k_1,\dots,k_N}\tilde{\Psi}^{k_1,\dots,k_N}$. □

3.2 Proof of Theorem I.4

We may suppose $c>\max(\mathcal{E}^N(R,\Psi),\mathcal{E}^N(R',\Psi'))$.

Let us first prove the equality $c=c'$. Indeed, $c'\leq c$ is obvious. Let be $r_n\in\mathcal{R}$ a sequence such that $m_n:=\max_{t\in[0;1]}E^N(r_n(t),Z)\to c'$ as $n\to+\infty$. For each $n\in\mathbb{N}$, we define

$$c_n=\inf_{\psi\in\Gamma_\Psi}\max_{t\in[0;1]}\mathcal{E}^N(r_n(t),\psi(t))$$

$$\Gamma_\Psi=\left\{\psi\in\mathcal{C}^0([0;1],SH^1_a(\mathbb{R}^{3N})),\ \psi(0)=\Psi,\ \psi(1)=\Psi'\right\}.$$

We may now apply the methods of [FG92] to obtain some sequences $t_k\in[0;1]$ and $(\Psi_k)_{k\geq 1}$ such that

1. $\lim_{k\to+\infty}\mathcal{E}^N(r^n(t_k),\Psi_k)=c_n$,
2. $H^N(r_n(t_k),Z)\cdot\Psi_k-\mathcal{E}^N(r_n(t_k),\Psi_k)\cdot\Psi_k\to 0$ in $L^2(\mathbb{R}^{3N})$,
3. $\mathcal{E}^N(r_n(t_k),\Psi_k)\leq\lambda_1^N(r_n(t_k),Z)+\varepsilon_k$, $\lim_{k\to+\infty}\varepsilon_k=0$

1, 2) correspond to the classical fact that one can obtain min-maxing sequences that are almost critical. On the other hand, 3) is a consequence of the less-known fact that one can obtain Palais-Smale sequences with Morse-type information related to the dimension of the homotopy-stable class used in the min-max method, which is 1 here (paths are deformations of $[0;1]$). Since we are in $\mathbb{C} = \mathbb{R}^2$, eigenvectors always have an even Morse index and this is why λ_1^N appears in 3) and not λ_2^N.

The fact that such a sequence (Ψ_k) is precompact in $H_a^1(\mathbb{R}^{3N})$ is now a simple consequence of Theorem I.1 - 4). Indeed, the compactness below the essential spectrum is nothing else but the Palais-Smale condition of $\mathcal{E}^N$ with Morse-type information introduced in [Gho93]. We have the following general lemma, whose proof is postponed until the end of the proof of Theorem I.4.

Lemma I.8. *We assume that $Z = (Z_1, ..., Z_M)$ is such that $N \leq |Z|$. Let (R^n, Ψ^n) be a sequence in $\Omega \times SH_a^1(\mathbb{R}^{3N})$ such that*

1. $R^n \to R \in \Omega$
2. $\lim_{n\to+\infty} \mathcal{E}^N(R^n, \Psi^n) = c$,
3. $H^N(R^n, Z) \cdot \Psi^n - \mathcal{E}^N(R^n, \Psi^n) \cdot \Psi^n \to 0$ *in* $L^2(\mathbb{R}^{3N})$,
4. *there exist* $d_0 \geq 1$ *and* $\varepsilon_n \to 0$ *such that* $\mathcal{E}^N(R^n, \Psi^n) \leq \lambda_{d_0}^N(R^n, Z) + \varepsilon_n$.

Then (Ψ^n) is precompact in $H_a^1(\mathbb{R}^{3N})$ and converges, up to a subsequence, to an eigenfunction Ψ of $H^N(R, Z)$ associated to $\lambda_d^N(R, Z)$ with $d \leq d_0$.

Applying this result, we obtain, by passing to the limit as $k \to +\infty$,

$$c \leq c_n = \lambda_1^N(r_n(t_n), Z) \leq m_n$$

for some $t_n \in [0;1]$ and so $c = c'$.

We now prove the alternative of the Theorem. We introduce

$$\mathcal{F}_c(R_0) = \left\{ (R, \Psi) \in \Omega \times SH_a^1(\mathbb{R}^{3N}),\ \sum_{i\neq j} |R_i - R_j| \leq R_0,\ \mathcal{E}^N(R, \Psi) \geq c \right\}$$

$$\Gamma(\alpha) = \left\{ \gamma \in \Gamma,\ \max_{t\in[0;1]} \mathcal{E}^N(\gamma(t)) \leq c + \alpha \right\}.$$

We have the following alternative:

either there exist $R_0 > 0$ and $\alpha > 0$ such that, $\mathcal{F}_c(R_0) \cap \gamma([0;1]) \neq \emptyset$ for all $\gamma \in \Gamma(\alpha)$,

or for all $R_0 > 0$ there exists a min-maxing sequence $\gamma_n \in \Gamma$ such that $\mathcal{F}_c(R_0) \cap \gamma_n([0;1]) = \emptyset$.

First Case : there exist $R_0 > 0$ and $\alpha > 0$ such that, $\mathcal{F}_c(R_0) \cap \gamma([0;1]) \neq \emptyset$ for all $\gamma \in \Gamma(\alpha)$.

Since (R, Ψ) and (R', Ψ') do not belong to $\mathcal{F}_c(R_0)$, we may apply the methods of [Gho93] ($\mathcal{F}_c(R_0)$ is a set which is dual to the homotopy-stable class $\Gamma(\alpha)$ with boundary $B = \{(R, \Psi), (R', \Psi')\}$) to obtain a sequence $(R^n, \Psi^n) \in \Omega \times SH_a^1(\mathbb{R}^{3N})$ such that

1. $\lim_{n\to+\infty} d((R^n, \Psi^n), \mathcal{F}_c(R_0)) = 0$,

2. $\lim_{n\to+\infty} \mathcal{E}^N(R^n,\Psi^n) = c$,

3. $\lim_{n\to+\infty} \nabla_R \mathcal{E}^N(R^n,\Psi^n) = 0$,

4. $H^N(R^n,Z)\cdot\Psi^n - \mathcal{E}^N(R^n,\Psi^n)\cdot\Psi^n \to 0$ in $L^2(\mathbb{R}^{3N})$,

5. $\mathcal{E}^N(R^n,\Psi^n) \leq \lambda_1^N(R^n,Z) + \varepsilon_n, \quad \lim_{n\to+\infty}\varepsilon_n = 0$

Remark that $2,3,4,5$) correspond to the fact that one can obtain min-maxing sequences that are almost critical, and with Morse-type information. On the other hand, 1) is the consequence of the duality theory developed in [Gho93] that enables to locate the critical points.

Due to 1), $\sum_{i\neq j}|R_i^n - R_j^n|$ is bounded. Up to a translation, we may suppose $R^n \to R \in \Omega$ (since $\lambda_1^N(R^n,Z) \to +\infty$ when $d(R^n,\partial\Omega)\to 0$ due to the nuclei/nuclei repulsion). Now Ψ^n converges up to a subsequence to a Ψ in $H^1(\mathbb{R}^{3N})$ by lemma I.8.

Second Case : for all $R_0 > 0$ there exists a min-maxing sequence $\gamma_n \in \Gamma$ such that $\mathcal{F}_c(R_0)\cap\gamma_n([0;1]) = \emptyset$.

Let (r_n) be a sequence in $\mathbb{R}$ such that $r_n \to +\infty$. For each r_n, there exists a γ_n such that, for instance,

$$c \leq \max_{t\in[0;1]} \mathcal{E}^N(\gamma_n(t)) \leq c + \frac{1}{n}$$

and $\mathcal{F}_c(r_n)\cap\gamma_n([0;1]) = \emptyset$. We now write $\gamma_n(t) = (R_n(t),\tilde{\Psi}_n(t))$ and fix n. The set $K_n = \{t\in[0;1],\ \sum|R_i^n(t) - R_j^n(t)| \leq r_n\}$ is a compact subset of $[0;1]$ such that $\max\mathcal{E}^N(\gamma(K_n)) < c$. We now introduce

$$\Gamma_\Psi = \left\{\psi\in\mathcal{C}^0([0;1],SH_a^1(\mathbb{R}^{3N})),\ \psi_{|K_n} \equiv \tilde{\Psi}^n_{|K_n}\right\}$$

which is an homotopy-stable class of dimension 1 with boundary $\tilde{\Psi}^n(K_n)$, and

$$c_n = \inf_{\psi\in\Gamma_\Psi}\max_{t\in[0;1]} \mathcal{E}^N(R_n(t),\psi(t)) \tag{I.13}$$

so that

$$c \leq c_n \leq c + \frac{1}{n}.$$

Applying the methods of [Gho93], we may find a sequence $t_k \in [0;1]\setminus K_n$ and $\Psi_n^k \in SH_a^1(\mathbb{R}^{3N})$ such that $t_k \to \bar{t}$ and

1. $\sum_{i\neq j}|R_{n,i}(t_k) - R_{n,j}(t_k)| \geq r_n$

2. $\lim_{k\to+\infty} \mathcal{E}^N(R_n(t_k),\Psi_n^k) = c_n$

3. $H^N(R_n(t_k),Z)\cdot\Psi_n^k - \mathcal{E}^N(R_n(t_k),\Psi_n^k)\cdot\Psi_n^k \to 0$ in $L^2(\mathbb{R}^{3N})$

4. $\mathcal{E}^N(R_n(t_k),\Psi_n^k) \leq \lambda_1^N(R_n(t_k),Z) + \alpha_k$ with $\lim_{k\to+\infty}\alpha_k = 0$

By lemma I.8, $(\Psi_n^k)_{k\in\mathbb{N}}$ is precompact in $H_a^1(\mathbb{R}^{3N})$ and converges, up to a subsequence, to some Ψ^n such that

1. $\sum_{i\neq j}|R_i^n - R_j^n| \geq r_n$

2. $\lim_{n\to+\infty} \mathcal{E}^N(R^n, \Psi^n) = c$

3. $H^N(R^n, Z)\cdot\Psi^n = \lambda_1^N(R^n, Z)\cdot\Psi^n$

where $R^n := R_n(\bar{t})$.

Since $\sum_{i\neq j}|R_i^n - R_j^n| \to +\infty$, there exists a $2 \leq p \leq M$, some $X_j^n \in \mathbb{R}^3$ with $j = 1, ..., p$ and a $R_0 > 0$ such that (changing the indices if necessary and up to a subsequence) $R^n = (X_1^n + r_1^n, ..., X_p^n + r_p^n)$, $\|r_j^n\| \leq R_0$, $Z = (z_1, ..., z_p)$, and $\lim_{n\to+\infty}|X_i^n - X_j^n| = +\infty$, $r_j^n \to r_j$.

Passing to the limit, we obtain, by Theorem I.2

$$c = \lim_{n\to+\infty} \lambda_1^N(R^n, Z) = \Lambda_1^N(r, z).$$

We now simply apply Theorem I.3 to obtain the convergence to a critical point at infinity of energy c, as defined in the corresponding section. □

Let us now prove Lemma I.8.

Proof of Lemma I.8 – Let m be an integer such that $\lambda_m^N(R, Z) > \lambda_{m-1}^N(R, Z) = \lambda_{d_0}^N(R, Z)$. Due to the fact that $\lambda_d^N(R^n, Z) \to \lambda_d^N(R, Z)$ as $n \to +\infty$, we have $\lambda_m^N(R^n, Z) > \lambda_{d_0}^N(R, Z) \geq c$ for n large enough.

Let be $V_n = \bigoplus_{i=1}^{m-1} ker\left(H(R^n, Z) - \lambda_i^N(R^n, Z)\right)$, and $(\psi_1^n, ..., \psi_{m-1}^n)$ an orthonormal basis of V^n, ψ_i^n being eigenfunctions of $H^N(R^n, Z)$ (Theorem I.1 - 4)). Due for instance to the uniform exponential decay of Theorem I.2, one easily sees that each ψ_i^n is precompact and converges up to a subsequence in $H_a^1(\mathbb{R}^{3N})$ to a ψ_i, with $span(\psi_i) = \bigoplus_{i=1}^{m-1} ker\left(H(R, Z) - \lambda_i^N(R, Z)\right)$.

Now we can write $\Psi^n = \Psi_{V^n} + \Psi_{(V^n)^\perp}$ with an obvious definition. Since $\mathcal{E}^N(R^n, \Psi^n)$ is bounded, it is a classical fact that (Ψ^n) is bounded in $H_a^1(\mathbb{R}^{3N})$ and so, up to a subsequence, $\Psi^n \rightharpoonup \Psi$ weakly in $H_a^1(\mathbb{R}^{3N})$. Since $dim(V_n) = m - 1$, (Ψ_{V_n}) is precompact in $H^1(\mathbb{R}^{3N})$ and converges to a $\Psi_V \in V$. By difference, $\Psi_{(V^n)^\perp} \rightharpoonup \Psi_{(V)^\perp}$ weakly in $H^1(\mathbb{R}^{3N})$.

Since $\lim_{n\to+\infty} \mathcal{E}^N(R^n, \Psi^n) = c$, $H^N(R^n, Z)\cdot\Psi^n - c\cdot\Psi^n \to 0$ in $L^2(\mathbb{R}^{3N})$, so we obtain

$$(H^N(R^n, Z) - c)\cdot\Psi_{V_n} + (H^N(R^n, Z) - c)\cdot\Psi_{(V_n)^\perp} \to 0$$

which implies $(H^N(R^n, Z) - c)\cdot\Psi_{V_n} \to 0$ and $(H^N(R^n, Z) - c)\cdot\Psi_{(V_n)^\perp} \to 0$ in $L^2(\mathbb{R}^{3N})$. Finally,

$$\mathcal{E}^N(R^n, \Psi_{(V_n)^\perp}) - c\|\Psi_{(V_n)^\perp}\|^2_{L^2(\mathbb{R}^{3N})} \to 0.$$

Because $\min \mathcal{E}^N(R^n, S(V_n)^\perp) = \lambda_m^N(R^n, Z) > \lambda_{d_0}^N(R^n, Z) \geq c$, this implies $\|\Psi_{(V_n)^\perp}\|_{L^2} \to 0$ and then $\|\Psi_{(V_n)^\perp}\|_{H^1} \to 0$. Thus Ψ^n converges in $H^1(\mathbb{R}^{3N})$ to a $\Psi = \Psi_V$ which is an eigenfunction of $H^N(R, Z)$ that belongs to V. □

3.3 The case of two interacting molecules

3.3.1 Proof of Lemma I.2

We shall use the following lemma:

Lemma I.9 (Multipole expansion). *There exists a constant C such that, for all R and $h \in \mathbb{R}^3$ with $R + h \neq 0$,*

$$\left|\frac{1}{|R+h|} - \left(\frac{1}{|R|} - \frac{e_R\cdot h}{|R|^2} + \frac{3(e_R\cdot h)^2 - |h|^2}{2|R|^3}\right)\right| \leq \frac{C\,|h|^3}{|R|^3|R+h|}$$

with $e_R = R/|R|$.

Proof of Lemma I.9 – It suffices to show

$$\left| \frac{1}{\sqrt{1-2xt+t^2}} - \left(1 + xt + \frac{t^2}{2}(3x^2-1)\right)\right| \leq \frac{C|t|^3}{\sqrt{1-2xt+t^2}}$$

for all $t \in \mathbb{R}$ and $x \in [-1;1]$ (take $x = -(e_R \cdot h)/|h|$ and $t = |h|/|R|$). We thus introduce $f(x,t) = 1 - \sqrt{1-2xt+t^2}\left(1 + xt + \frac{t^2}{2}(3x^2-1)\right)$. One easily computes $\frac{\partial f}{\partial x}(x,t) = \frac{3t^3(5x^2-2xt-1)}{2\sqrt{1-2xt+t^2}}$, so that

$$\max_{x\in[-1;1]} |f(x,t)| \leq \max\{f_1(t), f_2(t), f_3(t), f_4(t)\}$$

where $f_1(t) = |f(x_1(t),t)|\mathbb{1}_{-1\leq x_1\leq 1}(t)$, $f_2(t) = |f(x_2(t),t)|\mathbb{1}_{-1\leq x_2\leq 1}(t)$, $f_3(t) = |f(-1,t)|$, $f_4(t) = |f(1,t)|$, $x_1(t) = \frac{t-\sqrt{t^2+5}}{5}$ and $x_2(t) = \frac{t+\sqrt{t^2+5}}{5}$. It is now easy to conclude that $|f(x,t)| \leq C|t|^3$ for some constant $C > 0$. □

We are now able to prove Lemma I.2.

Proof of Lemma I.2 – Let $\xi_\alpha \in \mathcal{C}^\infty(\mathbb{R}^3,[0;1])$ be a cut-off function such that $\xi_\alpha \equiv 0$ on $\mathbb{R}^3 \setminus B(O,\alpha/3)$, $\xi_\alpha \equiv 1$ on $B(O,\alpha/3-1)$, $\|\nabla\xi_\alpha\|_\infty \leq 1$, $\|\Delta\xi_\alpha\|_\infty \leq 2$. We introduce $\tilde{\psi}_j^\alpha(x_1,...,x_{N_j}) = \prod_{k=1}^{N_j}\xi_\alpha(x_k)\psi_j(x_1,...,x_{N_j})$, $\psi_j^\alpha = \tilde{\psi}_j^\alpha/\|\tilde{\psi}_j^\alpha\|_{L^2}$ and $\tilde{\Psi}(\alpha,u,u') = (u\cdot\psi_1^\alpha)\wedge(\tau_{\alpha\vec{v}}\cdot u'\cdot\psi_2^\alpha)$. Due to the exponential decay of Theorem I.2, one has

$$\left|\mathcal{E}^N\left(R(\alpha,u,u'),\tilde{\Psi}(\alpha,u,u')\right) - \mathcal{E}^N\left(R(\alpha,u,u'),\Psi(\alpha,u,u')\right)\right| \leq Ce^{-a\alpha}$$

$$\left|E^{N_1}(r,z) - \mathcal{E}^{N_1}(r,\psi_1^\alpha)\right|,\ \left|E^{N_2}(r',z') - \mathcal{E}^{N_2}(r',\psi_2^\alpha)\right| \leq Ce^{-a\alpha}$$

for some $C, a > 0$. Let us recall that, by definition,

$$\psi\wedge\psi'(x_1,...,x_N) = \sqrt{\frac{N_1!N_2!}{N!}}\sum_{|C|=N_1}\sigma(C)\psi(x_C)\psi'(x_{\overline{C}})$$

where $x_C = (x_{i_1},...,x_{i_{N_1}})$ when $C = \{i_1 < \cdots < i_{N_1}\}$, $\sigma(C) = \pm 1$. Applying this equality to $\tilde{\Psi}$, we obtain on the right functions with disjoint supports. We shall therefore only study the expansion of

$$\mathcal{E}^N\left(R(\alpha,u,u'),(u\cdot\psi_1^\alpha)\otimes(\tau_{\alpha\vec{v}}\cdot u'\cdot\psi_2^\alpha)\right).$$

We have (using the notation $x^1 = (x_1,...,x_{N_1})$ and $x^2 = (x_{N_1+1},...,x_N)$)

$$\begin{aligned}\mathcal{E}^N\left(R(\alpha,u,u'),(u\psi_1^\alpha)\otimes(\tau_{\alpha\vec{v}}u'\psi_2^\alpha)\right) &= \mathcal{E}^{N_1}(ur,u\psi_1^\alpha) + \mathcal{E}^{N_2}(\alpha\vec{v}+u'r',\tau_{\alpha\vec{v}}u'\psi_2^\alpha)\\ &+ \sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\int_{\mathbb{R}^{3N}}\frac{|\psi_1^\alpha(x^1)|^2|\psi_2^\alpha(x^2)|^2}{|\alpha\vec{v}+u'\cdot x_j^2 - u\cdot x_i^1|}dx^1dx^2 + \sum_{i=1}^{m}\sum_{j=1}^{m'}\frac{z_iz_j'}{|\alpha\vec{v}+u'\cdot r_j' - u\cdot r_i|}\\ &- \int_{\mathbb{R}^{3N}}\left\{\sum_{i=1}^{m}\sum_{j=1}^{N_2}\frac{z_i|\psi_1^\alpha(x^1)|^2|\psi_2^\alpha(x^2)|^2}{|\alpha\vec{v}+u'\cdot x_j^1 - u\cdot r_i|} + \sum_{i=1}^{N_1}\sum_{j=1}^{m'}\frac{z_j'|\psi_1^\alpha(x^1)|^2|\psi_2^\alpha(x^2)|^2}{|\alpha\vec{v}+u'\cdot r_j' - u\cdot x_i^1|}\right\}dx^1dx^2\end{aligned}$$

so we obtain

$$\mathcal{E}^N\left(R(\alpha,u,u'),(u\psi_1^\alpha)\otimes(\tau_{\alpha\vec{v}}u'\psi_2^\alpha)\right)=\mathcal{E}^{N_1}(r,\psi_1^\alpha)+\mathcal{E}^{N_2}(r',\psi_2^\alpha)$$
$$+\iint_{\mathbb{R}^6}\frac{\rho_1^\alpha(x)\rho_2^\alpha(y)}{|\alpha\vec{v}+u'\cdot y-u\cdot x|}dx\,dy+\sum_{i=1}^m\sum_{j=1}^{m'}\frac{z_iz'_j}{|\alpha\vec{v}+u'\cdot r'_j-u\cdot r_i|}$$
$$-N_1\sum_{i=1}^m\int_{\mathbb{R}^3}\frac{z_i\rho_2^\alpha(y)}{|\alpha\vec{v}+u'\cdot y-u\cdot r_i|}dy-N_2\sum_{j=1}^{m'}\int_{\mathbb{R}^3}\frac{z'_j\rho_1^\alpha(x)}{|\alpha\vec{v}+u'\cdot r'_j-u\cdot x|}dx$$

where the ρ_k^α are the electronic densities associated to ψ_k^α, and finally

$$\mathcal{E}^N\left(R(\alpha,u,u'),(u\psi_1^\alpha)\otimes(\tau_{\alpha\vec{v}}u'\psi_2^\alpha)\right)=\mathcal{E}^{N_1}(r,\psi_1^\alpha)+\mathcal{E}^{N_2}(r',\psi_2^\alpha)$$
$$+\iint_{\mathbb{R}^6}\frac{\tilde{\rho}_1^\alpha(x)\tilde{\rho}_2^\alpha(y)}{|\alpha\vec{v}+u'\cdot y-u\cdot x|}dx\,dy$$

where $\tilde{\rho}_1^\alpha(x)=\rho_1^\alpha(x)-\sum_{i=1}^m z_i\delta_{r_i}(x)$ and $\tilde{\rho}_2^\alpha(y)=\rho_2^\alpha(y)-\sum_{j=1}^{m'}z'_i\delta_{r'_i}(y)$ are the total densities of charge associated to the distributions ρ_1^α and ρ_2^α. Now, by lemma I.9, we have

$$\iint_{\mathbb{R}^6}\frac{\tilde{\rho}_1^\alpha(x)\tilde{\rho}_2^\alpha(y)}{|\alpha\vec{v}+u'\cdot y-u\cdot x|}dx\,dy=\frac{q_1q_2}{\alpha}+q_2\frac{(uP_1^\alpha)\cdot\vec{v}}{\alpha^2}-q_1\frac{(u'P_2^\alpha)\cdot\vec{v}}{\alpha^2}$$
$$-\frac{3(uP_1^\alpha\cdot\vec{v})(u'P_2^\alpha\cdot\vec{v})-(uP_1^\alpha)\cdot(u'P_2^\alpha)}{\alpha^3}+\frac{3(q_2uQ_1u^T+q_1u'Q_2u'^T)v\cdot v}{2\alpha^3}$$
$$+O\left(\frac{1}{\alpha^3}\iint_{\mathbb{R}^6}\frac{|u'y-ux|^3|\tilde{\rho}_1^\alpha|(x)|\tilde{\rho}_2^\alpha|(y)}{|\alpha\vec{v}+u'\cdot y-u\cdot x|}dx\,dy\right)$$

But we have

$$\iint_{\mathbb{R}^6}\frac{|u'y-ux|^3|\tilde{\rho}_1^\alpha|(x)|\tilde{\rho}_2^\alpha|(y)}{|\alpha\vec{v}+u'\cdot y-u\cdot x|}dx\,dy\leq\frac{3C}{\alpha}\int_{\mathbb{R}^3}|x|^3(|\tilde{\rho}_1|+|\tilde{\rho}_2|)(x)dx.$$

It suffices to notice that $|P_k^\alpha-P_k|\leq Ce^{-a\alpha}$ and $|Q_k^\alpha-Q_k|\leq Ce^{-a\alpha}$ for some $C,a>0$ to end the proof. □

3.3.2 Proof of Theorem I.6

Let us suppose that we are in the second case of Theorem I.5, and that $c>\max\{\mathcal{E}^N(M),\mathcal{E}^N(M')\}$. By the proof of Theorem I.4, we obtain a sequence $\alpha_n\to+\infty$ and paths γ_n such that

$$\Lambda_1^N\leq c\leq\max_{t\in[0;1]}\mathcal{E}^N(\gamma_n(t))\leq c+\frac{1}{n}$$

and $\mathcal{E}^N(\alpha(t),u(t),u'(t),\Psi(t))<c$ when $\alpha(t)\leq\alpha_n$. Let t_1^n and t_2^n be respectively the minimum and the maximum of $\{t,\ \alpha(t)\geq\alpha_n\}$. By the definition of c, we have $0<t_1^n<t_2^n<1$. For the sake of simplicity, we introduce $u_j^n=u(t_j^n)$, $u'^n_j=u'(t_j^n)$, and $\Psi_j^n=\Psi(t_j^n)$.

The idea of the proof is now to connect $M_1^n=(\alpha_n,u_1^n,u'^n_1,\Psi_1^n)$ and $M_2^n=(\alpha_n,u_2^n,u'^n_2,\Psi_2^n)$ by a path on which α is constant, with a maximum energy that is below c.

We shall use the following lemma

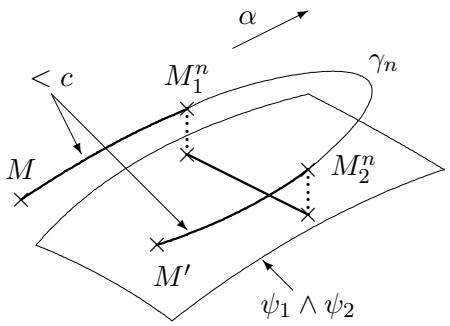


Figure I.3: A schematic representation of the proof

Lemma I.10. *We assume that* $(\alpha, u, u') \in (0; +\infty) \times (S\mathcal{O}_3(\mathbb{R}))^2$ *is fixed, and we abbreviate* $R = R(\alpha, u, u')$. *Let* Ψ_1 *and* Ψ_2 *be two wavefunctions of* $SH_a^1(\mathbb{R}^{3N})$. *Then there exists a continuous path* $\Psi : [0; 1] \to SH_a^1(\mathbb{R}^{3N})$ *such that* $\Psi(0) = \Psi_1$, $\Psi(1) = \Psi_2$, *and for all* $t \in [0; 1]$

$$\mathcal{E}^N(R, \Psi(t)) \leq \max\{\mathcal{E}^N(R, \Psi_1), \mathcal{E}^N(R, \Psi_2)\}.$$

Proof of Lemma I.10 – Let us denote by V the finite-dimensional eigenspace associated to the first eigenvalue $\lambda_1^N(R, Z)$ of $H^N(R, Z)$, and let P_V be the projection onto V. We may write $\Psi_1 = t_0\Psi_V + \sqrt{1 - t_0^2}\Psi_{V^\perp}$, where $t_0 \in [0; 1]$, $\Psi_V = \frac{P_V(\Psi_1)}{\|P_V(\Psi_1)\|_{L^2}}$ if $P_V(\Psi_1) \neq 0$, and Ψ_V is an arbitrary normalized function of V if $P_V(\Psi_1) = 0$ (then $t_0 = 0$). This enables to define a path connecting Ψ_1 to Ψ_V, on which the energy decreases, by varying $t \in [t_0; 1]$.

We may find with the same method a path from Ψ_2 to some Ψ'_V in V, with a decreasing energy. It remains now to take an arbitrary path in the sphere of V to connect Ψ_V and Ψ'_V (the sphere of V being pathwise-connected since we are in $\mathbb{C}$), on which $\mathcal{E}^N(R, \cdot)$ is constant. □

Now, let N_1 and N_2 be such that

$$E^{N_1}(r, z) + E^{N_2}(r', z') = \min\left\{E^{n_1}(r, z) + E^{n_2}(r', z'),\ n_1 + n_2 = N\right\} := \Lambda_1^N$$

and $(N_1 - |z|)(N_2 - |z'|) \neq 0$.

By lemma I.6, there exist ψ_1 and ψ_2, two ground states of respectively $H^{N_1}(r, z)$ and $H^{N_2}(r', z')$. We now introduce $\tilde{\Psi}_k^n = (u_k^n \cdot \psi_1) \wedge (\tau_{\alpha_n \vec{v}} \cdot u'^n_k \psi_2)$ for $k = 1, 2$. Since $S\mathcal{O}_3(\mathbb{R})$ is pathwise connected, one may find two paths $U_n, U'_n : [0; 1] \to S\mathcal{O}_3(\mathbb{R})$ such that $U_n(0) = u_1^n$, $U_n(1) = u_2^n$, $U'_n(0) = u'^n_1$ and $U'_n(1) = u'^n_2$. This enables to connect $\tilde{M}_1^n := (\alpha_n, u_1^n, u'^n_1, \tilde{\Psi}_1^n)$ and $\tilde{M}_2^n := (\alpha_n, u_2^n, u'^n_2, \tilde{\Psi}_2^n)$ by a path on which $\alpha \equiv \alpha_n$. Applying Lemma I.10, we may connect M_1^n to $\tilde{M}_1^n$, and M_2^n to $\tilde{M}_2^n$. We finally obtain an admissible path $\tilde{\gamma}_n$ connecting M and M', with a maximum that is attained "between" $\tilde{M}_1^n$ and $\tilde{M}_2^n$. It is now a consequence of Lemma I.2 that

$$\max_{t \in [0;1]} \mathcal{E}^N(\tilde{\gamma}_n(t)) = E^{N_1}(r, z) + E^{N_2}(r', z') + \frac{(N_1 - |z|)(N_2 - |z'|)}{\alpha_n} + O\left(\frac{1}{(\alpha_n)^2}\right).$$

Since $N_1 + N_2 = N = |z| + |z'|$, we have $(N_1 - |z|)(N_2 - |z'|) < 0$ and so

$$c \leq \max_{t\in[0;1]} \mathcal{E}^N(\tilde{\gamma}_n(t)) < E^{N_1}(r,z) + E^{N_2}(r',z')$$

for n large enough, which is a contradiction. □

3.3.3 Proof of Theorem I.7

As in the previous proofs, we may suppose that $c > \max\{\mathcal{E}^N(M), \mathcal{E}^N(M')\}$, and that there exists a sequence $\alpha_n \to +\infty$ and paths γ_n such that

$$c \leq \max_{t\in[0;1]} \mathcal{E}^N(\gamma_n(t)) \leq c + \frac{1}{n}$$

and $\mathcal{E}^N(\alpha(t), u(t), u'(t), \Psi(t)) < c$ when $\alpha(t) \leq \alpha_n$. We use the same definitions as above for t_1^n, t_2^n, $M_k^n = (u_k^n, u'^n_k, \Psi_k^n)$ for $k = 1, 2$. Applying Lemma I.10 if necessary, we may also assume that Ψ_k^n is a ground state of $H^N(R(\alpha_n, u_k^n, u'^n_k), Z)$.

Up to a subsequence, we may assume that $u_k^n \to u_k$ and $u'^n_k \to u'_k$ as n goes to $+\infty$ and that each $(\Psi_k^n)_n$ converges to some critical point at infinity of energy c (by Theorem I.3). But by *(H1), (H3)* (let us assume for instance that $E^{|z|}(r,z)$ is not degenerated), these points can be written $(u\psi) \wedge (u'\psi')$ where ψ is a fixed ground state of $H^{|z|}(r,z)$, and $\psi' \in ker\left(H^{|z'|}(r',z') - E^{|z'|}(r',z')\right)$, $\|\psi'\|_{L^2} = 1$. So finally, we may assume that each $(\Psi_k^n)_n$ converges to a critical point at infinity of the form $(u_k\psi)\wedge(u'_k\psi'_k)$.

Step 1: Rotating the molecules. As above, the idea is now to find a path from each M_k^n to some tensor product, with a non-increasing energy. But we cannot apply the method used before since a tensor product may now have an energy which is greater than c if the molecules have a 'bad' orientation. This is due to the fact that the first term in the expansion of the interaction energy will be the dipole/dipole term. Therefore, our first step will be to change the orientation of the molecules, so that the dipole/dipole interaction energy of a tensor product becomes negative.

Let $\xi_1^n \in \mathcal{C}^\infty(\mathbb{R}^3, [0;1])$ be a cut-off function such that $\xi_1^n \equiv 0$ on $\mathbb{R}^3 \setminus B(O, \alpha_n/3)$, $\xi_1^n \equiv 1$ on $B(O, \alpha_n/3 - 1)$, $\|\nabla\xi_1^n\|_\infty \leq 1$, $\|\Delta\xi_1^n\|_\infty \leq 2$, and let us denote by $\xi_2^n = \tau_{\alpha_n\vec{v}}\xi_1^n$ the translation of ξ_1^n, and by $\xi_3^n := 1 - \xi_1^n - \xi_2^n$. We now write as in the proof of Theorem I.2

$$\Psi_k^n = \prod_{l=1}^N (\xi_1^n + \xi_2^n + \xi_3^n)(x_l)\Psi_k^n = \sum_{1\leq k_l\leq 3} \xi_{k_1}^n(x_1)\cdots\xi_{k_N}^n(x_N)\Psi_k^n := \sum_{1\leq k_l\leq 3} (\Psi_k^n)^{k_1,\ldots,k_N}.$$

Since $k \in \{1,2\}$ and n will be fixed during this step, we shall forget the subscripts and write $\Psi = \sum_{1\leq k_l\leq 3} \Psi^{k_1,\ldots,k_N}$. We now introduce for all v_1, $v_2 \in S\mathcal{O}_3(\mathbb{R})$,

$$\Psi_{(v_1,v_2)}(x_1,\ldots x_N) := \sum_{1\leq k_l\leq 3} \Psi^{k_1,\ldots,k_N}(w_{k_1}x_1,\ldots w_{k_N}x_N),$$

$$M_{(v_1,v_2)} := (\alpha_n, v_1u, v_2u', \Psi_{(v_1,v_2)})$$

where $w_1 = v_1^T$, $w_3 = v_3 = I$ and $w_2(x) = v_2^T(x - \alpha_n\vec{v}) + \alpha_n\vec{v}$. We have also forgotten the indices k and n for the sake of simplicity. Obviously, $M_{(I,I)} = M_k^n$.

Lemma I.11. *There exists a path $V : [0,1] \to (S\mathcal{O}_3(\mathbb{R}))^2$ such that $V(0) = (I,I)$, $V(1)$ is a critical point of $(v_1,v_2) \mapsto \mathcal{E}^N(M_{(v_1,v_2)})$, and $\mathcal{E}^N(M_{V(t)}) \leq \mathcal{E}^N(M_{I,I})$ for all $t \in [0;1]$.*

Proof of Lemma I.11 – Since Ψ_k^n is locally Lipschitz by Theorem I.1, one easily shows that $F : (v_1, v_2) \mapsto \mathcal{E}^N(M_{(v_1,v_2)})$ is $\mathcal{C}^2$ on $(S\mathcal{O}_3(\mathbb{R}))^2$. We then simply consider the associated gradient flow to conclude. □

Applying this lemma, we may obtain some $X = (\alpha_n, vu, v'u', \tilde{\Psi})$, connected to M_k^n by a path with a non-increasing energy, and which is a critical point with respect to the rotation of the molecules as defined before. Indeed it is a critical point of the function

$$(H_1, H_2) \mapsto \mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', (\tilde{\Psi})_{(e^{H_1},e^{H_2})}\right)$$

defined on the space $A_3(\mathbb{R}) \times A_3(\mathbb{R})$, $A_3(\mathbb{R})$ being the space of 3×3 antisymmetric real matrices.

Step 2: Expression and expansion of the energy and its derivative. Now, we have

$$|\tilde{\Psi}|^2_{(e^{H_1},e^{H_2})} = \sum_{1\le k_l\le 2} |\tilde{\Psi}^{k_1,\dots,k_N}_{(e^{H_1},e^{H_2})}|^2 + \varphi$$

with $\|\varphi\|_{H^1} \le Ce^{-a\alpha_n}$ (by Theorem I.2), so we obtain

$$\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}_{(e^{H_1},e^{H_2})}\right) = \sum_{1\le k_l\le 2} \mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}^{k_1,\dots,k_N}_{(e^{H_1},e^{H_2})}\right) + O\left(e^{-a\alpha_n}\right).$$

Let us fix $1 \le k_1, \dots, k_N \le 2$ and denote $C_l = \{i,\ k_i = l\}$, $x^1 = (x_i)_{i\in C_1}$, $x^2 = (x_i)_{i\in C_2}$, $N_l = |C_l|$. We then have

$$\begin{aligned}
\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}^{k_1,\dots,k_N}_{(e^{H_1},e^{H_2})}\right) = {}& \\
&\langle H^{N_1}(r,z)_{C_1}\Psi^{k_1,\dots,k_N}, \Psi^{k_1,\dots,k_N}\rangle + \langle H^{N_2}(r',z')_{C_2}\Psi^{k_1,\dots,k_N}, \Psi^{k_1,\dots,k_N}\rangle \\
&+ \int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2 \left(\sum_{i\in C_1,\ j\in C_2} \frac{1}{|\alpha_n\vec{v} + e^{H_2}x_j - e^{H_1}x_i|} \right. \\
&- \sum_{i\in C_1}\sum_{j=1}^{m'} \frac{z'_j}{|\alpha_n\vec{v} + e^{H_2}r_j - e^{H_1}x_i|} - \sum_{i=1}^{m}\sum_{j\in C_2} \frac{z_i}{|\alpha_n\vec{v} + e^{H_2}x_j - e^{H_1}r_i|} \\
&\left. + \sum_{i=1}^{m}\sum_{j=1}^{m'} \frac{z_i z'_j}{|\alpha_n\vec{v} + e^{H_2}r_j - e^{H_1}r_i|} \right) dx^1 dx^2.
\end{aligned}$$

Differentiating this expression with regard to $H = (H_1, H_2)$, we find

$$\begin{aligned}
d_{H=0}\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}^{k_1,\dots,k_N}_{(e^{H_1},e^{H_2})}\right)(H_1,H_2) = {}& \\
&- \int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2 \left(\sum_{i\in C_1,\ j\in C_2} \frac{e_{\alpha_n\vec{v}+x_j-x_i}\cdot(H_2x_j - H_1x_i)}{|\alpha_n\vec{v} + x_j - x_i|^2} \right.
\end{aligned}$$

$$- \sum_{i\in C_1}\sum_{j=1}^{m'} \frac{z'_j e_{\alpha_n \vec{v}+r_j-x_i}\cdot(H_2 r_j - H_1 x_i)}{|\alpha_n\vec{v}+r_j-x_i|^2} - \sum_{i=1}^{m}\sum_{j\in C_2} \frac{z_i e_{\alpha_n\vec{v}+x_j-r_i}\cdot(H_2 x_j - H_1 r_i)}{|\alpha_n\vec{v}+x_j-r_i|^2}$$
$$+ \sum_{i=1}^{m}\sum_{j=1}^{m'} \frac{z_i z'_j e_{\alpha_n\vec{v}+r_j-r_i}\cdot(H_2 r_j - H_1 r_i)}{|\alpha_n\vec{v}+r_j-r_i|^2}\Bigg) dx^1 dx^2.$$

Now, we obtain, using lemma I.2,

$$\nabla_{H_1=0}\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}^{k_1,\dots,k_N}_{(e^{H_1},e^{H_2})}\right) =$$
$$- \frac{(N_2-|z'|)}{(\alpha_n)^2}\int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2\left(\sum_{i\in C_1} x_i - \sum_{i=1}^{m} z_j r_j\right)\vec{v}^T dx^1dx^2$$
$$- \frac{(N_2-|z'|)}{(\alpha_n)^3}\int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2(I-3\vec{v}\vec{v}^T)\left(\sum_{i\in C_1} x_i x_i^T - \sum_{i=1}^{m} z_i r_i r_i^T\right)dx^1dx^2$$
$$- \frac{1}{(\alpha_n)^3}\int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2(I-3\vec{v}\vec{v}^T)\cdot$$
$$\cdot\left(\sum_{i\in C_1} x_i - \sum_{i=1}^{m} z_i r_i\right)\left(\sum_{j\in C_2} x_j - \sum_{j=1}^{m'} z'_j r'_j\right)^T dx^1dx^2 + O\left(\frac{\int_{\mathbb{R}^{3N}}|\tilde{\Psi}^{k_1,\dots,k_N}|^2}{(\alpha_n)^4}\right),$$

and a similar expansion for the gradient with regard to H_2.

We may now expand the energy with a similar method and obtain

$$\mathcal{E}^N\left(\alpha_n, v, v', \tilde{\Psi}^{k_1,\dots,k_N}\right) = \langle H^{N_1}(r,z)_{C_1}\Psi^{k_1,\dots,k_N}, \Psi^{k_1,\dots,k_N}\rangle +$$
$$\langle H^{N_2}(r',z')_{C_2}\Psi^{k_1,\dots,k_N}, \Psi^{k_1,\dots,k_N}\rangle + \frac{(N_1-Z_1)(N_2-Z_2)}{\alpha_n}\int_{\mathbb{R}^{3N}}|\tilde{\Psi}^{k_1,\dots,k_N}|^2$$
$$+ \frac{(N_2-|z'|)}{(\alpha_n)^2}e_{\vec{v}}\cdot\int_{\mathbb{R}^{3N}}|(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2\left(\sum_{i\in C_1} x_i - \sum_{i=1}^{m} z_i r_i\right)dx^1dx^2$$
$$- \frac{(N_1-|z|)}{(\alpha_n)^2}e_{\vec{v}}\cdot\int_{\mathbb{R}^{3N}}|(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2\left(\sum_{j\in C_2} x_j - \sum_{j=1}^{m'} z'_j r'_j\right)dx^1dx^2$$
$$- \frac{(N_2-|z'|)}{2(\alpha_n)^3}\int_{\mathbb{R}^{3N}}|(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2\Bigg(\sum_{i\in C_1}(I-3\vec{v}\vec{v}^T)x_i\cdot x_i$$
$$- \sum_{i=1}^{m} z_i(I-3\vec{v}\vec{v}^T)r_i\cdot r_i\Bigg)dx^1dx^2$$
$$- \frac{(N_1-|z|)}{2(\alpha_n)^3}\int_{\mathbb{R}^{3N}}|(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2\left(\sum_{j\in C_2}(I-3\vec{v}\vec{v}^T)x_j\cdot x_j\right.$$

$$- \sum_{j=1}^{m'} z'_i (I - 3\vec{v}\vec{v}^T) r'_j \cdot r'_j \Bigg) \, dx^1 dx^2$$

$$+ \frac{1}{(\alpha_n)^3} \int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n \vec{v}})_{x^2} \tilde{\Psi}^{k_1,\dots,k_N}|^2 (I - 3\vec{v}\vec{v}^T) \left(\sum_{i \in C_1} x_i - \sum_{i=1}^{m} z_i r_i \right) \cdot$$

$$\cdot \left(\sum_{j \in C_2} x_j - \sum_{j=1}^{m'} z'_j r'_j \right) dx^1 dx^2 + O\left(\frac{\int_{\mathbb{R}^{3N}} |\tilde{\Psi}^{k_1,\dots,k_N}|^2}{(\alpha_n)^4} \right).$$

With a similar computation, it can be proved that

$$|\nabla_{H=0} \langle H^N(R^n, Z) \tilde{\Psi}^{k_1,\dots,k_N}, \tilde{\Psi}^{k'_1,\dots,k'_N} \rangle| = O\left(e^{-a\alpha_n}\right)$$

when $k'_l = 3$ for some l.

As a conclusion, we obtain, by summing these expressions,

$$\mathcal{E}^N \left(\alpha_n, e^{H_1} v, e^{H_2} v', \tilde{\Psi}^n_k \right) = A_n - \frac{B_n}{\alpha_n} + \frac{C_n}{(\alpha_n)^2} + \frac{D_n + E_n}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4} \right)$$

where

$$A_n = \sum_{N_1+N_2=N} \sum_{|C_1|=N_1} \langle \left(H^{N_1}(r,z)_{C_1} + H^{N_2}(r',z')_{C_2} \right) \Psi^{C_1,C_2}, \Psi^{C_1,C_2} \rangle$$

$$B_n = - \sum_{N_1+N_2=N} \sum_{|C_1|=N_1} (N_1 - |z|)(N_2 - |z'|) \|\tilde{\Psi}^{C_1,C_2}\|^2_{L^2} \geq 0$$

and with similar definitions for C_n, D_n, E_n (D_n contains the two terms involving $(N_1 - |z|)$ and $(N_2 - |z'|)$, and E_n is the third term).

Step 3: A boot-strap argument. We have

$$\begin{aligned}
A_n &\geq \sum_{N_1+N_2=N} \left(E^{N_1}(r,z) + E^{N_2}(r',z') \right) \sum_{|C_1|=N_1} \|\tilde{\Psi}^{C_1,C_2}\|^2_{L^2} \\
&= c \sum_{|C_1|=|z|} \|\tilde{\Psi}^{C_1,C_2}\|^2_{L^2} \\
&\quad + \sum_{\substack{N_1+N_2=N \\ N_1 \neq |z|}} \left(E^{N_1}(r,z) + E^{N_2}(r',z') \right) \sum_{|C_1|=N_1} \|\tilde{\Psi}^{C_1,C_2}\|^2_{L^2} \\
&= c + O\left(e^{-a\alpha_n}\right) + \sum_{\substack{N_1+N_2=N \\ N_1 \neq |z|}} \left(E^{N_1}(r,z) + E^{N_2}(r',z') - c \right) \sum_{|C_1|=N_1} \|\tilde{\Psi}^{C_1,C_2}\|^2_{L^2} \\
&\geq c + O\left(e^{-a\alpha_n}\right) + \varepsilon \sum_{\substack{N_1+N_2=N \\ N_1 \neq |z|}} \sum_{|C_1|=N_1} \|\tilde{\Psi}^{C_1,C_2}\|^2_{L^2} \\
&\geq c + \varepsilon' B_n + O\left(e^{-a\alpha_n}\right)
\end{aligned}$$

where $\varepsilon = \min\{E^{N_1}(r,z) + E^{N_2}(r',z'),\ N_1+N_2=N,\ N_1 \neq |z|\} - c > 0$ *(H1)*. Now, we have by the definition of M_k^n and X_k^n,

$$\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}_k^n\right) < c$$

so we obtain, for n large enough,

$$0 \leq \frac{\varepsilon'}{2}B_n \leq \left(\varepsilon' - \frac{1}{\alpha_n}\right)B_n \leq -\frac{C_n}{(\alpha_n)^2} - \frac{D_n+E_n}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4}\right). \tag{I.14}$$

Using the fact that $\tilde{\Psi}_k^n$ converges to some critical point at infinity, one easily sees that A_n, B_n, C_n and D_n are bounded as n goes to $+\infty$, and finally $B_n = O\left(\frac{1}{(\alpha_n)^2}\right)$. But we may write for instance, for $1 \leq k_l \leq 2$ with $|C_1| := N_1 \neq |z|$, by $(I.14)$

$$\begin{aligned}\left|(N_2-|z'|)e_{\vec{v}} \cdot \int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2 \left(\sum_{i\in C_1} x_i - \sum_{i=1}^m z_i r_i\right) dx^1dx^2\right| \\ \leq |N_2-|z'||(N_1+|z|)\alpha_n \int_{\mathbb{R}^{3N}} |(\tau_{-\alpha_n\vec{v}})_{x^2}\tilde{\Psi}^{k_1,\dots,k_N}|^2 dx^1dx^2 \\ = \alpha_n|N_2-|z'||O\left(\|\tilde{\Psi}^{k_1,\dots,k_N}\|_{L^2}^2\right)\end{aligned}$$

and so one can easily prove

$$C_n = O\left(\alpha_n B_n\right), \quad D_n = O\left((\alpha_n)^2 B_n\right). \tag{I.15}$$

Now $C_n = O\left(\frac{1}{\alpha_n}\right)$ and by $(I.14)$, $B_n = O\left(\frac{1}{(\alpha_n)^3}\right)$. Using one more time $(I.15)$, we obtain $C_n = O\left(\frac{1}{(\alpha_n)^2}\right)$ and $D_n = O\left(\frac{1}{\alpha_n}\right)$. As a consequence,

$$\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}_k^n\right) = A_n + \frac{E_n}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4}\right).$$

Using the same estimates, we also obtain an expansion of the form

$$\nabla_{H_1=0}\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}_{(H_1,H_2)}\right) = \frac{E_n'}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4}\right) = 0$$

where E_n' corresponds to the derivative of the dipole/dipole term. But now it is known that

$$\mathcal{P}_{\mathcal{A}_3(\mathbb{R})^\perp}\left[\nabla_{H_1=0}\mathcal{E}^N\left(\alpha_n, e^{H_1}v, e^{H_2}v', \tilde{\Psi}_{(H_1,H_2)}\right)\right] = 0$$

where $\mathcal{P}_{\mathcal{A}_3(\mathbb{R})^\perp}$ is the orthogonal projection onto $\mathcal{A}_3(\mathbb{R})^\perp$. We thus have $\mathcal{P}_{\mathcal{A}_3(\mathbb{R})^\perp}E_n' = O\left(\frac{1}{\alpha_n}\right)$ and so $\mathcal{P}_{\mathcal{A}_3(\mathbb{R})^\perp}E_n' \to 0$ as n goes to $+\infty$. A same result holds for the derivative with respect to H_2. Recall now that in fact $v = v_k^n$, $v' = v'^n_k$, $u = u_k^n$ and $u' = u'^n_k$. As above we may assume by extracting a sequence if necessary that v_k^n and v'^n_k converge to some v_k and v_k' as $n \to +\infty$.

We may now pass to the limit and obtain, by $(I.14)$ and the convergence of $\tilde{\Psi}_k^n$,

$$\begin{cases} (v_ku_kP)\cdot(v_k'u_k'P_k') - 3(v_ku_kP\cdot\vec{v})(v_k'u_k'P_k'\cdot\vec{v}) \leq 0 \\ \mathcal{P}_{\mathcal{A}_3(\mathbb{R})^\perp}\left[(I-3\vec{v}\vec{v}^T)(v_ku_kP)(v_k'u_k'P_k')^T\right] = 0 \\ \mathcal{P}_{\mathcal{A}_3(\mathbb{R})^\perp}\left[(I-3\vec{v}\vec{v}^T)(v_k'u_k'P_k')(v_ku_kP)^T\right] = 0 \end{cases} \tag{I.16}$$

for $k \in \{1,2\}$, and where $P := P_\psi$ and $P_k := P_{\psi'_k}$.

Step 4: Study of the critical points of the dipole/dipole interaction. We have the following result, which is proved in Appendix 2.

Lemma I.12. *Let be $P, P' \in \mathbb{R}^3 \setminus \{0\}$ and $e \in S^2$. The critical points of the function*

$$F^e_{P,P'} : (u,u') \mapsto (uP)\cdot(u'P') - 3(uP\cdot e)(u'P'\cdot e) = ((I-3ee^T)(uP))\cdot(u'P')$$

defined on $S\mathcal{O}_3(\mathbb{R})^2$ are given by

1. *$u_0P = \varepsilon|P|e$, $u'_0P' = \varepsilon|P'|e$ with $\varepsilon = \pm 1$. Then (u_0,u'_0) is a minimum of F, and $F^e_{P,P'}(u_0,u'_0) = -2|P|\,|P'|$.*
2. *$u_4P = \varepsilon|P|e$, $u'_4P' = -\varepsilon|P'|e$ with $\varepsilon = \pm 1$. Then (u_4,u'_4) is a maximum of F with a Morse index equal to 4, and $F^e_{P,P'}(u_4,u'_4) = 2|P|\,|P'|$.*
3. *$u_1P = |P|v$, $u'_1P' = -|P'|v$ for some $v \in S^2 \cap (e)^\perp$. Then (u_1,u'_1) has a Morse index equal to 1, and $F^e_{P,P'}(u_1,u'_1) = -|P|\,|P'|$.*
4. *$u_2P = |P|v$, $u'_2P' = |P'|v$ for some $v \in S^2 \cap (e)^\perp$. Then (u_2,u'_2) has a Morse index equal to 2, and $F^e_{P,P'}(u_1,u'_1) = |P|\,|P'|$.*

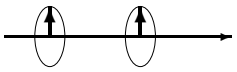
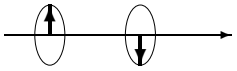
Index	Energy	
4	$2\|P\|\,\|P'\|$	
2	$\|P\|\,\|P'\|$	
1	$-\|P\|\,\|P'\|$	
0	$-2\|P\|\,\|P'\|$	

Figure I.4: The critical points of the dipole/dipole interaction

By $(I.16)$, $(v_ku_k, v'_ku'_k)$ is a critical point of $F^{\vec{v}}_{P,P'_k}$, with a non-positive energy. Since $P \neq 0$ and $P'_k \neq 0$ $(H1)$, we conclude by Lemma I.12 that $F^{\vec{v}}_{P,P'_k}(v_ku_k, v'_ku'_k) \leq -|P||P'_k| < 0$.

As a consequence, applying Lemma I.2, we obtain

$$\mathcal{E}^N\left(\alpha_n, v^n_ku^n_k, {v'}^n_k{u'}^n_k, (v^n_ku^n_k\cdot\psi)\wedge(\tau_{\alpha_n\vec{v}}\cdot {v'}^n_k{u'}^n_k\cdot\psi'_k)\right) = c + \frac{F^{\vec{v}}_{P,P'_k}(v^n_ku^n_k, {v'}^n_k{u'}^n_k)}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4}\right) < c$$

for n large enough, since

$$\lim_{n\to+\infty} F^{\vec{v}}_{P,P'_k}(v^n_k u^n_k, v'^n_k u'^n_k) = F^{\vec{v}}_{P,P'_k}(v_k u_k, v'_k u'_k) < 0.$$

Let us now denote by

$$\tilde{M}^n_k := \left(\alpha_n, v^n_k u^n_k, v'^n_k u'^n_k, (v^n_k u^n_k \cdot \psi) \wedge (\tau_{\alpha_n \vec{v}} \cdot v'^n_k u'^n_k \cdot \psi'_k)\right).$$

For n large enough, we may now construct a path connecting X^n_k to $\tilde{M}^n_k$ $k \in \{1,2\}$, with an energy below c, by Lemma I.10.

Step 5: Connecting the two tensor products. To end the proof, it remains to connect M^n_1 and M^n_2 by a path with a maximum energy below c. We fix here n and forget this subscript. Let us now introduce the function F defined by

$$F(\nu, \nu', \varphi) = F^{\vec{v}}_{P,P_\varphi}(\nu, \nu')$$

for $(\nu, \nu', \varphi) \in W := S\mathcal{O}_3(\mathbb{R})^2 \times V$ where V is the unit sphere of the finite dimensional space $ker\left(H^{|z'|}(r', z') - E^{|z'|}(r', z')\right)$. We now denote by $Y_k := (u_k v_k, u'_k v'_k, \psi_k)$ for $k \in \{1,2\}$. Notice that we have shown

$$\max(F(Y_1), F(Y_2)) \leq -\min_{\varphi\in V}\{|P||P_\varphi|\} < 0.$$

Now, let be

$$\tilde{c} := \inf_{Y\in\Gamma'} \max_{t\in[0,1]} F(Y(t)) \tag{I.17}$$

where

$$\Gamma' = \{Y \in \mathcal{C}^0([0,1], W),\ Y(0) = Y_1,\ Y(1) = Y_2\}.$$

We have the following

Lemma I.13. *We have* $\tilde{c} \leq -\min_{\varphi\in V}\{|P||P_\varphi|\} := C < 0$.

Proof of Lemma I.13 – Suppose that $\tilde{c} > max(F(Y_1), F(Y_2))$. By the methods of [Gho93], we may find a sequence $y_n \in W$ such that $F(y_n) \to \tilde{c}$, $\nabla F(y_n) \to 0$ and $d^2_{(\nu,\nu')}F \geq -e_n$ on a space of codimension at most 1, with $e_n \to 0$ as $n \to +\infty$. W being compact, up to a subsequence, $y_n = (\nu_n, \nu'_n, \varphi_n)$ converges to a critical point $y = (\nu, \nu', \varphi)$ of energy $\tilde{c}$, and with a Morse index with respect to (ν, ν') variations at most 1. By Lemma I.12, we necessary have $F(y) \leq -|P||P_\varphi|$. □

Let $Y_n \in \Gamma'$ be a path connecting Y_1 and Y_2 such that $\max_{t\in[0;1]} F(Y(t)) \leq \tilde{c} + \frac{1}{\alpha_n}$. We then have

$$\begin{aligned}\mathcal{E}^N &\left(\alpha_n, \nu_n(t), \nu'_n(t), (\nu_n(t)\cdot\psi) \wedge (\tau_{\alpha_n\vec{v}} \cdot \nu'_n(t) \cdot \psi_n(t))\right) \\ &= c + \frac{F(Y_n(t))}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4}\right) \leq c + \frac{C}{(\alpha_n)^3} + O\left(\frac{1}{(\alpha_n)^4}\right) < c\end{aligned}$$

for n large enough.

As a conclusion, we have built a path $\tilde{\gamma}_n$ for n large enough such that

$$\max_{t\in[0;1]} \mathcal{E}^N(\tilde{\gamma}_n(t)) < c$$

which is a contradiction and ends the proof of Theorem I.7. □

Appendix 1: Non-isotropic exponential decay of the electronic density

We want to prove here Lemma I.3, that we recall for the reader's convenience.

Lemma I.14. *Let Ψ_R be an eigenfunction associated to the eigenvalue $\lambda_d^N(R,Z)$ and ρ_R be the electronic density. We introduce $\varepsilon_R = \Sigma^N(R,Z) - \lambda_d^N(R,Z)$. Then*

1. *ρ_R satisfies the inequation*

$$-\frac{1}{2}\Delta\rho_R + V_R\rho_R + \varepsilon_R\rho_R \leq 0. \tag{I.18}$$

2. *With $R_1(\varepsilon) := max\left(R_0+1, R_0+\frac{2Np}{\varepsilon}\right)$ and $C(\varepsilon) := \cup_{j=1}^p\{x,\ |x-X_j| = R_1(\varepsilon)\}$, and if $r > 2R_1(\varepsilon_R)$, then we have*

$$\begin{aligned}\rho_R(x) &\leq \|\rho_R\|_{L^\infty(C(\varepsilon_R))}\sum_{j=1}^p e^{-\sqrt{\varepsilon_R/p}(|X_j-x|-R_1(\varepsilon_R))}\\ &\leq p\|\rho_R\|_{L^\infty(C(\varepsilon_R))}e^{-\sqrt{\varepsilon_R/p}(\delta(x)-R_1(\varepsilon_R))},\\ &\leq Me^{-\sqrt{\varepsilon_R/p}(\delta(x)-R_1(\varepsilon_R))}\end{aligned} \tag{I.19}$$

on $\mathcal{U}(R_1(\varepsilon_R))$, where $\delta(x) = \min\{|x-X_j|,\ j=1,...,d\}$, and $M = M(p,N,R_1(\varepsilon_R))$.

Proof – The fact that ρ_R is a solution to (I.10) is essentially proved in [HØ01] (the proof is written for atoms with no nuclei/nuclei interaction).

The proof of (I.19) follows ideas of [HO77]. We have

$$V_R(x) \geq -\sum_{j=1}^p \frac{N}{|x-X_j|-R_0}$$

for all $x \in \mathcal{U}(R_1(\varepsilon_R))$, so that ρ_R satisfies

$$-\frac{1}{2}\Delta\rho_R + \sum_{j=1}^p\left(\frac{\varepsilon_R}{p} - \frac{N}{|x-X_j|-R_0}\right)\rho_R \leq 0.$$

on $\mathcal{U}$. Let f be the radial positive solution of

$$-\frac{1}{2}\Delta f + \left(\frac{\varepsilon_R}{p} - \frac{N}{|x|-R_0}\right)f = 0$$

on $\mathbb{R}^3 \setminus B(0,R_1(\varepsilon_R))$, such that $f(x) = 1$ if $|x| = R_1(\varepsilon_R)$ (f is a Whittaker function, see [HO77]). Then $f_j(x) = f(x-X_j)$ is a positive solution of

$$-\frac{1}{2}\Delta f_j + \left(\frac{\varepsilon_R}{p} - \frac{N}{|x-X_j|-R_0}\right)f_j = 0$$

on $\mathbb{R}^3 \setminus B(X_j,R_1(\varepsilon_R))$ and $F = \sum_{j=1}^p f_j$ is a solution of

$$-\frac{1}{2}\Delta F + \sum_{j=1}^p\left(\frac{\varepsilon_R}{p} - \frac{N}{|x-X_j|-R_0}\right)F = 0.$$

such that $F(x) \geq 1$ if $x \in C(\varepsilon_R)$.

Since $\rho_R \in \mathcal{C}^\infty(\mathcal{U}(R_1(\varepsilon_R)))$ and $\lim_{|x|\to+\infty} \rho_R(x) = 0$, the maximum principle implies

$$\rho_R \leq \|\rho_R\|_{L^\infty(C(\varepsilon_R))} F(x).$$

But, on $\mathbb{R}^3 \setminus B(X_j, R_1(\varepsilon_R))$, we also have

$$\frac{\varepsilon_R}{p} - \frac{N}{|x - X_j| - R_0} \geq \frac{\varepsilon_R}{2p} > 0,$$

so that (maximum principle)

$$f_j(x) \leq \frac{R_1(\varepsilon_R)}{|X_i - x|} e^{-\sqrt{\varepsilon_R/p}(|x-X_j|-R_1(\varepsilon_R))} \leq e^{-\sqrt{\varepsilon_R/p}(|x-X_j|-R_1(\varepsilon_R))}$$

on $\mathbb{R}^3 \setminus B(X_j, R_1(\varepsilon_R))$ which implies

$$\begin{aligned} F(x) &\leq \sum_{j=1}^{p} e^{-\sqrt{\varepsilon_R/p}(|x-X_j|-R_1(\varepsilon_R))} \\ &\leq p e^{-\sqrt{\varepsilon_R/p}\left(\frac{1}{p}\sum_{j=1}^{p}|x-X_j|-R_1(\varepsilon_R)\right)} \\ &\leq p e^{-\sqrt{\varepsilon_R/p}(\delta(x)-R_1(\varepsilon_R))}. \end{aligned}$$

To end the proof, we now remark that since ρ_R is real analytic on $\mathcal{U}(R_0 + 1/2)$, there exists a constant M depending only on $R_1(\varepsilon_R)$ such that

$$\|\rho_R\|_{L^\infty(C(\varepsilon_R))} \leq M \|\rho_R\|_{L^1(\mathcal{U}(R_0+1/2)))} \leq M \|\rho_R\|_{L^1(\mathbb{R}^3)} = M \cdot N$$

and so

$$\rho_R \leq MNp e^{-\sqrt{\varepsilon_R/p}(\delta(x)-R_1(\varepsilon_R))}.$$

□

Appendix 2: Study of the critical points of the dipole/dipole interaction

We prove here Lemma I.12 that we recall.

Lemma I.15. *Let be $P, P' \in \mathbb{R}^3 \setminus \{0\}$ and $e \in S^2$. The critical points of the function*

$$F^e_{P,P'} : (u, u') \mapsto (uP) \cdot (u'P') - 3(uP \cdot e)(u'P' \cdot e) = ((I - 3ee^T)(uP)) \cdot (u'P')$$

defined on $S\mathcal{O}_3(\mathbb{R})^2$ are given by

1. *$u_0 P = \varepsilon|P|e$, $u'_0 P' = \varepsilon|P'|e$ with $\varepsilon = \pm 1$. Then (u_0, u'_0) is a minimum of F, and $F^e_{P,P'}(u_0, u'_0) = -2|P|\,|P'|$.*
2. *$u_4 P = \varepsilon|P|e$, $u'_4 P' = -\varepsilon|P'|e$ with $\varepsilon = \pm 1$. Then (u_4, u'_4) is a maximum of F with a Morse index equal to 4, and $F^e_{P,P'}(u_4, u'_4) = 2|P|\,|P'|$.*
3. *$u_1 P = |P|v$, $u'_1 P' = -|P'|v$ for some $v \in S^2 \cap (e)^\perp$. Then (u_1, u'_1) has a Morse index equal to 1, and $F^e_{P,P'}(u_1, u'_1) = -|P|\,|P'|$.*
4. *$u_2 P = |P|v$, $u'_2 P' = |P'|v$ for some $v \in S^2 \cap (e)^\perp$. Then (u_2, u'_2) has a Morse index equal to 2, and $F^e_{P,P'}(u_1, u'_1) = |P|\,|P'|$.*

Proof – It is simpler to study the function

$$(x, y) \mapsto |P||P'|x \cdot y - 3(x \cdot e)(y \cdot e)$$

defined on $S^2 \times S^2$, but we shall not use here this expression, to keep the point of view imposed by the proof of Theorem I.7.

If (u, u') is a critical point of $F^e_{P,P'}$, then $(0, 0)$ is a critical point of $f : (H, H') \mapsto F^e_{uP,u'P'}(e^H, e^{H'})$, defined on $\mathcal{A}_3(\mathbb{R}) \times \mathcal{A}_3(\mathbb{R})$. We may assume $u = u' = I$ and $|P| = |P'| = 1$ to simplify notations, and we denote by $M := I - 3ee^T$. Then we have

$$\left\{ \begin{array}{l} MP'P^T \in \mathcal{A}_3(\mathbb{R})^\perp \\ MP(P')^T \in \mathcal{A}_3(\mathbb{R})^\perp \end{array} \right. \iff \left\{ \begin{array}{l} MP'P^T = P(P')^T M \\ MP(P')^T = P'P^T M \end{array} \right. \tag{I.20}$$

Multiplying by M, we find for instance $M^2 P'P^T = P'P^T M^2$ and so $M^2 P' = (P^T M^2 P)P'$, showing that P and P' are eigenvectors of M^2. It is then easy to see that they are eigenvectors of M. Using $(I.20)$, we obtain $P = \pm P'$. The critical points are thus those given in the Lemma.

The second derivative is given by

$$d^2 f_{P,P'}(H, H') = 2(MHP) \cdot (H'P') + (MH^2 P) \cdot P' + (M(H')^2 P') \cdot P.$$

Suppose for instance that $P = e$. Let (e_2, e_3) be an orthogonal basis of $e^\perp$. Then, in the basis $(e_1 = e, e_2, e_3)$, $M = diag(-2, 1, 1)$. We introduce $H_{ij} := e_i e_j^T - e_j e_i^T$, so that $(H_{ij})_{1 \le i < j \le 3}$ is an orthogonal basis of $\mathcal{A}_3(\mathbb{R})$. The matrix of $d^2_{P,P} f$ is this basis is then

$$d^2_{P,P} f = \begin{pmatrix} 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 1 \\ 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{pmatrix} \sim \; diag(1, 1, 3, 3, 0, 0),$$

and obviously $d^2_{P,-P}f \sim diag(-1,-1,-3,-3,0,0)$.

If $P \perp e$, we use the same basis $(H_{ij})_{1\leq i<j\leq 3}$ but with $e_2 = P$ and find

$$d^2_{P,P}f = \begin{pmatrix} -1 & 0 & 0 & -2 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -2 & 0 & 0 & -1 & 0 & 0 \\ 0 & -2 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \sim \ diag(1,-2,-3,0,0,0),$$

and obviously $d^2_{P,-P}f \sim diag(-1,2,3,0,0,0)$. □

Bibliography

[Agm82] S. Agmon, *Lectures on exponential decay of solutions of second-order elliptic equations*, Princeton University Press, 1982.

[Bah89] A. Bahri, *Critical points at infinity in some variational problems*, Pitman Research Notes in Mathematics Series, vol. 182, Longman Scientific and Technical Ed., 1989.

[C+03] E. Cancès et al., *Computational quantum chemistry: a primer*, Handbook of Numerical Analysis, vol. X, pp. 3–270, Elsevier, 2003.

[CL92] I. Catto and P.-L. Lions, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part I.*, Comm. Part. Diff. Eq. **17** (1992), no. 7-8, 1051–1110.

[CL93a] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part II.*, Comm. Part. Diff. Eq. **18** (1993), no. 1-2, 305–354.

[CL93b] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part III.*, Comm. Part. Diff. Eq. **18** (1993), no. 3-4, 381–429.

[CL93c] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-fermi type theories. Part IV.*, Comm. Part. Diff. Eq. **18** (1993), no. 7-8, 1149–1159.

[FF96] J.B. Foresman and A. Frisch, *Exploring chemistry with electronic structure methods*, Gaussian Inc, 1996, Chapter 3 and references therein.

[FG92] G. Fang and N. Ghoussoub, *Morse-type information on Palais-Smale sequences obtained by min-max principles*, Manuscripta Math. **75** (1992), 81–95.

[FHØ02a] S. Fournais, M. and T. Hoffmann-Ostenhof, and T. Østergaard Sørensen, *Analyticity of the density of electronic wavefunctions*, Preprint, 2002.

[FHØ02b] ———, *The electron density is smooth away from the nuclei*, Commun. Math. Phys. **228** (2002), 401–415.

[FHØ02c] ———, *On the regularity of the density of electronic wavefunctions*, Mathematical results in quantum mechanics (Taxco, 2001), Contemp. Math., vol. 307, 2002, pp. 143–148.

[Fri03] G. Friesecke, *The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions*, Arch. Rat. Mech. Analysis **169** (2003), 35–71.

[Gho93] N. Ghoussoub, *Duality and perturbation methods in critical point theory*, Cambridge University Press, 1993.

[HO77] M. and T. Hoffmann-Ostenhof, *Schrödinger inequalities and asymptotic behavior of the electron density of atoms and molecules*, Phys. Rev. A **16** (1977), no. 5, 1782–1785.

[HØ01] M. and T. Hoffmann-Ostenhof and T. Østergaard Sørensen, *Electron wavefunctions and densities for atoms*, Ann. Henri Poincaré **2** (2001), no. 1, 77–100.

[HS00] W. Hunziker and I.M. Sigal, *The quantum N-body problem*, J. Math. Phys **41** (2000), no. 6, 3348–3509.

[Hun66] W. Hunziker, *On the spectra of Schrödinger multiparticle Hamiltonians*, Helv. Phys. Acta **39** (1966), 451–462.

[Kat51] T. Kato, *Fundamental properties of hamiltonian operators of Schrödinger type*, Trans. Am. Math. Soc **70** (1951), 195–221.

[Kat57] ———, *On the eigenfunctions of many particle systems in quantum mechanics*, Commun. Pure and Appl. Math. **10** (1957), 151–171.

[Lie84] E.H. Lieb, *Bound on the maximum negative ionization of atoms and molecules*, Phys. Rev. A. **29** (1984), no. 6, 3018–3028.

[Lie90] ———, *The stability of matter : from atoms to stars*, Bull. Amer. Math. Soc. **22** (1990), 1–49.

[Lio84] P.L. Lions, *The concentration-compactness principle in the calculus of variations. The locally compact case, part i*, Ann. Inst. Henri Poincaré **1** (1984), no. 2, 109–149.

[LS77] E.H. Lieb and B. Simon, *The Thomas-Fermi theory of atoms, molecules and solids*, Adv. Math. **23** (1977), 22–116.

[LT86] E.H. Lieb and W.E. Thirring, *Universal nature of Van Der Waals forces for coulomb systems*, Phys. Rev. A **34** (1986), 40–46.

[RS72] M. Reed and B. Simon, *Methods of modern mathematical physics*, vol. I – Functionnal Analysis, Academic Press, 1972.

[Sch87] H. B. Schlegel, *Ab inition methods in quantum chemistry*, vol. 1, pp. 249–286, Lawley Ed, 1987, and references cited therein.

[Sig82] I.M. Sigal, *Geometric methods in the quantum many-body problem. Non existence of very negative ions*, Commun. Math. Phys. **85** (1982), 309–324.

[Sim77] B. Simon, *Geometric methods in multiparticle quantum systems*, Commun. Math. Phys. **55** (1977), 259–274.

[Tel62] E. Teller, Rev. Mod. Phys. **34** (1962), 267.

[Van64] C. Van Winter, *Theory of finite systems of particles. I. The Green function*, Mat.-Fys. Skr. Danske Vid. Selsk. **2** (1964), no. 8.

[Zhi60] G. M. Zhislin, *Discussion of the spectrum of Schrödinger operators for systems of many particles. (in Russian)*, Trudy Moskovskogo matematiceskogo obscestva **9** (1960), 81–120.

Mathieu LEWIN
CEREMADE, CNRS UMR 7534,
Université Paris IX Dauphine,
Place du Maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France.
lewin@ceremade.dauphine.fr

(Submitted July 7, 2003, accepted January 27, 2004)

B.II

Sur quelques méthodes variationnelles faisant intervenir des valeurs propres de matrices symétriques.

Résumé

Dans ce chapitre, nous étudions deux principes variationnels portant sur les valeurs propres d'une matrice symétrique dépendant d'un paramètre. Ces deux méthodes sont, en dimension finie, exactement similaires à nos deux études des Parties A et B.

Dans ces deux précédents travaux, nous avons, pour des raisons de perte de régularité, modifié la définition utilisée en pratique. Grâce à des exemples et contre-exemples, nous comparons ici les deux approches.

Dans ce chapitre, nous voulons mentionner quelques résultats simples concernant des méthodes variationnelles faisant intervenir des valeurs propres de matrices symétriques.

Dans les Parties A et B de cette thèse, nous avons étudié des problèmes pratiques dans lesquels les chimistes sont amenés à considérer des minima ou des min-max de valeurs propres d'un opérateur auto-adjoint réel. Dans les deux cas, un problème de régularité nous a conduit à utiliser une définition différente de celle proposée au départ. Nous voulons dans ce chapitre comparer notre approche avec celle utilisée en pratique, en nous plaçant dans un cadre simplifié.

Nous nous limitons donc ici à l'étude de matrices, c'est-à-dire à la dimension finie, même si les problèmes mathématiques étudiés sont posés en dimension infinie. En fait, les résultats que nous allons présenter peuvent être facilement étendus à la dimension infinie, mais nous préférons rester dans ce cadre simple qui est par ailleurs bien évidemment celui des calculs numériques.

Après quelques remarques d'ordre général sur l'optimisation de valeurs propres, nous étudions un lemme du col faisant intervenir la première valeur propre d'une matrice symétrique, en liaison avec l'étude précédente des réactions chimiques. Certaines remarques peuvent être à notre avis utiles d'un point de vue numérique. Ensuite, nous donnons quelques exemples permettant de clarifier la méthode que nous avons utilisée dans la Partie A qui est reliée à la minimisation de la seconde valeur propre d'une matrice symétrique. Dans ce cadre commun, nous verrons que les deux études effectuées présentent des difficultés similaires que nous avons par ailleurs résolues différemment.

Pour faciliter la compréhension du lecteur, rappelons brièvement le cadre des deux études précédentes. L'objet principal est l'opérateur auto-adjoint réel

$$H^N(R,Z) = \sum_{i=1}^{N} \left(-\frac{1}{2}\Delta_{x_i} + V(x_i) \right) + \sum_{1\leq i<j\leq N} \frac{1}{|x_i - x_j|} + \sum_{1\leq i<j\leq M} \frac{Z_i Z_j}{|R_i - R_j|} \qquad \text{(II.1)}$$

avec

$$V(x) = -\sum_{j=1}^{M} \frac{Z_j}{|x - R_j|},$$

qui dépend des positions $R = (R_1,...,R_M)$ des noyaux de charge $Z_1,...,Z_M$. Ses valeurs propres seront notées ici $\lambda_d^N(R,Z)$.

Au Chapitre B.I, nous étudions un lemme du col entre deux minima de la première valeur propre $\lambda_1^N(R,Z)$ de $H^N(R,Z)$, les variables étant les positions des noyaux $R_1,...,R_M \in \mathbb{R}^3$.

Dans la Partie A, ces positions $R_1,...,R_M$ sont fixées et nous étudions les modèles de type multi-configurations qui permettent le calcul d'états excités. Nous définissons un certain ensemble $\mathcal{G}_D$ contenant des sous-espaces de dimension D de $H_a^1(\mathbb{R}^{3N})$ (paramétrés par des fonctions $\varphi_1,...,\varphi_K \in H^1(\mathbb{R}^3,\mathbb{R})$). Lorsque $V \in \mathcal{G}_D$, $\lambda_d(V)$ est la $d^{\text{ième}}$ valeur propre de l'opérateur $\mathcal{P}_V H^N(R,Z)\mathcal{P}_V$ où $\mathcal{P}_V$ est le projecteur orthogonal sur le sous-espace V. Nous étudions ensuite la minimisation des valeurs propres $\lambda_d(V)$ lorsque V varie dans $\mathcal{G}_d$ (nous nous intéresserons surtout ici au cas $d = 2$, c'est-à-dire au premier état excité).

Dans ces deux cas, il est très fréquent que les valeurs propres soient dégénérées au point recherché, engendrant une perte de régularité qui pose des problèmes à la fois théoriques (on n'obtient pas un point critique, donc par exemple *a priori* pas de régularité de la solution) et numériques (les algorithmes sont basés sur le calcul des dérivées).

1 Sur l'optimisation de valeurs propres

Plaçons-nous dans le cadre suivant.

Soit X un espace de Banach, et $A : X \mapsto \mathcal{S}_N(\mathbb{R})$ une application de classe $\mathcal{C}^2$. $\mathcal{S}_N(\mathbb{R})$ est l'ensemble des matrices symétriques réelles de taille $N \times N$. Nous noterons $\lambda_1(x) \leq \cdots \leq \lambda_N(x)$ les N valeurs propres de $A(x)$ pour chaque $x \in X$. Pour simplifier l'énoncé du résultat suivant, nous noterons aussi $\lambda_0(x) = -\infty$ et $\lambda_{N+1}(x) = +\infty$ pour tout x. Le Théorème 13 montrera que les fonctions λ_i sont continues. En revanche, en toute généralité, la régularité des λ_i ne peut être vraiment plus précise. Elles sont lipschitziennes si A l'est mais il n'y a aucune raison pour qu'elles soient $\mathcal{C}^1$ aux points de dégénérescence, même si A est analytique réelle par exemple.

Exemple 1: pour $A(x) = \begin{pmatrix} x & 0 \\ 0 & -x \end{pmatrix}$, on a $\lambda_1(x) = -|x|$ et $\lambda_2(x) = |x|$.

Cette perte de régularité due à la dégénérescence des valeurs propres peut poser des problèmes importants, comme nous l'avons dit. Or il existe de nombreuses situations pratiques où une optimisation de valeurs propres est nécessaire. Évidemment, toutes les situations ne sont pas identiques puisque par exemple la minimisation de λ_1 conduira rarement à une dégénérescence, alors qu'un maximum local de λ_1, ou un minimum d'une autre valeur propre sera fréquemment dégénéré (comme dans l'exemple simple ci-dessus).

Historiquement, le premier problème de cette nature remonte à Lagrange en 1773, qui voulait déterminer la forme de la colonne (à symétrie de révolution) la plus solide ayant un volume et une longueur fixés (ainsi que des conditions "de bord" à ses extrémités). Ce problème peut s'exprimer comme la maximisation de la première valeur propre d'un opérateur différentiel d'ordre 4. Comme cette valeur propre est dégénérée à l'optimum, le calcul numérique de celui-ci a été sujet à de nombreuses controverses et des algorithmes spécifiques prenant en compte cette dégénérescence ont dû être utilisés. Nous renvoyons aux travaux d'Overton *et al.* [LO96, Ove92, Ove88] pour plus de détails et d'autres exemples de tels problèmes.

Malheureusement, les travaux mathématiques de ces derniers auteurs se concentrent plutôt sur des modèles convexes ou même le plus souvent linéaires ($A(x) = \sum_k x_k A_k$), une restriction qui limite l'utilisation de ces théories et ne convient pas à nos problèmes physiques.

Pour toute la suite, nous considérons une partie connexe par arcs $C \subset X$ (souvent, une sphère par exemple). Nous supposons que A est $\mathcal{C}^2$ sur un voisinage de C, et qu'il existe m tel que $m \leq A_{|C}$, c'est-à-dire que $m \leq A(x)$ pour tout $x \in C$, au sens des matrices symétriques. Les λ_i sont alors des fonctions minorées sur C.

Avant d'étudier les deux principes de min-max mentionnés ci-dessus, rappelons quelques propriétés classiques concernant le comportement des valeurs et espaces propres en fonction du paramètre $x \in C$. Nous renvoyons à [Kat95, Chapter 2 - Section 5], [HS96, Section 15.2] et [RS78, Sections XII.1-2] pour plus de détails (plutôt pour le cas analytique). Le résultat principal est le suivant (il généralise [Kat95, Thm 5.1] qui traite le cas de $\mathbb{R}$ ou $\mathbb{C}$)

Théorème 13. *On suppose que A est de classe $\mathcal{C}^k$, $k \geq 0$. Alors les fonctions $x \in C \mapsto \lambda_i(x) \in \mathbb{R}$ sont continues, et Lipschitziennes si $k \geq 1$. Considérons un $x_0 \in C$ et $1 \leq d_1 \leq d_2 \leq N$ des entiers tels que*

$$\lambda_{d_1-1}(x_0) < \lambda_{d_1}(x_0) = \cdots = \lambda_{d_2}(x_0) < \lambda_{d_2+1}(x_0).$$

Soit alors $P(x)$ le projecteur orthogonal sur la somme directe des espaces propres associés aux valeurs propres $\lambda_{d_1}(x)$,...,$\lambda_{d_2}(x)$, défini sur un voisinage $\mathcal{V}_{x_0}$ de x_0. Alors la fonction $x \mapsto P(x)$ est de classe $\mathcal{C}^k$ sur $\mathcal{V}_{x_0}$.

Il est important de remarquer que les projecteurs spectraux ne sont pas en général continus individuellement lorsqu'il y a une dégénérescence, mais uniquement globalement, comme le montre l'exemple suivant, tiré de [Kat95, Exemple 5.3]

Exemple 2 : Posons

$$A(x) = e^{-\frac{1}{x^2}} \begin{pmatrix} \cos\frac{2}{x} & \sin\frac{2}{x} \\ \sin\frac{2}{x} & -\cos\frac{2}{x} \end{pmatrix}, \text{ et } A(0) = 0.$$

A est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$. Ses valeurs propres sont $\pm e^{-1/x^2}$. Les projecteurs associés aux deux valeurs propres sont

$$\begin{pmatrix} \cos^2\frac{1}{x} & \cos\frac{1}{x}\sin\frac{1}{x} \\ \cos\frac{1}{x}\sin\frac{1}{x} & \sin^2\frac{1}{x} \end{pmatrix} \text{ et } \begin{pmatrix} \sin^2\frac{1}{x} & -\cos\frac{1}{x}\sin\frac{1}{x} \\ -\cos\frac{1}{x}\sin\frac{1}{x} & \cos^2\frac{1}{x} \end{pmatrix}$$

qui n'ont pas de limite en 0, alors qu'évidemment la somme des deux projecteurs est l'identité.

Notons que lorsque $X = \mathbb{R}$ et A est analytique réelle, un résultat de Rellich [RS78, Theorem XII.3] fournit l'existence de fonctions β_{d_1},...,β_{d_2} analytiques sur un voisinage de x_0 et qui coïncident avec les valeurs propres ordonnées λ_{d_1},...,λ_{d_2}. Ce résultat est faux en dimension supérieure (voir Exemple 5).

2 Un lemme du col faisant intervenir λ_1

Pour ce premier exemple, nous supposons qu'il existe x_0 et x_0' dans $C \subset X$ tels que

$$\lambda_1(x_0) = \lambda_1(x_0') = \min_C \lambda_1,$$

et définissons

$$c' := \min_{\gamma' \in \Gamma'} \max_{t \in [0;1]} \lambda_1(\gamma'(t))$$

$$\Gamma' = \{\gamma' \in \mathcal{C}^0([0;1],C),\ \gamma'(0) = x_0,\ \gamma'(1) = x_0'\}.$$

Il s'agit donc d'un lemme du col dans l'esprit de A. Ambrosetti et P.H. Rabinowitz [AR73].

Voici un exemple d'une telle situation en chimie quantique.

Pour des molécules neutres ($Z = N$), il a été démontré par Lieb et Thirring [LT86] que l'infimum

$$\lambda_1^N := \min_{R=(R_i)\in\Omega} \lambda_1^N(R,Z)$$

$$\Omega := \{(R_1,...,R_M) \in (\mathbb{R}^3)^M,\ R_i \neq R_j\}$$

est atteint, prouvant la stabilité des molécules neutres. Supposons maintenant qu'il existe $R = (R_1,...,R_M)$ et $R' = (R_1',...,R_M')$ dans Ω tels que $\lambda_1^N(R,Z) = \lambda_1^N(R',Z) = \lambda_1^N$. Nous pouvons alors définir

$$c' = \min_{r\in\mathcal{R}} \max_{t\in[0;1]} \lambda_1^N(r(t),Z)$$

avec

$$\mathcal{R} = \{r \in \mathcal{C}^0([0;1],\Omega),\ r(0) = R,\ r(1) = R'\}.$$

Si le problème est posé en dimension infinie, il a toutefois la même forme que le cas général exposé plus haut.

Intuitivement, un chemin optimal s'interprète comme une réaction chimique adiabatique menant d'un minimum à l'autre. c' s'interprète comme le seuil énergétique à dépasser pour que la réaction puisse avoir lieu. Les calculs de c' et de l'optimum, appelé *état de transition*, sont très importants en chimie.

Comme indiqué précédemment, des problèmes de régularité peuvent survenir lorsque λ_1^N dégénère, ce qui trouble grandement l'efficacité des algorithmes.

Revenons maintenant au cas général. Voici tout d'abord un exemple très simple où une dégénérescence intervient au point col cherché :

Exemple 3 : Pour

$$A(x) = \begin{pmatrix} \sin x & 0 \\ 0 & -\sin x \end{pmatrix},$$

on a $\lambda_1(x) = -|\sin x|$ et $\lambda_2(x) = |\sin x|$. Le minimum de λ_1 est par exemple atteint en $x_0 = -\pi/2$ et $x_0' = \pi/2$. On voit alors aisément que $c' = 0$, atteint en $x_1 = 0$, un point où λ_1 n'est pas dérivable. x_1 n'est donc pas un point critique de λ_1.

Pour contourner cette difficulté, nous avons choisi dans notre travail sur le modèle chimique de relâcher les fonctions d'onde dans le principe de min-max. Dans le cas général, v_0 et v_0' étant deux vecteurs propres normalisés associés respectivement à $\lambda_1(x_0)$ et $\lambda_1(x_0')$, cela revient à définir

$$c_{v_0,v_0'} = \min_{\gamma \in \Gamma_{v_0,v_0'}} \max_{t\in[0;1]} E(\gamma(t)) \tag{II.2}$$

avec

$$\Gamma_{v_0,v_0'} = \{\gamma \in \mathcal{C}^0([0;1],C \times S^{N-1}),\ \gamma(0) = (x_0,v_0),\ \gamma(1) = (x_0',v_0')\}$$

$$E(x,v) = \langle A(x)v,v\rangle.$$

Dans la suite, nous noterons V_0 et V_0' les espaces propres associés respectivement à $\lambda_1(x_0)$ et $\lambda_1(x_0')$.

Remarquons que E a maintenant la régularité de A, et qu'un tel min-max donnera donc sous de bonnes hypothèses un point critique de niveau $c_{v_0,v_0'}$. Nous ne voulons pas préciser ces hypothèses, mais préférons simplement maintenant comparer $c_{v_0,v_0'}$ et c'. Une inégalité évidente est donnée par le

Lemme 2. *Pour tous vecteurs propres normalisés $(v_0,v_0') \in V_0 \times V_0'$, on a*

$$c' \leq c_{v_0,v_0'}.$$

Preuve – Soit $\gamma \in \Gamma$; on note $\gamma(t) = (x(t),v(t))$. On a alors $\lambda_1(x(t)) \leq E(\gamma(t))$, donc $c' \leq \max_t \lambda_1(x(t)) \leq \max_t E(\gamma(t))$. □

L'idée intuitive est que cette nouvelle définition peut permettre en quelque sorte de "lisser" les discontinuités de λ_1, en autorisant une plus grande variation des vecteurs propres proche de l'optimum. Reprenons l'exemple précédent :

Exemple 3 (suite) : on pose à nouveau

$$A(x) = \begin{pmatrix} \sin x & 0 \\ 0 & -\sin x \end{pmatrix},$$

$x_0 = -\pi/2$ et $x_0' = \pi/2$. On paramétrise S^1 usuellement en posant $v = (\cos\theta, \sin\theta)$, d'où l'énergie

$$E(x,\theta) = \sin x(\cos^2\theta - \sin^2\theta).$$

On voit aisément que $\nabla E(0,\theta) = 0$ si et seulement si $\theta = \pi/4\,[\pi/2]$ ce qui fournit l'existence de points critiques de E au niveau 0. On pose $\theta_0 = 0$ (i.e. $v_0 = (1,0)$) et $\theta_0' = \pi/2$ (i.e. $v_0' = (0,1)$). On relie ensuite simplement (x_0,θ_0) à (x_0',θ_0') par un segment de droite en posant $\gamma(t) = (x(t),\theta(t)) = (-\pi/2 + \pi t, \pi t/2)$, $t \in [0;1]$. On obtient alors un chemin optimal passant par le point critique $(0,\pi/4)$ de niveau $c_{v_0,v_0'} = 0$ (voir la figure II.1).

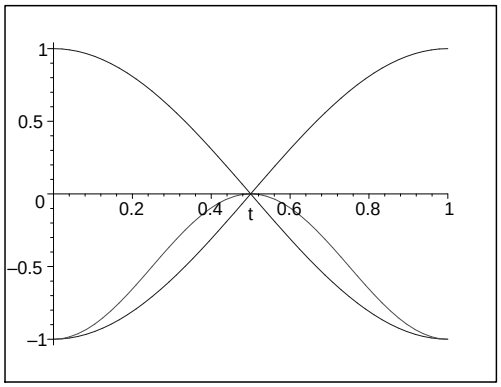


FIG. II.1 – *Les deux valeurs propres et l'énergie le long du chemin de l'exemple.*

Remarque. En fait, on peut construire très facilement des chemins optimaux pour (II.2) où $v(t)$ reste un vecteur propre associé à $\lambda_1(x(t))$ pour tout t, tout en restant continu. Pour cela, il suffit d'ajouter aux dégénérescences un petit bout de chemin sur lequel x reste constant, pour permettre à $v(t)$ de tourner et éviter ainsi les sauts. Pour l'exemple très simple proposé ci-dessus, on obtiendrait un profil de l'énergie le long du chemin comme dans la figure 2 ci-dessous.

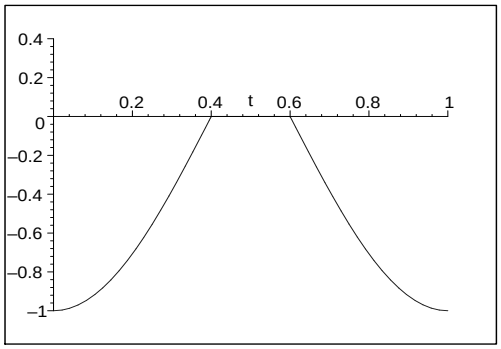


FIG. II.2 – *L'énergie le long d'un chemin où $v(t)$ reste un vecteur propre associé à $\lambda_1(x(t))$.*

Pour savoir maintenant si on a égalité ou pas dans le cas général, nous devons traiter séparément les cas de $\mathbb{R}$ et $\mathbb{C}$.

2.1 Le cas d'une matrice symétrique réelle agissant sur $\mathbb{C}^N$

Commençons tout d'abord par un cas simple, mais qui est la version "en dimension finie" du problème d'origine chimique exposé plus haut. En effet, dans ce modèle, le Hamiltonien $H^N(R,Z)$ est un opérateur **réel agissant sur** $L^2(\mathbb{R}^{3N},\mathbb{C})$. En d'autres termes, il agit séparément sur parties réelles et imaginaires et, en se ramenant à $\mathbb{R}$, on est donc conduit à étudier un opérateur de la forme $B = A \otimes A$, dont les valeurs propres sont toutes de multiplicités paires.

Dans ce cadre, on a alors le résultat suivant :

Proposition 3. *Si* $A = A' \otimes A' = \begin{pmatrix} A' & 0 \\ 0 & A' \end{pmatrix}$, *où* A' *est une fonction de classe* $\mathcal{C}^1$ *à valeurs dans* $\mathcal{S}_N(\mathbb{R})$, *alors on a pour tous vecteurs normalisés* $(v_0,v_0') \in V_0 \times V_0'$

$$c_{v_0,v_0'} = c'.$$

Preuve – Soit $\varepsilon > 0$ et $x \in \Gamma'$ un chemin vérifiant $c' \leq \max_{t\in[0;1]} \lambda_1(x(t)) \leq c' + \varepsilon$. On définit maintenant

$$c_x := \min_{v\in V} \max_{t\in[0;1]} E(x(t),v(t))$$

avec

$$V = \{v \in \mathcal{C}^0([0;1],S^{N-1}),\ v(0) = v_0,\ v(1) = v_0'\}.$$

Grâce aux méthodes développées dans [Gho93, FG92], on peut trouver une suite $(t_n,v_n) \in [0;1] \times S^{N-1}$ telle que

$$E(x(t_n),v_n) \to c_x, \quad A(x(t_n))v_n - E(x(t_n),v_n)v_n \to 0$$

et

$$E(x(t_n),v_n) \leq \lambda_1(x(t_n)) + \varepsilon_n, \quad \lim_{n\to+\infty} \varepsilon_n = 0.$$

Les deux premières assertions sont classiques (suites de Palais-Smale). La dernière provient du fait moins connu que lorsque le min-max prend en compte des déformations de chemins (donc de dimension 1), on peut alors obtenir également une information sur la dérivée seconde. Normalement, on aurait dû avoir $\lambda_2(x(t_n))$ dans l'estimation à droite, mais d'après les hypothèses, on a $\lambda_2 \equiv \lambda_1$. On a donc finalement $E(x(t_n),v_n) \leq \max_{t\in[0;1]} \lambda_1(x(t)) + \varepsilon_n \leq c' + \varepsilon + \varepsilon_n$. En passant à la limite, on obtient donc

$$c_{v_0,v_0'} \leq c_x \leq c' + \varepsilon$$

car $(x,v) \subset \Gamma_{v_0,v_0'}$, ce qui termine la preuve. □

2.2 Le cas réel

Il est facile de voir que l'égalité $c_{v_0,v_0'} = c'$ n'est pas vraie en toute généralité dans le cas réel.

Exemple 4 : Considérons $A(x) = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$ pour tout $x \in \mathbb{R}$. Evidemment, on a alors $c' = 1$. Prenons maintenant $x_0 = x_0' = 0$ et $v_0 = (1,0)$, $v_0' = (-1,0)$. Soit $\gamma = (x,v)$ un chemin quelconque de Γ. Alors nécessairement il existe un $\bar{t} \in [0;1]$ tel que $v(\bar{t}) = (0,\pm 1)$, mais alors $E(x(\bar{t}),v(\bar{t})) = 2$, ce qui montre que $c_{v_0,v_0'} = 2 > 1 = c'$.

Cet exemple montre que le choix des vecteurs propres v_0 et v_0' joue un rôle crucial. En effet, il est très facile de voir que si on avait posé $v_0'' = (1,0)$, on aurait trouvé $c_{v_0,v_0''} = c' = 1$.

En revanche, nous pouvons montrer que, modulo un bon choix des vecteurs propres de départ, les deux définitions coïncident :

Proposition 4. *Dans le cas général, on a*

$$c' = \min_{\substack{(v_0,v_0') \in V_0 \times V_0' \\ \|v_0\| = \|v_0'\| = 1}} c_{v_0,v_0'}$$

où V_0 et V_0' sont les espaces propres associés respectivement à $\lambda_1(x_0)$ et $\lambda_1(x_0')$.

Preuve – Soit $\varepsilon > 0$ et $x \in \Gamma'$ un chemin vérifiant $c' \leq \max_{t\in[0;1]} \lambda_1(x(t)) \leq c' + \varepsilon$. Nous traitons deux cas séparément.

• Premier cas : $\lambda_1(x(t))$ est non dégénérée pour tout $t \in [0;1]$. Soit alors $v_0 \in V_0$ un vecteur normalisé. Dans ce cas, grâce à la continuité des vecteurs propres, on peut trouver un chemin continu $v : [0;1] \to S^{N-1}$ menant de v_0 à un certain $v_0' \in V_0'$. On a alors $E(x(t),v(t)) = \lambda_1(x(t))$ pour tout $t \in [0;1]$, d'où $c_{v_0,v_0'} \leq \max_t E(x(t),v(t)) \leq c' + \varepsilon$.

• Second cas : $\lambda_1(x(t))$ est dégénérée pour un certain $t \in [0;1]$. Soit alors

$$K_\varepsilon = \{t \in [0;1],\ \lambda_2(x(t)) \leq \lambda_1(x(t)) + \varepsilon\}.$$

K_ε étant un compact non vide de $[0;1]$, on peut le recouvrir par une union finie de segments disjoints

$$K_\varepsilon \subset \bigcup_{i=1}^{M} [a_i;b_i] := K'$$

avec $0 \leq a_1 < b_1 < a_2 < \cdots < a_M < b_M \leq 1$, et $K' \subset K_{2\varepsilon}$. Pour simplifier, nous allons supposer que $a_1 > 0$ et $b_M < 1$ (la démonstration est tout à fait similaire sinon), et on pose alors $b_0 = 0$, $a_{M+1} = 1$.

Sur chaque $[b_i;a_{i+1}]$ pour $i = 0,...,M$, $\lambda_2(x(t)) \geq \lambda_1(x(t))+\varepsilon$, donc $\lambda_1(x(t))$ est toujours non dégénérée. On peut donc construire un chemin de vecteurs propres : pour cela, on choisit pour chaque i un vecteur propre normalisé v_{b_i} associé à $\lambda_1(x(b_i))$ et on déduit du premier cas qu'il existe chemin $v_i : [b_i;a_{i+1}] \to S^{N-1}$ tel que $A(x(t))\, v_i(t) = \lambda_1(x(t))\, v_i(t)$, $v_i(b_i) = v_{b_i}$. On pose alors $v_{a_{i+1}} := v_i(a_{i+1})$.

Ensuite, on définit sur chaque $[a_i;b_i]$

$$c_i = \min_{v\in V_i} \max_{t\in[a_i;b_i]} E(x(t),v(t))$$

$$V_i = \{\gamma \in \mathcal{C}^0([a_i;b_i],S^{N-1}),\ \gamma(a_i) = v_{a_i},\ \gamma(b_i) = v_{b_i}\}.$$

Grâce aux méthodes développées dans [FG92, Gho93], on peut montrer comme précédemment que

$$c_i \leq \max_{t\in[a_i;b_i]} \lambda_2(x(t)) \leq \max_{t\in[a_i;b_i]} \lambda_1(x(t)) + 2\varepsilon \leq c' + 3\varepsilon$$

par construction des suites (a_i) et (b_i), et par définition de x. Pour chaque i, on peut donc trouver un chemin $w_i \in V_i$ tel que $\max_{t\in[a_i;b_i]} E(x(t),w_i(t)) \leq c' + 4\varepsilon$.

En recollant les v_i et les w_i, on obtient un chemin continu $t \mapsto (x(t),v(t))$ tel que

$$\max_{t\in[0;1]} E(x(t),v(t)) \leq c' + 4\varepsilon$$

ce qui termine la preuve. □

Remarque. Dans la preuve on voit que si λ_1 dégénère au moins une fois le long du chemin $t \mapsto x(t)$, alors on peut choisir v_0 et v_0' de manière quelconque (prendre $v_{a_{M+1}} = v_0'$ au lieu de v_{b_M}). Ainsi, si tous les chemins réalisant presque le min-max passent par un point où λ_1 dégénère, alors les choix de v_0 et v_0' peuvent être quelconques et on obtiendra toujours $c' = c_{v_0,v_0'}$. C'est le cas si par exemple $\dim V_0 > 1$ ou $\dim V_0' > 1$.

3 Un min-max relié à la minimisation des λ_i

Dans la section précédente, nous avons présenté un lemme du col appliqué à λ_1 qui pouvait conduire à une dégénérescence de cette valeur propre. Comme nous l'avons dit, la minimisation d'autres valeurs propres peut également conduire très fréquemment à des dégénérescences (voir les exemples précédents). Dans cette partie, nous étudions donc des problèmes du type

$$c_i' := \min_{x \in C} \lambda_i(x), \quad i > 1.$$

Voici un exemple important :

Exemple 5 (Rellich) : On pose

$$A(x,y) = \begin{pmatrix} -\sin x & \sin y \\ \sin y & \sin x \end{pmatrix}, \qquad E(x,y,v) = \langle A(x,y)v, v \rangle.$$

On a alors $\lambda_2(x,y) = \sqrt{\sin^2 x + \sin^2 y}$, dont le minimum est atteint en $(x_0,y_0) = (0,0)$. Or on calcule $\frac{\partial E}{\partial x}(0,0,v) = -v_1^2 + v_2^2$ et $\frac{\partial E}{\partial y}(0,0,v) = 2v_1v_2$. On voit donc qu'il n'existe pas de point critique de E de la forme $(0,0,v) \in \mathbb{R}^2 \times S^1$.

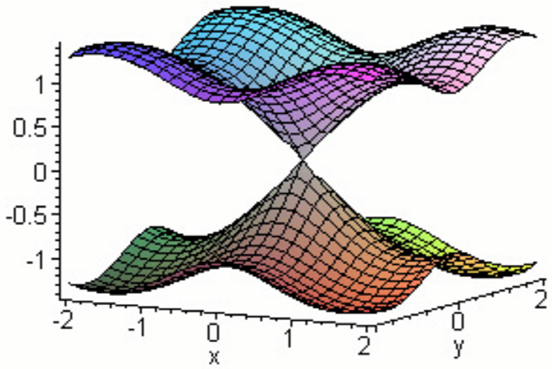


FIG. II.3 – *Les deux valeurs propres de $A(x,y)$.*

Dans cet exemple, on voit que l'étude d'un procédé de min-max portant sur l'énergie $E(x,y,v)$ ne pourra en aucun cas conduire à l'existence d'un point critique de E au niveau 0, car il n'en existe aucun. Voir la figure II.3 pour une représentation des deux valeurs propres.

Avant de présenter la méthode de min-max de niveau c_i que nous avons utilisée pour remplacer la minimisation de λ_i, nous allons maintenant donner un exemple d'une telle situation en chimie quantique.

Dans cet exemple, on considère une molécule comprenant comme précédemment M noyaux de charges Z_1,...,Z_M, dont les positions sont cette fois fixées, et N électrons. On cherche à calculer numériquement les valeurs propres (λ_i) du Hamiltonien $H^N(R,Z)$. Pour cela, on se fixe un entier K assez grand et on restreint $H^N(R,Z)$ à des espaces variables de dimension $\binom{K}{N}$. Ces espaces sont paramétrés par des K-uplets $(\varphi_1,...,\varphi_K) \in (H^1(\mathbb{R}^3,\mathbb{R}))^K$ tels que $\int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}$ en posant

$$V_{\varphi_1,...,\varphi_K} = \text{vect}(|\varphi_{i_1},...,\varphi_{i_N}\rangle,\ 1 \leq i_k \leq K).$$

Pour $\Phi = (\varphi_1,...,\varphi_K)$, on nomme ensuite H_Φ la matrice de $H^N(R,Z)$ dans la base $(|\varphi_{i_1},...,\varphi_{i_N}\rangle)_{1\leq i_k\leq K}$ et $(\lambda_i^K(\Phi))_{i=1,...,\binom{K}{N}}$ ses valeurs propres. On pose alors

$$\mu_i^K := \min\left\{\lambda_i^K(\Phi),\ \Phi \in (H^1(\mathbb{R}^3))^K,\ \int_{\mathbb{R}^3} \varphi_i\varphi_j = \delta_{ij}\right\}$$

pour $i \leq \binom{K}{N}$.

Sous cette forme, on retrouve bien pour chaque K un problème de minimisation de valeurs propres d'une matrice symétrique réelle de dimension $\binom{K}{N} \times \binom{K}{N}$.

Remarquons que l'on a

$$c_i' = \min_{x\in C} \lambda_i(x) = \min_{x\in C} \min_{A\in\mathcal{A}_i} \max_{v\in A} E(x,v)$$

$$\mathcal{A}_i = \{f(S^{i-1}),\ f \in \mathcal{C}^0(S^{i-1},S^{N-1}) \text{ impaire}\},$$

d'après le principe classique de Courant-Fisher et le théorème de Borsuk-Ulam. Ainsi, on peut écrire

$$c_i' = \min_{B\in\mathcal{B}_i} \max_{(x,v)\in B} E(x,v)$$

$$\mathcal{B}_i = \{(x,f(S^{i-1})),\ x \in C,\ f \in \mathcal{C}^0(S^{i-1},S^{N-1}) \text{ impaire}\}.$$

En d'autres termes, on effectue un min-max sur des parties B qui sont des déformations de cercles, mais où x est constant. Le problème est que $\mathcal{B}_i$ n'est pas invariant par homotopie, une hypothèse classique dans cette théorie qui est nécessaire pour démontrer l'existence d'un point critique. En d'autres termes en empêchant une variation de x, on interdit les déformations permettant de faire baisser la dérivée de E par rapport à x.

C'est pourquoi, nous avons "augmenté" l'ensemble $\mathcal{B}_i$ en ajoutant ces déformations (d'où l'inégalité $c_i \leq c_i'$), mais en conservant les propriétés d'imparité par rapport à v.

Plus précisément, on considère alors l'action du groupe $\mathbb{Z}_2$ sur $C \times S^{N-1}$ définie par $(-)\cdot(x,v) := (x,-v)$ et on munit S^{i-1} pour tout $i \leq N$ de l'action usuelle. On rappelle qu'une fonction $f : S^{i-1} \to C \times S^{N-1}$ est dite $\mathbb{Z}_2$-équivariante lorsque $f(-z) = (-)\cdot f(z)$. L'énergie E est elle $\mathbb{Z}_2$-invariante, c'est-à-dire $E(x,v) = E(x,-v)$.

On définit alors le principe de min-max classique [Gho93] dans cette situation :

$$c_i = \min_{F\in\mathcal{F}_i} \max_{(x,v)\in F} E(x,v) \tag{II.3}$$

$$\mathcal{F}_i = \{f(S^{i-1}),\ f \in \mathcal{C}^0(S^{i-1},C \times S^{N-1})\ \mathbb{Z}_2 - \text{équivariante}\}.$$

Comme annoncé, on a évidemment l'inégalité suivante :

Lemma II.1. *On a toujours $c_i \leq c_i'$.*

Preuve : Il suffit de remarquer que $\mathcal{B}_i \subset \mathcal{F}_i$. □

Comme maintenant $\mathcal{F}_i$ est invariant par les $\mathbb{Z}_2$-homotopies, on peut, sous certaines hypothèses que nous ne préciserons pas ici, démontrer l'existence d'un point critique de E de niveau c_i. Nous allons maintenant fournir deux exemples de situations pour lesquelles on obtient par exemple $c_2 < c_2'$.

Pour cela, nous devons noter que (II.3) s'écrit également

$$c_2 = \min_{(x,v)\in C\times S^{N-1}} \min_{\gamma\in\Gamma_{(x,v)}} \max_{t\in[0;1]} E(\gamma(t))$$

$$\Gamma_{(x,v)} = \left\{\gamma \in \mathcal{C}^0([0;1],C\times S^{N-1}),\ \gamma(0) = (x,v),\ \gamma(1) = (x,-v)\right\}.$$

En effet, lorsque $i = 2$, on doit déformer des cercles S^1 mais avec l'hypothèse d'invariance par rapport à v, il suffit de déformer des chemins continus menant d'un certain (x,v) à $(x,-v)$. Au premier minimum près portant sur le "point d'attache du chemin" (x,v), on voit donc que, pour le calcul de c_2, on a affaire à un principe du col du type de celui étudié dans la partie précédente.

Exemple 6 : Soit A l'application matricielle définie sur $S^1 \simeq \mathbb{R}/2\pi\mathbb{Z}$ par

$$A(\theta) = \begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix}.$$

On a alors $\lambda_1(\theta) = -1$ et $\lambda_2(\theta) = 1$ pour tout θ. Posons maintenant, pour $t \in [0;1]$, $\theta(t) = \pi/2 + 2\pi t$ et $v(t) = (\cos\pi t, \sin\pi t)$ qui est un vecteur propre associé à $\lambda_1(\theta(t))$. Si $\gamma(t) = (\theta(t),v(t))$, alors on a $\gamma(0) = (\pi/2,v_0)$ et $\gamma(1) = (\pi/2+2\pi,-v_0) = (\pi/2,-v_0)$ avec $v_0 = (1,0)$. Comme $E(\gamma(t)) = -1$ pour tout $t \in [0;1]$, on déduit que $c_2' = -1 < c_2 = 1$.

On déduit même de cet exemple que le fait que l'inégalité $c_i < c_i'$ soit stricte n'est pas lié au fait que les valeurs propres dégénèrent ou pas, puisqu'ici elles sont constantes et distinctes.

Pour compléter cet exposé, reprenons l'exemple cité au début de cette section, et calculons c_2 dans ce cas.

Exemple 5 (suite) : On pose à nouveau

$$A(x,y) = \begin{pmatrix} -\sin x & \sin y \\ \sin y & \sin x \end{pmatrix}.$$

Un point critique (x,y,v) de E vérifie les équations

$$\begin{cases} (v_2^2 - v_1^2)\cos x = 0 \\ v_1 v_2 \cos y = 0 \\ A(x,y)v = E(x,y,v)v. \end{cases}$$

On montre aisément que ces points critiques existent en les points $(\pi/2 + k\pi,\pi/2 + k'\pi)$, $(k\pi,\pi/2 + k'\pi)$ et $(\pi/2 + k\pi,k'\pi)$ (v est l'un des vecteurs propres à chaque fois). Ensuite, on relie ces points en faisant le tour du point (0,0) selon le schéma de la figure 4. Un calcul très simple montre que l'on est en fait ramené à l'exemple précédent, pour lequel on a démontré $c_2 = -1$ (prendre $v(t)$ un vecteur propre associé à $\lambda_1(x(t),y(t))$ tout le long du chemin, comme indiqué sur la figure). Finalement, on a bien $c_2 = -1 < c_2' = 0$.

On peut voir que les points critiques situés en $(\pi/2 + k\pi,\pi/2 + k'\pi)$ qui sont d'énergie $-\sqrt{2}$ correspondent à des minima globaux de λ_1, alors que les autres sont des points col

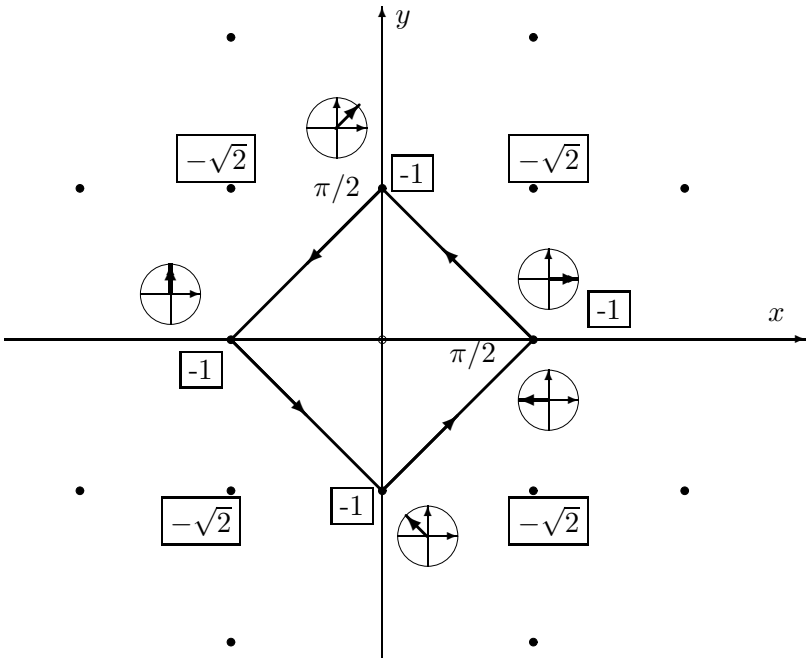


FIG. II.4 – *Le chemin construit pour démontrer que* $c_2 = -1$.

d'énergie -1. Tous les maxima globaux de λ_1 sont dégénérés et ne fournissent pas de points critiques pour E.

Avec le min-max (II.3), nous avons donc obtenu un point col de λ_1, qui n'a rien à voir avec les minima de λ_2, étudiés au départ. Ceci n'est pas très étonnant vu que la minimisation de λ_2 ne peut fournir aucun point critique.

Bibliographie

[AR73] A. Ambrosetti and P.H. Rabinowitz, *Dual variation methods in critical point theory and applications*, J. Functional Analysis **14** (1973), 349–381.

[FG92] G. Fang and N. Ghoussoub, *Morse-type information on Palais-Smale sequences obtained by min-max principles*, Manuscripta Math. **75** (1992), 81–95.

[Gho93] N. Ghoussoub, *Duality and perturbation methods in critical point theory*, Cambridge University Press, 1993.

[HS96] P.D. Hislop and I.M Sigal, *Introduction to spectral theory with application to Schrödinger operators*, Springer (New York), 1996.

[Kat95] T. Kato, *Perturbation theory for linear operators*, second ed., Springer, 1995.

[LO96] A.S. Lewis and M.L. Overton, *Eigenvalue optimization*, Acta Num. **5** (1996), 149–190.

[LT86] E.H. Lieb and W.E. Thirring, *Universal nature of Van Der Waals forces for coulomb systems*, Phys. Rev. A **34** (1986), 40–46.

[Ove88] M.L. Overton, *On minimizing the maximum eigenvalue of a symmetric matrix*, SIAM J. Matrix Anal. Appl. **9** (1988), 256–268.

[Ove92] ———, *Large scale optimization of eigenvalues*, SIAM J. Optim. **2** (1992), 88–120.

[RS78] M. Reed and B. Simon, *Methods of modern mathematical physics*, vol. IV – Analysis of Operators, Academic Press, 1978.

Troisième partie

La prise en compte des phénomènes relativistes : le modèle Bogoliubov-Dirac-Fock

C.I

Stability of the Polarized Vacuum in the Bogoliubov-Dirac-Fock approximation

Ce chapitre reprend le texte intégral d'un article écrit en collaboration avec Christian Hainzl et Éric Séré. Il a fait l'objet d'une prépublication au Ceremade.

Résumé

D'après les idées de Dirac, on doit imaginer que le vide est une "mer" d'électrons virtuels occupant les énergies négatives du spectre de l'operateur de Dirac libre. Sa charge (infinie) est supposée non mesurable. Cependant, lorsque l'on introduit un champ extérieur, ces électrons virtuels réagissent et le vide se polarise.

Dans cet article, nous étudions le modèle Bogoliubov-Dirac-Fock qui est obtenu à partir de l'électrodynamique quantique, et a été introduit par Chaix et Iracane [CI89]. L'énergie BDF tient compte de la polarisation du vide, et est bornée inférieurement. Un vide BDF-stable est par définition un minimimum de cette énergie. S'il existe, un tel minimimum est solution d'une équation auto-consistante.

Nous démontrons ici l'existence d'un minimimum de l'énergie BDF en présence d'un champ extérieur électrostatique, grâce à un théorème de type point fixe. Ce minimum est interprété comme le vide polarisé.

Stability of the Polarized Vacuum in the Bogoliubov-Dirac-Fock approximation

Christian Hainzl[a,b], Mathieu Lewin[a] & Éric Séré[a]

Abstract

According to Dirac's ideas, the vacuum consists of infinitely many virtual electrons which completely fill up the negative part of the spectrum of the free Dirac operator D^0. In the presence of an external field, these virtual particles react and the vacuum becomes polarized.

In this paper, following Chaix and Iracane [CI89], we consider the Bogoliubov-Dirac-Fock model, which is derived from QED. The corresponding BDF-energy takes the polarization of the vacuum into account and is bounded from below. A BDF-stable vacuum is defined to be a minimizer of this energy. If it exists, such a minimizer is a solution of a self-consistent equation.

We show the existence of a minimizer of the BDF-energy in the presence of an external electrostatic field, by means of a fixed-point approach. This minimizer is interpreted as the polarized vacuum.

1 Introduction

The relativistic quantum theory of electrons and positrons is based on the free Dirac operator, which is defined by

$$D^0 = -i\sum_{k=1}^{3} \alpha_k \partial_k + \beta := -i\boldsymbol{\alpha}\cdot\nabla + \beta \tag{I.1}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and

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

with

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

We follow here the notation of Thaller's book [Tha92], and of [BBHS99]. We have chosen a system of units such that $\hbar = c = 1$, and also such that the mass m_e of the electron is normalized to 1.

The operator D^0 acts on 4-spinors, i.e. functions $\Psi \in \mathcal{H} := L^2(\mathbb{R}^3, \mathbb{C}^4)$. It is self-adjoint on $\mathcal{H}$, with domain $H^1(\mathbb{R}^3, \mathbb{C}^4)$ and form domain $H^{1/2}(\mathbb{R}^3, \mathbb{C}^4)$. Moreover, it is defined to ensure

$$(D^0)^2 = -\Delta + 1.$$

The spectrum of D^0 is $(-\infty; -1] \cup [1; \infty)$. The fact that the spectrum of D^0 is not bounded from below is the main reason for all the problems that occur in this theory. To explain why a free electron does not dissolve into the lower continuum, Dirac's idea was to postulate that the vacuum contains infinitely many virtual electrons which completely fill up the negative part of the spectrum of D^0. This model of the vacuum is called the *Dirac Sea*.

In what follows, the projector associated with the negative part of the spectrum of D^0 will be denoted by P^0:

$$P^0 := \chi_{(-\infty;0)}(D^0).$$

We then have

$$D^0 P^0 = P^0 D^0 = -\sqrt{1-\Delta} P^0 = -P^0 \sqrt{1-\Delta},$$
$$D^0(1-P^0) = (1-P^0)D^0 = \sqrt{1-\Delta}(1-P^0) = (1-P^0)\sqrt{1-\Delta},$$

and

$$\mathcal{H} = \mathcal{H}^0_- \oplus \mathcal{H}^0_+,$$

where $\mathcal{H}^0_- := P^0\mathcal{H}$ and $\mathcal{H}^0_+ := (1-P^0)\mathcal{H}$.

According to what we said above, we can identify P^0 with the vacuum since it projects onto the Dirac sea. It is often called the *bare vacuum* [CI89].

Let us now add the potential created by some nuclei of total charge Z. In our system of units, the Dirac operator with external potential φ is

$$D^{\alpha\varphi} := D^0 - \alpha\varphi, \tag{I.2}$$

where $\varphi = Zn * \frac{1}{|\cdot|}$ is the Coulomb potential created by the smeared out nuclei of total density $Zn \geq 0$ such that $\int_{\mathbb{R}^3} n = 1$. The constant α is often called the *Sommerfeld fine structure constant*, and its physical value is approximately $\frac{1}{137}$. Nevertheless in the present paper it will rather play the role of a small dimensionless coupling parameter.

Dirac postulated that the charge of the Dirac sea is not measurable. However, in the presence of the external field created by the nuclei, the virtual electrons should react, by occupying the negative energy state of an other Dirac operator: *the vacuum is polarized*. This polarization of the *dressed vacuum*, which takes the form of a local density of charge, is measurable in practice. Our main goal in this paper is to study a model for the polarized vacuum which was derived by Chaix and Iracane [CI89] from no-photon Quantum Electrodynamics .

It is natural to expect that the vacuum should be described by the projector $P^{\alpha\varphi}$ associated with the negative part of the spectrum of $D^{\alpha\varphi}$ (this choice is called the *Furry picture*). We shall see that this is only an approximation: we describe below a more exact model.

Mathematically speaking, we shall say that a *vacuum* is an orthogonal projector P with the additional requirement that $Q = P - P^0$ is in $\mathfrak{S}_2(\mathcal{H})$, the space of all Hilbert-Schmidt operators on $\mathcal{H}$. As explained in Appendix, this condition guarantees the existence of a dressed vacuum in the Fock space defined with respect to the splitting $\mathcal{H} = \mathcal{H}^0_- \oplus \mathcal{H}^0_+$, but we shall not give much more detail here.

Since the model takes the free case as reference, it can be seen using the formalism of Quantum Electrodynamics (further details are given in the Appendix) that the polarization of a dressed vacuum P can be described by the difference

$$Q = P - P^0$$

which is interpreted as the *one-body density matrix of the vacuum*. This means that the total Hamiltonian is now the formal operator

$$D_Q := D^{\alpha\varphi} + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}. \tag{I.3}$$

The two terms $\alpha\rho_Q * \frac{1}{|\cdot|}$ and $\alpha\frac{Q(x,y)}{|x-y|}$ are called respectively the non-exchange and exchange potentials and they vanish when $P = P^0$. The function

$$\rho_Q(x) := \mathrm{Tr}_{\mathbb{C}^4}\left(Q(x,x)\right)$$

is the *charge density* of the vacuum P. Remark that since we only assume Q to be Hilbert-Schmidt, ρ_Q is not *a priori* a well-defined function, but an adequate framework will be given later.

Let us now explain how this P is chosen. According to Dirac's ideas, a "correct" P should be the projector associated with the negative part of the spectrum of the one-body Hamiltonian of the system. Therefore, it should be a solution of the *self-consistent equation*

$$\boxed{P = \chi_{(-\infty;0)}(D_Q) = \chi_{(-\infty;0)}\left(D^{\alpha\varphi} + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}\right).} \tag{I.4}$$

Remark that if $\varphi = 0$ (no external potential), then P^0 is already a solution to this equation since $Q = 0$ in this particular case.

This equation can be interpreted as the Euler-Lagrange equation associated with the minimization of the Bogoliubov-Dirac-Fock (BDF) energy defined by Chaix and Iracane [CI89] (see also [Cha90]). The study of this functional was our original motivation for solving (I.4). In this paper, we first give a rigorous meaning to formula ($I.4$) and then show the existence of a solution by a fixed-point argument. We then prove that our solution is a minimizer of the BDF energy.

Notice that the use of a fixed-point method to solve a self-consistent equation is very common in quantum chemistry and physics. The numerical algorithms used in practice are based on this idea. For a mathematical existence result using the Schauder fixed-point theorem, see the resolution of the Hartree equations in [Wol73]. For the determination of projectors describing the vacuum, the fixed-point approach has been used for the first time by E.H. Lieb and H. Siedentop [LS00]. We use the Banach fixed-point theorem as in [LS00], but our model is different and the necessary estimates are much more delicate.

A model also inspired by [CI89] was studied by V. Bach, J.-M. Barbaroux, B. Helffer and H. Siedentop in [BBHS99] (see also [BRS01]) . Although our model coincide with [BBHS99] when there is no external potential (i.e. $Z = 0$), it is very different in the presence of an external electrostatic field (i.e. $Z \neq 0$). Indeed, the authors of [BBHS99] have neglected the terms describing the polarization of the vacuum and, as this is explained in [CI89], the resulting energy is not bounded from below. Therefore a max-min procedure inspired by [Mit81] was considered, leading to the solution $P = P^{\alpha\varphi}$. Since we keep the

vacuum polarization terms, the BDF energy is bounded from below and we can define the polarized vacuum as the minimum of this energy. We find a P which solves (I.4) and is therefore different from $P^{\alpha\varphi}$.

When studying equation (I.4), a first problem occurs with the definition of the density ρ_Q. Let us explain this fact in the Furry picture, that is to say when P is chosen to be $P^{\alpha\varphi}$. Remark that $P^{\alpha\varphi}$ is the projector which is obtained after the first iteration of the fixed-point algorithm if we start at P^0.

It is known since the very beginning of QED [Dir34, Hei34, FO34, Ueh35, Ser35] that the density $\rho^{\alpha\varphi}$ associated with $Q^{\alpha\varphi} = P^{\alpha\varphi} - P^0$ is *never well-defined*, since it diverges pointwise logarithmically. In physics literature [FW49, Sch49, Dys49a, Dys49b] (see also the books [IZ80, Mil94, Wei96]), a procedure called *charge renormalization* aims at extracting the main information from $\rho^{\alpha\varphi}$ and a renormalized density $\rho^{\alpha\varphi}_{\rm ren}$ is used to replace the ill-defined density in the Hamiltonian

$$D^{\rm ren}_{Q^{\alpha\varphi}} = D^{\alpha\varphi} + \alpha\rho^{\alpha\varphi}_{\rm ren} * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}.$$

This procedure has been recently clarified by C. Hainzl and H. Siedentop in [HS03], where it is in addition shown that $D^{\rm ren}_{Q^{\alpha\varphi}}$ is then a well defined self-adjoint operator, under some reasonable assumption on φ. Some interesting features of $\rho^{\alpha\varphi}_{\rm ren}$, in the case of strong external fields, were obtained by Hainzl in [Hai03]. We do not want to give a precise definition of $\rho^{\alpha\varphi}_{\rm ren}$ here and we refer the reader to [HS03, Hai03, MPS98].

For the same reason as for $\rho^{\alpha\varphi}$, it can be seen that if P is a solution to equation $(I.4)$, then ρ_Q is necessarily logarithmically divergent. However our approach is different from [HS03] since we do not renormalize ρ_Q *a priori*: the interpretation of the solution as a minimizer of the BDF energy becomes unclear if a renormalization is used. Instead, we impose a momentum cut-off. In other words, from now on, we replace the ambiant space $\mathcal{H}$ by

$$\mathcal{H}_\Lambda := \left\{ f \in \mathcal{H},\ supp(\widehat{f}) \subset B(0,\Lambda) \right\}.$$

Since D^0 is a multiplication operator in Fourier space, $\mathcal{H}_\Lambda$ is invariant under P^0 and we keep the notation P^0 to avoid the introduction of a restricted operator.

As we shall see it later on, our results will be valid under a condition of the form $\alpha\sqrt{\ln\Lambda} \leq C$ for some constant $C \gtrsim 0.33$. If α has the physical value $1/137$, this leads to an extremely large Λ, which corresponds to scales that are far beyond the reach of experimental and theoretical physics at the present time.

We would like to emphasize that in QED calculations (see, e.g., [IZ80, Equation (7.18)]) a momentum cut-off is necessary. It appears in the *charge renormalization* which consists of a replacement of the *bare* charge e by $e^2_{\rm eff} \simeq e^2(1 - \frac{2\alpha}{3\pi}\ln\Lambda)$. This charge renormalization is the physical explanation for the renormalization of $\rho^{\alpha\varphi}$ mentioned above.

The paper is organized as follows. In the next section, we define the Bogoliubov-Dirac-Fock model and state our main results. For the sake of clarity, we have brought all the proofs together in Sections 3 and 4. In the Appendix, we explain how the BDF energy can be deduced from no-photon QED.

Acknowledgement: *C.H. wishes to thank Heinz Siedentop for suggesting him the possibility of studying a self-consistent model of the polarized vacuum, during his time as post-doc at the LMU (Munich). The authors acknowledge support from the European*

Union's IHP network Analysis and Quantum *HPRN-CT-2002-00277. E.S. acknowledges support from the Institut Universitaire de France.*

2 Model and main results

In this section, we study the Bogoliubov-Dirac-Fock model introduced in [CI89, Cha90]. Our system of notation is similar to [BBHS99], with the difference that we keep all the terms describing the vacuum polarization. As explained in [CI89], although these terms are negligible in most practical atomic computations, they are essential if one wants to deal with an energy which is *bounded from below* and *intrinsic*. By intrinsic, we mean that it does not depend on the choice of a particular normal ordering of the second-quantized Hamiltonian (see the Appendix for details). Remark also that the vacuum polarization terms play an essential role within the treatment of muonic atoms [GMR85].

2.1 Supertrace-class operators

In order to define the BDF-functional properly, we introduce the concept of supertrace-class operators, in the spirit of [Tha92, Section 5.7]. In this section only, we work in an abstract Hilbert space $\mathcal{H}$.

Definition I.1. *Let P be a projector such that P and $1-P$ have infinite rank, and $A \in \mathfrak{S}_2(\mathcal{H})$. We shall say that A is* supertrace-class with respect to P *if and only if $A_{++} := (1-P)A(1-P)$ and $A_{--} := PAP$ are trace-class. Then we define*

$$\mathrm{str}_P(A) := \mathrm{tr}(A_{++}) + \mathrm{tr}(A_{--}).$$

We denote by $\mathfrak{S}_1^P(\mathcal{H})$ the set of all Hilbert-Schmidt operators which are supertrace-class with respect to P.

Remark that if A is a trace-class operator, then $A \in \mathfrak{S}_1^P(\mathcal{H})$ and $\mathrm{tr}(A) = \mathrm{str}_P(A)$ for any projector P. We want to mention the following result, which will be useful in the proofs of our main theorems, and which is proved in Section 3.

Theorem I.1. *Let P and P' be two projectors such that $P - P' \in \mathfrak{S}_2(\mathcal{H})$. Then A is supertrace-class with respect to P if and only if it is supertrace-class with respect to P', and in this case* $\mathrm{str}_P(A) = \mathrm{str}_{P'}(A)$.

2.2 The Bogoliubov-Dirac-Fock model

The Bogoliubov-Dirac-Fock energy is defined by

$$\mathcal{E}(\Gamma) = \mathrm{str}_{P^0}(D^0\Gamma) - \alpha\int \rho_\Gamma\varphi + \frac{\alpha}{2}\iint \frac{\rho_\Gamma(x)\rho_\Gamma(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\Gamma(x,y)|^2}{|x-y|}dx\,dy \quad \text{(I.5)}$$

on the set

$$\mathcal{B}_\Lambda := \left\{\Gamma \in \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda),\ -P^0 \le \Gamma \le 1-P^0,\ \rho_\Gamma \in \mathcal{C}\right\}$$

with

$$\mathcal{C} = \left\{\rho \in L^2(\mathbb{R}^3),\ \int_{\mathbb{R}^3} \frac{|\widehat{\rho}(k)|^2}{|k|^2}\,dk < \infty\right\}, \qquad \|\rho\|_{\mathcal{C}}^2 := \int_{\mathbb{R}^3} \frac{1+|k|^2}{|k|^2}|\widehat{\rho}(k)|^2\,dk.$$

In (I.5), ρ_Γ is the well-defined diagonal of $\Gamma \in \mathfrak{S}_2(\mathcal{H}_\Lambda)$

$$\rho_\Gamma(x) = \mathrm{Tr}_{\mathbb{C}^4}\left(\Gamma(x,x)\right) = \frac{1}{(2\pi)^3}\iint_{|p|,|q|\leq\Lambda} \mathrm{Tr}_{\mathbb{C}^4}\left(\widehat{\Gamma}(p,q)\right)e^{ix(p-q)}dp\,dq \tag{I.6}$$

so that

$$\widehat{\rho}_\Gamma(k) = \frac{1}{(2\pi)^{3/2}}\int_{|p|\leq\Lambda} \mathrm{Tr}_{\mathbb{C}^4}\left(\widehat{\Gamma}(p+k/2,p-k/2)\right)dp$$

showing that $\widehat{\rho}_\Gamma \in L^1$, and so $\rho_\Gamma \in \mathcal{C}^0_0$. Notice that (I.6) shows that the function $\Gamma \in \mathfrak{S}_2(\mathcal{H}_\Lambda) \mapsto \rho_\Gamma \in \mathcal{C}^0_0 \cap L^2(\mathbb{R}^3)$ is continuous.

In the next subsection, we prove that $\mathcal{E}$ is well-defined and bounded from below on $\mathcal{B}_\Lambda$. We then show the existence of a minimizer, which can be interpreted as a stable polarized vacuum in the BDF approximation.

To this end, let us first explain how this energy is used in [CI89]. In the BDF setting, a state of the system is a pair (P,γ) where

- P is a projector such that $Q = P - P^0 \in \mathfrak{S}_2(\mathcal{H}_\Lambda)$, representing the dressed vacuum,
- $\gamma \in \mathfrak{S}_1(\mathcal{H}_\Lambda)$ is the density matrix of a BDF-state built with electrons and positrons defined by P. This means that

  $$\gamma = \gamma_+ - \gamma_-$$

  where γ_+ and γ_- are projectors of finite rank with $\mathrm{Ran}(\gamma_+) \subset (1-P)\mathcal{H}_\Lambda$ and $\mathrm{Ran}(\gamma_-) \subset P\mathcal{H}_\Lambda$. It therefore satisfies $-P \leq \gamma \leq 1-P$.

Details are given in Appendix for the interested reader. The density matrix γ corresponds to Slater determinants involving electrons and positrons defined by means of the vacuum P. The energy of such a state (P,γ) is then given by

$$\mathcal{E}^{\mathrm{BDF}}(P,\gamma) = \mathcal{E}(Q+\gamma) \tag{I.7}$$

where we recall that $Q = P - P^0$. Remark that the real physical object is not really the pair (P,γ), but the sum $\Gamma = Q + \gamma = P - P^0 + \gamma$. It fulfills $-P^0 \leq \Gamma \leq 1 - P^0$. Therefore the set $\mathcal{B}_\Lambda$, on which $\mathcal{E}$ is defined, is seen to be the convex hull of the BDF states taking the special form $P - P^0 + \gamma$. For such models, it is a very common idea to extend the functional to the convex hull of the states under consideration (see for instance the notion of quasi-free states defined in [BLS94] and used in [BBHS99]). To explain why $\mathcal{B}_\Lambda \subset \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda)$, we now state the following

Lemma I.1. *Let P be a projector on $\mathcal{H}_\Lambda$, such that $Q = P - P^0$ is a Hilbert-Schmidt operator. Then Q is supertrace-class with respect to P^0. Moreover, $\mathrm{str}_{P^0}(Q)$ is an integer which satisfies*

$$\forall n \geq 1,\ \mathrm{str}_{P^0}(Q) = \mathrm{tr}(Q^{2n+1}).$$

Proof – We have $Q^2 = P - PP^0 - P^0P + P^0 = (1-P^0)Q(1-P^0) - P^0QP^0$. This implies that $(1-P^0)Q(1-P^0)$ and $-P^0QP^0$ are non-negative trace-class operators.

We now use the proof of [ASS94, Theorem 4.1]. Since $Q \in \mathfrak{S}_2$, we infer $Q^3 \in \mathfrak{S}_1$ and so (P,P^0) is a Fredholm pair, in the language of [ASS94]. Therefore, $\mathrm{tr}(Q^3)$ is an integer and satisfies $\mathrm{tr}(Q^3) = \mathrm{tr}(Q^{2n+1})$ for all $n \geq 1$. Now we have

$$(P-P^0)^3 = (P-P^0)^2P - (P-P^0)^2P^0 = P(P-P^0)P + P^0(P-P^0)P^0$$

Applying this result to $1-P$ and $1-P^0$, we find

$$(P-P^0)^3 = (1-P)(P-P^0)(1-P) + (1-P^0)(P-P^0)(1-P^0).$$

Summing up this two identities, we obtain by Theorem I.1

$$2\,\mathrm{tr}(Q^3) = \mathrm{str}_P(Q) + \mathrm{str}_{P^0}(Q) = 2\,\mathrm{str}_{P^0}(Q) \quad \square$$

Remark: Since $\mathrm{str}_{P^0}(Q)$ is an integer, it can be interpreted as the charge of the dressed vacuum P (see [Hai03] for results in this direction).

In Chaix-Iracane [CI89], formula (4.8), the expression $\mathcal{E}(Q+\gamma)$ is expanded to get

$$\begin{aligned}\mathcal{E}^{\mathrm{BDF}}(P,\gamma) &= \mathrm{tr}\left(\left(D^{\alpha\varphi} + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|x-y|}\right)\gamma\right) \\ &\qquad + \frac{\alpha}{2}\iint \frac{\rho(x)\rho(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint\frac{|\gamma(x,y)|^2}{|x-y|}dx\,dy + \mathcal{E}(Q) \qquad \text{(I.8)} \\ &= \mathrm{tr}\,(D_Q\gamma) + \frac{\alpha}{2}\iint \frac{\rho(x)\rho(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint\frac{|\gamma(x,y)|^2}{|x-y|}dx\,dy + \mathcal{E}(Q),\end{aligned}$$

since $\gamma \in \mathfrak{S}_1(\mathcal{H}_\Lambda)$ which implies $\int_{\mathbb{R}^3}\left(\rho_Q * \frac{1}{|\cdot|} - \varphi\right)\rho = \mathrm{tr}\left[\left(\rho_Q * \frac{1}{|\cdot|} - \varphi\right)\gamma\right]$.

In [BBHS99, Formula (21)], the polarization potentials appearing in D_Q and the energy of the vacuum $\mathcal{E}(Q)$ were neglected by the authors who used the following functional

$$\mathcal{E}_P^{[\mathrm{BBHS99}]}(\gamma) = \mathrm{tr}\,(D^{\alpha\varphi}\gamma) + \frac{\alpha}{2}\iint\frac{\rho(x)\rho(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint\frac{|\gamma(x,y)|^2}{|x-y|}dx\,dy,$$

with the constraint $-P \le \gamma \le 1-P$. This energy is easily seen to be non intrinsic and unbounded from below: $\inf_P \inf_{-P\le\gamma\le 1-P}\mathcal{E}_P^{[\mathrm{BBHS99}]}(\gamma) = -\infty$ (see the properties of the Dirac-Fock functional defined in [ES99]). Therefore, a procedure which takes the form $\sup_P \inf_{-P\le\gamma\le 1-P}\mathcal{E}_P^{[\mathrm{BBHS99}]}(\gamma)$ inspired by [Mit81] was considered in [BBHS99], leading to the solution $P = P^{\alpha\varphi}$. In fact, as explained in [CI89], the vacuum polarization potentials and the energy of the vacuum $\mathcal{E}(Q)$ are unavoidable if one wants to deal with a bounded-below and intrinsic energy.

2.3 Study of the BDF energy

We are now able to state the following result (recall that $\varphi = Zn * \frac{1}{|\cdot|}$).

Theorem I.2. *Let be $n \in \mathcal{C}$ and $Z \in \mathbb{R}$.*

1. *$\mathcal{E}$ is well-defined on*

$$\mathcal{A}_\Lambda := \left\{\Gamma \in \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda),\ \rho_\Gamma \in \mathcal{C}\right\} \supset \mathcal{B}_\Lambda.$$

2. *If $0 \le \alpha \le \frac{4}{\pi}$, then $\mathcal{E}$ is bounded from below on $\mathcal{B}_\Lambda$, independently of Λ.*

3. *If $0 \le \alpha \le \frac{4}{\pi}$ and $Z = 0$, then $\mathcal{E}$ is non-negative on $\mathcal{B}_\Lambda$ [BBHS99], 0 being the unique minimizer.*

Without external potential, the non-negativity of the functional $\mathcal{E}$ was first proved by Chaix, Iracane, and Lions [CIL89] and later, with improvement in the parameter range, stated by Bach, Barbaroux, Helffer, and Siedentop [BBHS99]. This result is optimal in the sense that the functional becomes unbounded from below when $Z = 0$: $\inf_{\Gamma\in\mathcal{B}_\Lambda} \mathcal{E}(\Gamma) = -\infty$ if $\alpha > 4/\pi$ as shown by Hundertmark, Röhrl, and Siedentop [HRS00].

Proof of Theorem I.2 – Let us first show that $\mathcal{E}$ is well defined on $\mathcal{A}_\Lambda$. If $\Gamma \in \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda)$, then we have $P^0D^0\Gamma P^0 = D^0P^0\Gamma P^0 = D^0\Gamma_{++} \in \mathfrak{S}_1(\mathcal{H}_\Lambda)$ since P^0 commutes with D^0 and $|D^0| \leq \sqrt{1+\Lambda^2}$. With a similar argument for $1-P^0$, we obtain that $D^0\Gamma \in \mathfrak{S}_1^{P^0}(\mathcal{H}_\Lambda)$. Therefore, $\mathrm{str}_{P^0}(D^0\Gamma)$ is well-defined and

$$\mathrm{str}_{P^0}(D^0\Gamma) = \mathrm{tr}(D^0\Gamma_{++}) + \mathrm{tr}(D^0\Gamma_{--}) = \mathrm{tr}(|D^0|\Gamma_{++}) - \mathrm{tr}(|D^0|\Gamma_{--}). \tag{I.9}$$

On the other hand we have by Hardy's inequality

$$\iint \frac{|\Gamma(x,y)|^2}{|x-y|}dx\,dy \leq \frac{\pi}{2}\,\mathrm{tr}(|D^0|\Gamma^2)$$

showing that this last term is well-defined since $|D^0|$ is bounded on $\mathcal{H}_\Lambda$ and $\Gamma \in \mathfrak{S}_2(\mathcal{H}_\Lambda)$. Finally, let us introduce $D(f,g) := \iint \frac{f(x)g(y)}{|x-y|}dx\,dy$ which defines a scalar product on $\mathcal{C}$. We have

$$\int \rho_\Gamma\varphi = D(\rho_\Gamma, Zn) \quad \text{and} \quad \iint \frac{\rho_\Gamma(x)\rho_\Gamma(y)}{|x-y|}dx\,dy = D(\rho_\Gamma,\rho_\Gamma)$$

which show that the terms in ρ_Γ are well-defined as soon as n and ρ_Γ belong to $\mathcal{C}$.

Now, if $\Gamma \in \mathcal{A}_\Lambda$ fulfills the additional property $-P^0 \leq \Gamma \leq 1-P^0$, then it is easily shown that $\Gamma_{++} := (1-P^0)\Gamma(1-P^0) \geq 0$ and $\Gamma_{--} := P^0\Gamma P^0 \leq 0$, and thus $\mathrm{str}_{P^0}(D^0\Gamma) \geq 0$, by $(I.9)$. On the other hand, following an argument of [BBHS99], we derive from $0 \leq \Gamma + P^0 \leq 1$ that $\Gamma^2 + P^0\Gamma + \Gamma P^0 \leq \Gamma$ or equivalently $\Gamma^2 \leq \Gamma_{++} - \Gamma_{--}$. Therefore

$$\iint \frac{|\Gamma(x,y)|^2}{|x-y|}dx\,dy \leq \frac{\pi}{2}\,\mathrm{tr}(|D^0|(\Gamma_{++}-\Gamma_{--})) = \frac{\pi}{2}\,\mathrm{str}_{P^0}(D^0\Gamma).$$

If we now simply remark that

$$-\alpha\int\rho_\Gamma\varphi + \frac{\alpha}{2}\iint\frac{\rho_\Gamma(x)\rho_\Gamma(y)}{|x-y|}dx\,dy = \frac{\alpha}{2}\Big(D(\rho_\Gamma,\rho_\Gamma) - 2D(\rho_\Gamma,Zn)\Big) \geq -\frac{\alpha Z^2}{2}D(n,n),$$

we finally obtain

$$\mathcal{E}(\Gamma) \geq \Big(1-\frac{\alpha\pi}{4}\Big)\,\mathrm{str}_{P^0}(D^0\Gamma) - \frac{\alpha Z^2}{2}D(n,n) \geq -\frac{\alpha Z^2}{2}D(n,n)$$

when $\alpha \leq 4/\pi$, which easily ends the proof. □

Even if it is more convenient to study the functional $\mathcal{E}$ on the convex hull $\mathcal{B}_\Lambda$, we are interested in minimizers which belong to the BDF class. For the study of the vacuum, this leads to the following definition

Definition I.2. *We say that a projector P is a* BDF-stable vacuum *if and only if $P - P^0$ is a minimizer of $\mathcal{E}$ on $\mathcal{B}_\Lambda$.*

Theorem I.2 implies that P^0 is a BDF-stable vacuum when there is no external potential [CI89, BBHS99], which corresponds to Dirac's ideas. But if we consider a non-vanishing external potential $\varphi = Zn * \frac{1}{|\cdot|}$, then P^0 obviously cannot be BDF-stable, since it is easy to create a state $-P^0 \leq \gamma \leq 1 - P^0$ such that $\mathcal{E}(\gamma) < 0 = \mathcal{E}(0)$. This means that the vacuum is necessarily polarized.

To state the following result, we introduce the notation

$$\mathcal{P}_\Lambda = \left\{ P \text{ projector}, \ P - P^0 \in \mathcal{A}_\Lambda \right\}.$$

Theorem I.3 (BDF-Stability). *Let be $P \in \mathcal{P}_\Lambda$, $n \in \mathcal{C}$ and $Z \in \mathbb{R}$. We assume that there exists a positive constant d such that*

$$d|D_Q| \geq |D^0| \quad \textit{with} \quad \alpha d \frac{\pi}{4} \leq 1, \tag{I.10}$$

where D_Q is defined in (I.3). Then, the following assertions are equivalent

1. *P fulfills the equation*

$$P = \chi_{(-\infty;0)}(D_Q) = \chi_{(-\infty;0)} \left(D^{\alpha\varphi} + \alpha \rho_Q * \frac{1}{|\cdot|} - \alpha \frac{Q(x,y)}{|x-y|} \right). \tag{I.11}$$

2. *P is BDF-stable, i.e. $P - P^0$ is a minimizer of $\mathcal{E}$ on $\mathcal{B}_\Lambda$.*

The main ideas of the proof, which is given in Section 3, are taken from [BBHS99], with additional difficulties coming from the fact that we do not consider trace-class operators. Remark that it is easily seen from (I.11) that $P^{\alpha\varphi}$ is not a BDF-stable vacuum.

2.4 Existence of a BDF-stable vacuum

We may now state our main Theorem.

Theorem I.4 (Existence of a BDF-stable vacuum). *Let be $2 - \sqrt{2} \leq b < 1$, $Z \in \mathbb{R}$ and $n \in \mathcal{C}$. Then for all Λ, there exists an $\alpha_b(\Lambda) > 0$ such that for all α which satisfies*

$$4\sqrt{\pi}\alpha Z \|n\|_{\mathcal{C}} \leq b \quad \textit{and} \quad \alpha \leq \alpha_b(\Lambda)$$

where

$$\alpha_b(\Lambda) \sim_{\Lambda\to\infty} \frac{C(1-b)}{\sqrt{\log \Lambda}},$$

there exists a unique stable vacuum P in some ball $B(P^0, R_0(\alpha,\Lambda)) \subset \mathcal{P}_\Lambda$ containing $P^{\alpha\varphi}$, which is a solution of

$$P = \chi_{(-\infty;0)} \left(D^0 + \alpha(\rho_Q - Zn) * \frac{1}{|\cdot|} + \alpha \frac{Q(x,y)}{|x-y|} \right)$$

with $Q = P - P^0$. Moreover, we have

$$\mathrm{str}_{P^0}(P - P^0) = 0.$$

Remarks: It can be easily shown that when $\alpha \to 0$ and $\alpha Z = \kappa$ is fixed, our solution P_α converges to the projector in the Furry picture P^φ where $\varphi = \kappa n * \frac{1}{|\cdot|}$. This justifies the use of the Furry picture for high-Z atoms in numerical calculations.

A precise definition of the constant C appearing in this result is given in the proof, but we do not know an explicit formula. However, we have obtained numerically the bound $C \geq 0.798$.

$\mathcal{P}_\Lambda$ is naturally equipped with the metric

$$d(P_1, P_2) = \|P_1 - P_2\|_{\mathfrak{S}_2(\mathcal{H}_\Lambda)}$$

which is weaker than the one we use in the proof (and which defines the ball in Theorem I.4).

3 Proof of Theorem I.3

In this section, we prove Theorem I.3. To this end, we first need to prove Theorem I.1.

3.1 Proof of Theorem I.1

Let be P and P' two projectors such that $P' - P \in \mathfrak{S}_2(\mathcal{H})$, and a Hilbert-Schmidt operator A which is supertrace-class with respect to P. This means that PAP and $(1-P)A(1-P)$ are trace-class.

Let us first show that $P'AP'$ is trace-class. To this end, we write

$$\begin{aligned} P'AP' &= (P' - P + P)A(P' - P + P) \\ &= (P' - P)A(P' - P) + (P' - P)AP + PA(P' - P) + PAP. \end{aligned} \tag{I.12}$$

This shows that $P'AP'$ is trace-class since the last term is in $\mathfrak{S}_1$ by assumption, $P' - P$ and A are in $\mathfrak{S}_2$, and P is bounded. The same computation shows that $(1 - P')A(1 - P')$ is trace-class.

We now compute

$$\begin{aligned} \mathrm{tr}[P'AP'] &= \mathrm{tr}[(P' - P)A(P' - P)] + \mathrm{tr}[(P' - P)AP] + \mathrm{tr}[PA(P' - P)] + \mathrm{tr}[PAP] \\ &= \mathrm{tr}[A\left((P' - P)(P' - P) + P(P' - P) + (P' - P)P\right)] + \mathrm{tr}[PAP] \\ &= \mathrm{tr}[A(P' - P)] + \mathrm{tr}[PAP], \end{aligned}$$

where we have used the formula $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, valid for $A, B \in \mathfrak{S}_2$. The same computation gives

$$\mathrm{tr}[P'_+AP'_+] = \mathrm{tr}[A(P'_+ - P_+)] + \mathrm{tr}[P_+AP_+] = -\mathrm{tr}[A(P' - P)] + \mathrm{tr}[P_+AP_+],$$

where we have used the notation $P_+ = 1 - P$ and $P'_+ = 1 - P'$. Summing this two results, we obtain the formula $\mathrm{str}_P[A] = \mathrm{str}_{P'}[A]$. □

3.2 Preliminaries

To prove Theorem I.3, we need the following

Lemma I.2. *Assume that $\varphi = \rho * \frac{1}{|\cdot|}$ for some $\rho \in \mathcal{C}$. Then*

$$\|\nabla\varphi\|_{H^1} = 4\pi\|\rho\|_{\mathcal{C}}, \qquad \|\varphi\|_{L^\infty} \leq C_\infty 4\pi\|\rho\|_{\mathcal{C}}, \qquad \|\varphi\|_{L^6} \leq C_6 4\pi\|\rho\|_{\mathcal{C}}$$

where $C_\infty := \frac{1}{2\pi^{1/2}}$ and C_6 is the Sobolev constant for the inequality $\|u\|_{L^6(\mathbb{R}^3)} \leq C_6\|\nabla u\|_{L^2(\mathbb{R}^3)}$.

Proof – We have

$$\int_{\mathbb{R}^3} \frac{1+|k|^2}{|k|^2} |\widehat{\rho}(k)|^2 \, dk = \frac{1}{(4\pi)^2} \int_{\mathbb{R}^3} |k|^2 (1+|k|^2) |\widehat{\varphi}(k)|^2 \, dk,$$

and so

$$\|\varphi\|_{L^\infty} \leq \frac{1}{(2\pi)^{3/2}} \|\hat{\varphi}\|_{L^1} \leq \frac{1}{(2\pi)^{3/2}} \left(\int_{\mathbb{R}^3} \frac{dk}{|k|^2(1+|k|^2)} \right)^{1/2} 4\pi \|\rho\|_{\mathcal{C}}.$$

The rest is easily obtained by the Sobolev inequalities. □

Lemma I.3. *Let P be a projector such that $P - P^0 \in \mathcal{A}_\Lambda$. Then D_Q is bounded.*

Proof – Due to the cut-off in Fourier space, D^0 is bounded on $\mathcal{H}_\Lambda$. On the other hand, if $\varphi = \rho * \frac{1}{|\cdot|}$ for some $\rho \in \mathcal{C}$, then $\varphi \in L^\infty$ by Lemma I.2 and so this is also a bounded operator. Let us now denote $R(x,y) = \frac{Q(x,y)}{|x-y|}$. We then have

$$|Rf(x)|^2 = \left| \int_{\mathbb{R}^3} \frac{Q(x,y) f(y)}{|x-y|} \, dy \right|^2 \leq \left(\int_{\mathbb{R}^3} \frac{|Q(x,y)|^2}{|x-y|} \, dy \right) \times \left(\int_{\mathbb{R}^3} \frac{|f(y)|^2}{|x-y|} \, dy \right)$$

and since, by a Hardy-type inequality,

$$\int_{\mathbb{R}^3} \frac{|f(y)|^2}{|x-y|} \, dy \leq \frac{\pi}{2} \langle f, |D^0| f \rangle$$

$$\iint_{\mathbb{R}^6} \frac{|Q(x,y)|^2}{|x-y|} \, dx \, dy \leq \frac{\pi}{2} \operatorname{tr}(|D^0| \, Q^2) \tag{I.13}$$

this shows that $R \leq C \, |D^0|^{1/2}$ and so R is bounded. □

Lemma I.4. *Let be P a projector such that $P - P^0 \in \mathcal{A}_\Lambda$. Then $D_Q \Gamma \in \mathfrak{S}_1^{P_0}(\mathcal{H}_\Lambda)$ for all $\Gamma \in \mathcal{A}_\Lambda$ and we have*

$$\operatorname{str}_{P^0}(D^0 \Gamma) + \alpha \int_{\mathbb{R}^3} \rho_\Gamma \left((\rho_Q - Zn) * \frac{1}{|\cdot|} \right) + \alpha \operatorname{tr} \left(\frac{Q(x,y)}{|x-y|} \Gamma \right) = \operatorname{str}_{P^0}(D_Q \Gamma).$$

Proof – Remark that $R\Gamma = \frac{Q(x,y)}{|x-y|}\Gamma$ is trace-class, since $\frac{Q(x,y)}{|x-y|^{1/2}}$ and $\frac{\Gamma(x,y)}{|x-y|^{1/2}}$ are in $\mathfrak{S}_2$ by (I.13). Let us now define $D = D^{-\alpha(\rho_Q - Zn) * \frac{1}{|\cdot|}} = D^0 + \alpha(\rho_Q - Zn) * \frac{1}{|\cdot|}$ and $P' = \chi_{(-\infty;0)}(D)$. By the result of Klaus-Scharf [KS77] (see also [HS98] and the proof of Theorem I.4), it is known that $P' - P^0 \in \mathfrak{S}_2(\mathcal{H}_\Lambda)$. Thus $P' D \Gamma P' = D P' \Gamma P' \in \mathfrak{S}_1(\mathcal{H}_\Lambda)$ since $\Gamma \in \mathfrak{S}_1^{P_0} = \mathfrak{S}_1^{P'}$ by Theorem I.1, and D is bounded by the proof of Lemma I.3. Therefore, $D_Q \Gamma = D\Gamma + R\Gamma$ is in $\mathfrak{S}_1^{P_0}$.

To show the expected equality, we show

$$\operatorname{str}_{P^0}(D^0 \Gamma) + \alpha \int_{\mathbb{R}^3} \rho_\Gamma \varphi'_Q = \operatorname{str}_{P^0}(D\Gamma) \tag{I.14}$$

where $\varphi'_Q = \rho'_Q * \frac{1}{|\cdot|}$, $\rho'_Q = \rho_Q - Zn \in \mathcal{C}$. This will end the proof since the other term is trace-class. The general idea of the proof is to approximate Γ by a trace-class operator for which this equality is true, and to pass to the limit. However, the behaviour of the associated density in the space $\mathcal{C}$ is not obvious and to overcome this difficulty, we shall also approximate the density ρ'_Q to obtain a potential in $L^2(\mathbb{R}^3)$. We thus start by choosing

a sequence ρ_j which converges as $j \to +\infty$ to ρ'_Q in $\mathcal{C}$, such that $\varphi_j = \rho_j * \frac{1}{|\cdot|}$ is in $L^2(\mathbb{R}^3)$. We can choose for instance $\widehat{\rho_j}(k) = \widehat{\rho'_Q}(k)\chi(|k| \geq 1/j)$. We now show

$$\mathrm{str}_{P^0}(D^0\Gamma) + \alpha \int_{\mathbb{R}^3} \rho_\Gamma \varphi_j = \mathrm{str}_{P^0}(D_j\Gamma) \tag{I.15}$$

for all $\Gamma \in \mathcal{A}_\Lambda$, and where $D_j = D^0 - \alpha\varphi_j$. To this end, we may find a sequence Γ^n_{+-} of finite rank operator which converges to $\Gamma_{+-} = (1-P^0)\Gamma P^0$ in $\mathfrak{S}_2$. Then

$$\Gamma^n := \begin{pmatrix} \Gamma_{++} & \Gamma^n_{+-} \\ (\Gamma^n_{+-})^* & \Gamma_{--} \end{pmatrix}$$

converges to Γ in $\mathfrak{S}_2$. Since $\Gamma^n \in \mathfrak{S}_1$ for all $n \geq 0$, we have

$$\mathrm{str}_{P^0}(D^0\Gamma^n) + \alpha \int_{\mathbb{R}^3} \rho_{\Gamma^n} \varphi_j = \mathrm{str}_{P^0}(D_j\Gamma^n). \tag{I.16}$$

By (I.6), the function $Q \in \mathfrak{S}_2(\mathcal{H}_\Lambda) \mapsto \rho_Q \in L^2(\mathbb{R}^3)$ is continuous. Therefore, $\rho_{\Gamma^n} \to \rho_\Gamma$ in $L^2(\mathbb{R}^3)$. Since $\varphi_j \in L^2(\mathbb{R}^3)$, we may now pass to the limit in (I.16) and obtain

$$\lim_{n\to\infty} \left(\mathrm{str}_{P^0}(D^0\Gamma^n) + \alpha \int_{\mathbb{R}^3} \rho_{\Gamma^n}\varphi_j \right) = \mathrm{str}_{P^0}(D^0\Gamma) + \alpha \int_{\mathbb{R}^3} \rho_\Gamma \varphi_j,$$

where we have used that

$$\mathrm{str}_{P^0}(D^0\Gamma^n) = \mathrm{tr}(D^0\Gamma^n_{++}) + \mathrm{tr}(D^0\Gamma^n_{--}) = \mathrm{tr}(D^0\Gamma_{++}) + \mathrm{tr}(D^0\Gamma_{--}) = \mathrm{str}_{P^0}(D^0\Gamma).$$

Let us now pass to the limit in the right hand side. Indeed, we can write, by Theorem I.1,

$$\mathrm{str}_{P^0}(D_j\Gamma^n) = \mathrm{str}_{P'_j}(D_j\Gamma^n) = \mathrm{tr}(D_jP'_j\Gamma^nP'_j) + \mathrm{tr}(D_j(1-P'_j)\Gamma^n(1-P'_j))$$

where $P'_j = \chi_{(-\infty;0)}(D_j)$ and since $P'_j - P^0 \in \mathfrak{S}_2$ by [KS77]. Now, using (I.12), it is easily seen that $P'_j\Gamma^nP'_j \to P'_j\Gamma P'_j$ and $(1-P'_j)\Gamma^n(1-P'_j) \to (1-P'_j)\Gamma(1-P'_j)$ in $\mathfrak{S}_1$ as $n \to \infty$, since this terms can be expanded as a sum of trace-class operators and products of at least two Hilbert-Schmidt operators converging strongly in $\mathfrak{S}_2$. Since D_j is bounded by the proof of Lemma I.3, we obtain that $\mathrm{str}_{P'_j}(D_j\Gamma^n) \to_{n\to\infty} \mathrm{str}_{P'_j}(D_j\Gamma) = \mathrm{str}_{P^0}(D_j\Gamma)$ by Theorem I.1.

As a conclusion, we have proved (I.15) for all $\Gamma \in \mathcal{A}_\Lambda$. To finish the proof, it remains to pass to the limit as $j \to +\infty$. Since

$$\int_{\mathbb{R}^3} \rho_\Gamma \varphi_j = D(\rho_\Gamma, \rho_j)$$

and $\rho_\Gamma \in \mathcal{C}$ (recall that $\Gamma \in \mathcal{A}_\Lambda$), $\rho_j \to \rho'_Q$ strongly in $\mathcal{C}$ as $j \to \infty$, we may pass to the limit in the left hand side of (I.15). To pass to the limit in the right hand side, we use again the fact that

$$\mathrm{str}_{P^0}(D_j\Gamma) = \mathrm{str}_{P'_j}(D_j\Gamma) = \mathrm{tr}(D_jP'_j\Gamma P'_j) + \mathrm{tr}(D_j(1-P'_j)\Gamma(1-P'_j)).$$

By the results of Klaus-Scharf [KS77] (see also the proof of Theorem I.4), it is known that $P'_j - P' \to 0$ in $\mathfrak{S}_2$, since $\rho_j \to \rho'_Q$ in $\mathcal{C}$. Using again (I.12), it is then easily seen that $P'_j\Gamma P'_j \to P'\Gamma P'$ and $(1-P'_j)\Gamma(1-P'_j) \to (1-P')\Gamma(1-P')$ in $\mathfrak{S}_1$ as $j \to \infty$. Since $D_j \to D$ in $\mathfrak{S}_\infty$ by Lemma I.3, we may thus pass to the limit and obtain the desired equality (I.14). □

3.3 Proof of Theorem I.3

We are now able to prove Theorem I.3.

We start by proving 1) ⇒ 2). We thus consider a projector P that satisfies the assumption of the Theorem, and is also a solution to the equation $P = \chi_{(-\infty;0)}(D_Q)$. We fix some $\Gamma \in \mathcal{B}_\Lambda$ and show that $\mathcal{E}(\Gamma) \geq \mathcal{E}(Q)$. To this end, we write $\mathcal{E}(\Gamma) = \mathcal{E}(Q + \Gamma')$ where $\Gamma' = \Gamma - Q = \Gamma + P^0 - P$. By assumption, Γ fulfills $-P^0 \leq \Gamma \leq 1 - P^0$, and so Γ' fulfills $-P \leq \Gamma' \leq 1 - P$. Using Lemma I.4, we may expand $\mathcal{E}(Q + \Gamma')$ and obtain

$$\mathcal{E}(Q+\Gamma') = \mathrm{str}_{P^0}\left(D_Q\Gamma'\right) + \frac{\alpha}{2}\iint \frac{\rho_{\Gamma'}(x)\rho_{\Gamma'}(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\Gamma'(x,y)|^2}{|x-y|}dx\,dy + \mathcal{E}(Q).$$

It is thus sufficient to prove that

$$\mathrm{str}_{P^0}\left(D_Q\Gamma'\right) + \frac{\alpha}{2}\iint \frac{\rho_{\Gamma'}(x)\rho_{\Gamma'}(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\Gamma'(x,y)|^2}{|x-y|}dx\,dy \geq 0.$$

Indeed we now essentially follow the proof of Theorem I.2: we have

$$\iint \frac{\rho_{\Gamma'}(x)\rho_{\Gamma'}(y)}{|x-y|}dx\,dy \geq 0$$

and

$$\iint \frac{|\Gamma'(x,y)|^2}{|x-y|}dx\,dy \leq \frac{\pi}{2}\,\mathrm{tr}(|D^0|(\Gamma')^2) \leq \frac{\pi d}{2}\,\mathrm{tr}(|D_Q|(\Gamma')^2)$$

since by assumption $|D^0| \leq d|D_Q|$. From $-P \leq \Gamma' \leq 1 - P$, we obtain as before $(\Gamma')^2 \leq P\Gamma'P - (1-P)\Gamma'(1-P)$. We thus have

$$\iint \frac{|\Gamma'(x,y)|^2}{|x-y|}dx\,dy \leq \frac{\pi d}{2}\,\mathrm{tr}\left[|D_Q|(P\Gamma'P - (1-P)\Gamma'(1-P))\right] = \frac{\pi d}{2}\,\mathrm{str}_P(D_Q\Gamma')$$

since P commutes with D_Q. Using Theorem I.1 and the fact that $P - P^0 \in \mathfrak{S}_2$, we deduce that $\mathrm{str}_{P^0}\left(D_Q\Gamma'\right) = \mathrm{str}_P\left(D_Q\Gamma'\right) \geq 0$ and

$$\mathrm{str}_{P^0}\left(D_Q\Gamma'\right) + \frac{\alpha}{2}\iint \frac{\rho_{\Gamma'}(x)\rho_{\Gamma'}(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\Gamma'(x,y)|^2}{|x-y|}dx\,dy \geq \left(1 - \alpha d\frac{\pi}{4}\right)\mathrm{str}_P(D_Q\Gamma') \geq 0$$

when $\alpha d\frac{\pi}{4} \leq 1$, which ends the proof of this first part.

We now show 2) ⇒ 1). Let be P which satisfies the assumption of the Theorem, and such that $Q = P - P^0$ is a minimizer of $\mathcal{E}$ in $\mathcal{B}_\Lambda$. We therefore have for all $\gamma \in \mathfrak{S}_1 \cap \mathcal{A}_\Lambda$ such that $-P \leq \gamma \leq 1 - P$

$$\mathcal{E}(Q + \gamma) \geq \mathcal{E}(Q)$$

or equivalently

$$\mathcal{E}'(\gamma) := \mathrm{tr}\left(D_Q\gamma\right) + \frac{\alpha}{2}\iint \frac{\rho_\gamma(x)\rho_\gamma(y)}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|\gamma(x,y)|^2}{|x-y|}dx\,dy \geq 0. \tag{I.17}$$

We now follow the proof of Theorem 4 by Bach *et al.* [BBHS99]. Their proof is done with D_Q replaced by $D^{\alpha\varphi}$ but they also mention that it can be extended to a more general case, provided $0 \notin \sigma(D_Q)$ and P, $1 - P$ leave the domain of D_Q invariant, which is the case here. Let us however explain the general ideas of this proof for the reader's convenience.

If P does not commute with $\chi_{(-\infty;0)}(D_Q)$, then there exist two normalized vectors $f \in (1-P)\mathcal{H}_\Lambda$ and $g \in P\mathcal{H}_\Lambda$ such that the real part of $\langle f, D_Q g\rangle$ is not 0. We now let

$$\gamma_\varepsilon = \varepsilon^2 \left(|f\rangle\langle f| - |g\rangle\langle g|\right) + \varepsilon\sqrt{1-\varepsilon^2}\left(|f\rangle\langle g| + |g\rangle\langle f|\right)$$

which is a trace-class operator satisfying $-P \leq \gamma \leq 1-P$. Computing now the energy (I.17) of γ_ε yields

$$\mathcal{E}'(\gamma_\varepsilon) = 2\varepsilon\, \Re\langle f, D_Q g\rangle + O(\varepsilon^2)$$

which contradicts (I.17) since this can be made negative for ε small enough. This implies that P commutes with D_Q, and so it commutes with all spectral projections of D_Q, and namely with $P' = \chi_{(-\infty;0)}(D_Q)$ and $1-P'$.

We now show that $P = P' = \chi_{(-\infty;0)}(D_Q)$. Suppose that there exists a $u \in (1-P)\mathcal{H}_\Lambda \cap P'\mathcal{H}_\Lambda$. We then let $\gamma_0 = |u\rangle\langle u|$. The corresponding energy is then

$$\mathcal{E}'(\gamma_0) = \langle u, D_Q u\rangle < 0,$$

since the exchange and non-exchange terms cancel. This contradicts (I.17). Since we may show with the same argument that $P\mathcal{H}_\Lambda \cap (P'\mathcal{H}_\Lambda)^c = \emptyset$, we obtain that $P = P' = \chi_{(-\infty;0)}(D_Q)$. □

4 Proof of Theorem I.4

In this section, we prove Theorem I.4 by using a Banach fixed-point method.

4.1 Preliminaries

We start by defining the norms and spaces that will be used to apply this well-known result. In fact, the main challenge concerning the fixed-point argument consisted in finding suitable Banach spaces.

Norms and spaces

We choose the following norms

$$\|Q\|_{\mathcal{Q}} := \left(\iint E(p-q)^2 E(p+q)|\widehat{Q}(p,q)|^2\, dp\, dq\right)^{1/2},$$

$$\|R\|_{\mathcal{R}} := \left(\iint \frac{E(p-q)^2}{E(p+q)}|\widehat{R}(p,q)|^2\, dp\, dq\right)^{1/2},$$

$$\|\rho\|_{\mathcal{C}} := \left(\int \frac{E(k)^2}{|k|^2}|\widehat{\rho}(k)|^2\, dk\right)^{1/2},$$

$$\|\varphi\|_{\mathcal{Y}} := \left(\int |k|^2 E(k)^2 |\widehat{\varphi}(k)|^2\, dk\right)^{1/2},$$

where

$$E(k) = \sqrt{1+|k|^2}$$

and denote by $\mathcal{Q}$, $\mathcal{R}$, $\mathcal{C}$ and $\mathcal{Y}$ the associated Hilbert spaces. The dual space $\mathcal{C}'$ of $\mathcal{C}$ will be also useful and we introduce

$$\|\zeta\|_{\mathcal{C}'} := \int \frac{|k|^2}{E(k)^2} |\hat{\zeta}(k)|^2 \, dk.$$

In the following, it will be easier to use the norm $\|R\|_{\mathcal{R}}$ where $R = \frac{Q(x,y)}{|x-y|}$ in our estimates and a relation with $\|Q\|_{\mathcal{Q}}$ will then be needed. To this end, we first need the following Lemma, which will be useful throughout the rest of the proof.

Lemma I.5. *For all ξ and η in $\mathbb{R}^3$, we have*

$$\forall s \geq 0, \;\; E(\xi)^s \leq 2^{\delta(s)} \left(E(\xi - \eta)^s + E(\eta)^s \right) \tag{I.18}$$

$$\forall s \in \mathbb{R}, \;\; E(\xi)^s \leq 2^{|s|} E(\xi - \eta)^s E(\eta)^{|s|}, \tag{I.19}$$

with

$$\delta(s) = \begin{cases} s & \text{if} \quad 0 \leq s < 1 \\ s - 1 & \text{if} \quad s \geq 1 \end{cases}.$$

Proof – We have $E(\xi) = \|(1, \xi)\| \leq \|(1, \xi - \eta)\| + \|(0, \eta)\| \leq \|(1, \xi - \eta)\| + \|(1, \eta)\|$, proving $(I.18)$ when $s = 1$. Now, for $s \geq 1$, the result follows from the convexity of $x \mapsto x^s$ on $[1; +\infty)$, using $(a/2 + b/2)^s \leq a^s/2 + b^s/2$. When $s \in [0; 1)$, we simply write $(a+b)^s = (a+b)^{s+1}/(a+b) \leq 2^s a^s a/(a+b) + 2^s b^s b/(a+b) \leq 2^s (a^s + b^s)$, ending the proof of $(I.18)$. On the other hand, the Peetre inequality $(I.19)$, is easily obtained by reduction to the case $s = 1$. □

Remark: A trivial consequence of $(I.18)$ is the following inequality

$$\frac{1}{E(p) + E(q)} \leq \min \left(\frac{1}{E(p+q)}, \frac{1}{E(p-q)} \right). \tag{I.20}$$

Now we can give a connection between $\|R\|_{\mathcal{R}}$ and $\|Q\|_{\mathcal{Q}}$ when $R = \frac{Q(x,y)}{|x-y|}$ (we also recall the easy relation between $\|\rho\|_{\mathcal{C}}$ and $\|\varphi\|_{\mathcal{Y}}$ when $\varphi = \rho * \frac{1}{|\cdot|}$).

Lemma I.6. *If $\rho \in \mathcal{C}$ and $Q \in \mathcal{Q}$, then we have $\varphi = \rho * \frac{1}{|\cdot|} \in \mathcal{Y}$ and $R(x,y) = \frac{Q(x,y)}{|x-y|} \in \mathcal{R}$ and more precisely*

$$\begin{aligned} \|\varphi\|_{\mathcal{Y}} &= 4\pi \|\rho\|_{\mathcal{C}}, \\ \|R\|_{\mathcal{R}} &\leq C_R \|Q\|_{\mathcal{Q}}, \\ \|R \cdot |D_0|^{-1/2}\|_{\mathfrak{S}_2} &\leq 2 \|R\|_{\mathcal{R}}, \end{aligned} \tag{I.21}$$

with

$$C_R := \frac{1}{2\pi^2} \inf_{\theta \in (0;2)} \sup_{x \in \mathbb{R}^3} \left(E(2x)^\theta \int_{\mathbb{R}^3} \frac{du}{E(2u)^{1+\theta} |u - x|^2} \right). \tag{I.22}$$

Proof of Lemma I.6 – We have

$$\widehat{R}(p, q) = \frac{1}{2\pi^2} \int_{\mathbb{R}^3} \frac{\widehat{Q}(p - l, q - l)}{|l|^2} \, dl$$

so we obtain, for some fixed $\theta \in (0;2)$

$$\begin{aligned}
\|R\|^2_{\mathcal{R}} &= \iint \frac{E(p-q)^2}{E(p+q)} |\widehat{R}(p,q)|^2 \, dp\, dq \\
&= 8 \iint \frac{E(2v)^2}{E(2u)} |\widehat{R}(u+v, u-v)|^2 \, du\, dv \\
&= \frac{8}{(2\pi^2)^2} \iiiint \frac{E(2v)^2}{E(2u)} \frac{\widehat{Q}(l+v, l-v) \cdot \widehat{Q}(l'+v, l'-v)}{|l-u|^2 |l'-u|^2} \, du\, dv\, dl\, dl' \\
&= \frac{8}{(2\pi^2)^2} \iint du\, dv \frac{E(2v)^2}{E(2u)} \iint \frac{E(2l)^{\frac{1+\theta}{2}} \widehat{Q}(l+v, l-v)\, E(2l')^{\frac{1+\theta}{2}} \widehat{Q}(l'+v, l'-v)}{E(2l')^{\frac{1+\theta}{2}} |l-u|\, |l'-u|\, E(2l)^{\frac{1+\theta}{2}} |l-u|\, |l'-u|} dl\, dl' \\
&\leq \frac{8}{(2\pi^2)^2} \iiiint \frac{E(2v)^2}{E(2u)} \frac{E(2l)^{1+\theta}}{E(2l')^{1+\theta}} \frac{|\widehat{Q}(l+v, l-v)|^2}{|l-u|^2 |l'-u|^2} \, du\, dv\, dl\, dl' \\
&\leq 8 \iint E(2v)^2 E(2l) |\widehat{Q}(l+v, l-v)|^2 K_\theta(l) \, dv\, dl
\end{aligned}$$

where

$$K_\theta(l) := \frac{E(2l)^\theta}{(2\pi^2)^2} \iint \frac{1}{E(2u) \cdot E(2l')^{1+\theta} |l-u|^2 |l'-u|^2} \, du\, dl'.$$

Now, let us introduce

$$C_\theta := \sup_{x \in \mathbb{R}^3} \left(E(2x)^\theta \int_{\mathbb{R}^3} \frac{du}{E(2u)^{1+\theta} |u-x|^2} \right).$$

Remark that

$$\int_{\mathbb{R}^3} \frac{du}{E(2u)^{1+\theta} |u-x|^2} \leq \frac{1}{2^{1+\theta} |x|^\theta} \int_{\mathbb{R}^3} \frac{du}{|u|^{1+\theta} |u-e_x|^2}$$

where $e_x := x/|x|$, showing that $C_\theta < \infty$ when $\theta \in (0;2)$. Now we have

$$\begin{aligned}
K_\theta(l) &= \frac{E(2l)^\theta}{(2\pi^2)^2} \int du \frac{1}{E(2u)^{1+\theta} |l-u|^2} \left(E(2u)^\theta \int \frac{1}{E(2l')^{1+\theta} |l'-u|^2} dl' \right) \\
&\leq \frac{E(2l)^\theta}{(2\pi^2)^2} \int du \frac{1}{E(2u)^{1+\theta} |l-u|^2} \times C_\theta \leq \left(\frac{C_\theta}{2\pi^2} \right)^2
\end{aligned}$$

and so

$$\|R\|^2_{\mathcal{R}} \leq 8 \left(\frac{C_\theta}{2\pi^2} \right)^2 \iint E(2v)^2 E(2l) |\widehat{Q}(l+v, l-v)|^2 \, dv\, dl \leq \left(\frac{C_\theta}{2\pi^2} \right)^2 \|Q\|^2_{\mathcal{Q}}$$

which ends the proof of $(I.21)$.

Finally, we remark that we have

$$\begin{aligned}
\|R \cdot |D_0|^{-1/2}\|^2_{\mathfrak{S}_2} = \iint \frac{|\widehat{R}(p,q)|^2}{E(p)} \, dp\, dq &\leq 2 \iint \frac{|\widehat{R}(p,q)|^2}{E(2p)} \, dp\, dq \\
&\leq 4 \iint \frac{E(p-q) |\widehat{R}(p,q)|^2}{E(p+q)} \, dp\, dq \leq 4 \|R\|^2_{\mathcal{R}}
\end{aligned}$$

by Lemma I.5 □

An estimate from below for D_Q

We now state a Lemma in which we give a lower estimate for the operator

$$D_{Q,\rho} := D^{\alpha\varphi} + \alpha\rho * \frac{1}{|r|} - \alpha\frac{Q(x,y)}{|x-y|}$$

by D^0, in terms of the spaces introduced above.

For our result, we are interested in $D_Q = D_{Q,\rho_Q}$ but this definition with a different ρ will be useful in the following.

Lemma I.7. *Assume that $(Q,\rho) \in \mathcal{Q}\times\mathcal{C}$ are such that*

$$\alpha\left(2\sqrt{\pi}\|\rho - Zn\|_{\mathcal{C}} + 2C_R\|Q\|_{\mathcal{Q}}\right) < 1. \tag{I.23}$$

Then $D_{Q,\rho}$ is a bounded operator which satisfies

$$|D_{Q,\rho}| \geq \left(1 - \alpha\left(2\sqrt{\pi}\|\rho - Zn\|_{\mathcal{C}} + 2C_R\|Q\|_{\mathcal{Q}}\right)\right)|D^0|. \tag{I.24}$$

Proof – We have, with $\varphi' = (\rho - Zn) * \frac{1}{|\cdot|}$,

$$\|\varphi' u\|_{L^2} \leq \|\varphi'\|_{L^\infty}\|u\|_{L^2} \leq 2\pi^{1/2}\|\rho - Zn\|_{\mathcal{C}}\||D_0|\cdot u\|_{L^2}$$

by Lemma I.2, and

$$\|Ru\|_{L^2} = \|R|D_0|^{-1/2}|D_0|^{1/2}u\|_{L^2} \leq \|R|D_0|^{-1/2}\|_{\mathfrak{S}_2}\||D_0|^{1/2}u\|_{L^2} \leq 2C_R\|Q\|_{\mathcal{Q}}\||D_0|\cdot u\|_{L^2}$$

by Lemma I.6. This shows that $|\varphi' - R| \leq (2\pi^{1/2}\|\rho_Q - Zn\|_{\mathcal{C}} + 2C_R\|Q\|_{\mathcal{Q}})|D^0|$, the square root being monotone. This proves that $D_{Q,\rho}$ is bounded since D_0 is bounded on $\mathcal{H}_\Lambda$, and gives the expected inequality. □

Remark that Lemma I.7 will be useful when we shall apply Theorem I.3 (see the condition (I.10) in the statement). It also implies $0 \notin \sigma(D_{Q,\rho})$, a fact that will be used to compute the projection $\chi_{(-\infty;0)}(D_{Q,\rho})$.

Expansion by Cauchy's formula

We want to solve the equation

$$Q = \chi_{(-\infty;0)}\left(D^0 + \alpha(\rho_Q - Zn) * \frac{1}{|r|} - \alpha\frac{Q(x,y)}{|x-y|}\right) - \chi_{(-\infty;0)}(D^0) := F_1(Q).$$

If $\alpha\left(2\sqrt{\pi}\|\rho_Q - Zn\|_{\mathcal{C}} + 2C_R\|Q\|_{\mathcal{Q}}\right) < 1$, then $0 \notin \sigma(D_Q)$ by Lemma I.7. We may thus use the method of [HS03] and expand F_1 by Cauchy's formula

$$F_1(Q) = -\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\eta\left(\frac{1}{D_Q + i\eta} - \frac{1}{D^0 + i\eta}\right) = \sum_{n=1}^{\infty}\alpha^n Q_n$$

where

$$Q_n = -\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\eta\frac{1}{D^0 + i\eta}\left((R_Q - \varphi'_Q)\frac{1}{D^0 + i\eta}\right)^n$$

$$\rho'_Q = \rho_Q - Zn, \quad \varphi'_Q = \rho'_Q * \frac{1}{|r|}, \quad R_Q(x,y) = \frac{Q(x,y)}{|x-y|}.$$

We shall write

$$Q_n = \sum_{k,l/\ k+l=n} Q_{k,l}$$

$$Q_{k,l} = \frac{(-1)^{l+1}}{2\pi} \sum_{I\cup J=\{1,\dots,n\},\ |I|=k,\ |J|=l} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0+i\eta} \prod_{j=1}^{n} \left(R_j \frac{1}{D^0+i\eta} \right)$$

where $R_j = R_Q$ if $j \in I$ and $R_j = \varphi'_Q$ if $j \in J$ ($Q_{k,l}$ is the sum of all the terms containing k R_Q's and l φ'_Q's). We also denote $\rho_{k,l} := \rho_{Q_{k,l}}$.

Hence our equation can be written

$$\begin{cases} Q = \sum_{n=1}^{\infty} \alpha^n Q_n(Q, \rho_Q) \\ \rho_Q = \sum_{n=1}^{\infty} \alpha^n \rho_n(Q, \rho_Q), \end{cases} \tag{I.25}$$

where we recall that Q_n and ρ_n depend on both Q and ρ_Q. In order to have a better condition on α and Λ, we shall now change the second equation for the density, by taking into account the special form of the first order term ρ_1. To this end, we need to compute this term explicitely.

The first order density

Recall that

$$Q_{0,1} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0+i\eta} \varphi'_Q \frac{1}{D^0+i\eta}$$

so that

$$\widehat{Q_{0,1}}(p,q) = (2\pi)^{-5/2} \int_{-\infty}^{+\infty} d\eta \frac{1}{\boldsymbol{\alpha}\cdot p+\beta+i\eta} \widehat{\varphi'_Q}(p-q) \frac{1}{\boldsymbol{\alpha}\cdot q+\beta+i\eta}.$$

We now introduce

$$\begin{aligned} M(p,q) &:= \frac{1}{\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{\boldsymbol{\alpha}\cdot p+\beta+i\eta} \cdot \frac{1}{\boldsymbol{\alpha}\cdot q+\beta+i\eta} \\ &= \frac{1}{\pi} \int_{-\infty}^{+\infty} d\eta \frac{\boldsymbol{\alpha}\cdot p+\beta-i\eta}{p^2+1+\eta^2} \cdot \frac{\boldsymbol{\alpha}\cdot q+\beta-i\eta}{q^2+1+\eta^2} \\ &= \frac{1}{E(p)+E(q)} \left(\frac{(\boldsymbol{\alpha}\cdot p+\beta)}{E(p)} \frac{(\boldsymbol{\alpha}\cdot q+\beta)}{E(q)} - 1 \right). \end{aligned} \tag{I.26}$$

Hence

$$\widehat{Q_{0,1}}(p,q) = \frac{1}{2^{5/2}\pi^{3/2}} \widehat{\varphi'_Q}(p-q) M(p,q).$$

This enables us to compute

$$\begin{aligned} \widehat{\rho_{0,1}}(k) &= \frac{1}{(2\pi)^{3/2}} \int_{|l|\le\Lambda} \mathrm{Tr}_{\mathbb{C}^4} \left(\widehat{Q_{0,1}}(l+k/2, l-k/2) \right) dl \\ &= \frac{1}{16\pi^3} \widehat{\varphi'_Q}(k) \int_{|l|\le\Lambda} \mathrm{Tr}_{\mathbb{C}^4} \left(M(l+k/2, l-k/2) \right) dl \\ &= -\frac{1}{4\pi^2} \widehat{\varphi'_Q}(k) |k|^2 B'_\Lambda(k) = -\frac{1}{\pi} \widehat{\rho'_Q}(k) B'_\Lambda(k) \\ &= \frac{1}{\pi} (Z\widehat{n}(k) - \widehat{\rho_Q}(k)) B'_\Lambda(k) \end{aligned} \tag{I.27}$$

where

$$B'_\Lambda(k) = -\frac{1}{\pi|k|^2}\int_{|l|\leq\Lambda}\frac{(l+k/2)\cdot(l-k/2)+1-E(l+k/2)E(l-k/2)}{E(l+k/2)E(l-k/2)(E(l+k/2)+E(l-k/2))}dl.$$

This function is computed in [PR36]

$$B'_\Lambda(k) = \int_0^{\frac{\Lambda}{E(\Lambda)}}\frac{z^2-z^4/3}{1-z^2}\frac{dz}{1+|k|^2(1-z^2)/4}$$

and so for instance

$$B'_\Lambda(k) \leq B'_\Lambda(0) = \int_0^{\frac{\Lambda}{E(\Lambda)}}\frac{z^2-z^4/3}{1-z^2}\,dz = \frac{2}{3}\log(\Lambda) - \frac{5}{9} + \frac{2}{3}\log 2 + O(1/\Lambda^2).$$

In the following, we shall use the notation $B_\Lambda := \frac{B'_\Lambda(0)}{\pi} \sim_{\Lambda\to\infty} \frac{2}{3\pi}\log(\Lambda)$.

Equation

We are now able to introduce the function on which we shall apply the fixed-point theorem. According to what we said above, the equation in ρ_Q can be written

$$\rho_Q = \sum_{n=1}^{\infty}\alpha^n\rho_n(Q,\rho_Q) \tag{I.28}$$

or equivalently (we forget the dependence in Q and ρ_Q for simplicity)

$$\widehat{\rho_Q}(k) = \frac{\alpha B'_\Lambda(k)}{\pi}(Z\widehat{n}(k) - \widehat{\rho_Q}(k)) + \alpha\widehat{\rho_{1,0}}(k) + \sum_{n=2}^{\infty}\alpha^n\widehat{\rho_n}(k)$$

and

$$\widehat{\rho_Q}(k) = \frac{\alpha B'_\Lambda(k)}{\pi+\alpha B'_\Lambda(k)}Z\widehat{n}(k) + \frac{\pi}{\pi+\alpha B'_\Lambda(k)}\left(\alpha\widehat{\rho_{1,0}}(k) + \sum_{n=2}^{\infty}\alpha^n\widehat{\rho_n}(k)\right), \tag{I.29}$$

which is the equation that we want to solve for the density.

Since the equations in Q and ρ_Q are now different, we have to introduce the following space:

$$\mathcal{X} = \mathcal{Q}\times\mathcal{C}$$

consisting of all the pairs (Q,ρ) such that $Q\in\mathcal{Q}$ and $\rho\in\mathcal{C}$. Notice that in this space, ρ can be different from ρ_Q. However, we shall find a solution of the equations in this space, which satisfies $\rho = \rho_Q$. We also introduce on $\mathcal{X}$ the norm

$$\|(Q,\rho)\| = C_R\sqrt{2}\|Q\|_{\mathcal{Q}} + 2\sqrt{\pi}\|\rho\|_{\mathcal{C}},$$

where we recall that C_R is defined in Lemma I.6.

We now introduce the function $F:\mathcal{X}\to\mathcal{X}$ defined by

$$F(Q,\rho) = \Big(F_Q(Q,\rho)\ ,\ F_\rho(Q,\rho)\Big)$$

where

$$F_Q(Q,\rho) = \chi_{(-\infty;0)}(D_{Q,\rho}) - P^0 = \sum_{n=1}^{\infty} \alpha^n Q_n(Q,\rho) \tag{I.30}$$

$$\widehat{F_\rho(Q,\rho)} = \frac{\alpha B'_\Lambda}{\pi + \alpha B'_\Lambda} Z\widehat{n} + \frac{\pi}{\pi + \alpha B'_\Lambda}\left(\alpha\widehat{\rho_{1,0}}(Q,\rho) + \sum_{n=2}^{\infty} \alpha^n \widehat{\rho_n}(Q,\rho)\right) \tag{I.31}$$

where $Q_n(Q,\rho)$ and $\rho_n(Q,\rho)$ are defined in (I.25) (replace ρ_Q by ρ). Remark that $\rho_n = \rho_{Q_n}$ for all $n \geq 2$. In the proof of Theorem I.4, we solve the fixed-point equation in $\mathcal{X}$

$$F(Q,\rho) = (Q,\rho).$$

4.2 Proof of Theorem I.4

To prove our main Theorem, we need the following estimates

Proposition I.8. *Assume that $(Q,\rho) \in \mathcal{X}$ is such that $0 \notin \sigma(D_{Q,\rho})$. Then we have*

$$\|F(Q,\rho)\| \leq \kappa_1(\Lambda)\alpha\|(Q,\rho)\| + \left(\frac{\alpha B_\Lambda}{1+\alpha B_\Lambda} + \alpha\kappa_1(\Lambda)\right) 2\sqrt{\pi} Z\|n\|_{\mathcal{C}} + \sum_{n=2}^{+\infty} \kappa_n \left[\alpha(\|(Q,\rho)\| + 2\sqrt{\pi}Z\|n\|_{\mathcal{C}})\right]^n \tag{I.32}$$

$$\|F'(Q,\rho)\| \leq \kappa_1(\Lambda)\alpha + \alpha \sum_{n=2}^{+\infty} n\kappa_n \left[\alpha(\|(Q,\rho)\| + 2\sqrt{\pi}Z\|n\|_{\mathcal{C}})\right]^{n-1} \tag{I.33}$$

where

$$\kappa_1(\Lambda) = \max\left(\frac{C_R\sqrt{2}}{\sqrt{\pi}}\sqrt{\log\Lambda}\,,\ \sqrt{2}C_R + \frac{\sqrt{\log\Lambda}}{2^{3/2}\sqrt{\pi}}\right) \sim_{\Lambda\to\infty} \frac{C_R\sqrt{2}}{\sqrt{\pi}}\sqrt{\log\Lambda}$$

and $(\kappa_n)_{n\geq 2}$ is a sequence of positive numbers independent of Λ and which satisfies $\kappa_n \sim_{n\to\infty} K\sqrt{n}$ for some constant K.

To prove this proposition, we have to do some tedious estimates that we postpone until the end of the proof of Theorem I.4.

We are now able to prove Theorem I.4

Let us introduce the function $f(x) = \sum_{n=2}^{\infty} \kappa_n x^n$, which is a power serie with a radius of convergence equal to 1. The estimates $(I.32)$ and $(I.33)$ can be written

$$\|F(Q,\rho)\| \leq \kappa_1(\Lambda)\alpha\|(Q,\rho)\| + \left(\frac{\alpha B_\Lambda}{1+\alpha B_\Lambda} + \alpha\kappa_1(\Lambda)\right) 2\sqrt{\pi}Z\|n\|_{\mathcal{C}} + f\big(\alpha\|(Q,\rho)\| + 2\sqrt{\pi}\alpha Z\|n\|_{\mathcal{C}}\big) \tag{I.34}$$

$$\|F'(Q,\rho)\| \leq \kappa_1(\Lambda)\alpha + \alpha f'\big(\alpha\|(Q,\rho)\| + 2\sqrt{\pi}\alpha Z\|n\|_{\mathcal{C}}\big).$$

To apply Banach fixed-point theorem, we now have to find a ball $B(0,R) \subset \mathcal{X}$ which is invariant under the function F and on which F is a contraction. Let $R > 0$ be some fixed radius. We have

$$\sup_{(Q,\rho)\in B(0,R)} \|F'(Q,\rho)\| \leq \kappa_1(\Lambda)\alpha + \alpha f'\big(\alpha R + 2\sqrt{\pi}\alpha Z\|n\|_{\mathcal{C}}\big) := \mu.$$

Moreover, we also have

$$\|F(Q,\rho)\| \leq \|F(Q,\rho) - F(0,0)\| + \|F(0,0)\| \leq \mu\|(Q,\rho)\| + \|F(0,0)\|.$$

Therefore a condition for the ball $B(0,R)$ to be invariant under the action of F is $\|F(0,0)\| \leq (1-\mu)R$. Additionally due to Lemma I.7 we assume $\alpha\sqrt{2}R < 1$ as well as

$$\frac{\alpha\pi}{4(1-\alpha\sqrt{2}R)} \leq 1,$$

due to Theorem I.3 and Lemma I.7 which is equivalent to

$$\alpha \leq \frac{1}{\pi/4 + R\sqrt{2}}.$$

As a conclusion, if (α, R) fulfills

$$\begin{cases} \left(\dfrac{\alpha B_\Lambda}{1+\alpha B_\Lambda} + \kappa_1(\Lambda)\right) 2\sqrt{\pi}Z\|n\|_{\mathcal{C}} + f(2\sqrt{\pi}Z\|n\|_{\mathcal{C}}) + \kappa_1(\Lambda)\alpha R + \alpha R f'(\alpha R + 2\sqrt{\pi}Z\|n\|_{\mathcal{C}}) \leq R \\ \alpha \leq \dfrac{1}{\pi/4 + R\sqrt{2}}, \end{cases} \tag{I.35}$$

then we are able to apply the Banach Fixed-Point Theorem on $B(0,R)$. Remark that these inequalities also contain the conditions $\mu < 1$ and $\alpha\sqrt{2}R < 1$. Notice also that if (α, R) is a solution to (I.35), then (α', R) is a solution to (I.35) for all $\alpha' \leq \alpha$, since the function which appears on the left of (I.35) is increasing in α.

Now, if we assume that $4\sqrt{\pi}\alpha Z\|n\|_{\mathcal{C}} \leq b$, we obtain that if (α, R) fulfills

$$\begin{cases} \dfrac{b}{2\alpha} + \kappa_1(\Lambda)b/2 + f(b/2) + \kappa_1(\Lambda)\alpha R + \alpha R f'(\alpha R + b/2) \leq R \\ \alpha \leq \dfrac{1}{\pi/4 + R\sqrt{2}}, \end{cases} \tag{I.36}$$

then it also fulfills (I.35), f being increasing. Remark that we have used the estimate

$$\frac{\alpha B_\Lambda}{1+\alpha B_\Lambda} 2\sqrt{\pi}Z\|n\|_{\mathcal{C}} \leq \frac{b}{2\alpha}.$$

The first inequation of (I.36) is simpler when it is written in terms of the variables α and $x := \alpha R$. It becomes

$$\frac{b}{2\alpha} + \kappa_1(\Lambda)b/2 + f(b/2) + \kappa_1(\Lambda)x + xf'(x + b/2) \leq \frac{x}{\alpha} \tag{I.37}$$

which implies $b/2 \leq x$. Since necessarily $x < 1 - b/2$ (recall that f is defined on $[0;1)$), we see that we need $b < 1$. Now, let us call $a_\Lambda(x)$ the solution of (I.37) where the $\leq$ is replaced by $=$, i.e.

$$a_\Lambda(x) = \frac{x - b/2}{\kappa_1(\Lambda)b/2 + f(b/2) + \kappa_1(\Lambda)x + xf'(x + b/2)}, \tag{I.38}$$

a smooth function on $[b/2; 1-b/2)$. Since $a_\Lambda(b/2) = \lim_{x\to 1-b/2} a_\Lambda(x) = 0$, we deduce that a_Λ attains its maximum on $(b/2; 1-b/2)$ and we may denote by x_Λ the largest $x \in (b/2; 1-b/2)$ such that

$$a_\Lambda(x_\Lambda) = A_b(\Lambda) := \max_{x\in(b/2;1-b/2)} a_\Lambda(x). \tag{I.39}$$

We now define $R_\Lambda := x_\Lambda/A_\Lambda$ and

$$\alpha_b(\Lambda) := \min\left(A_b(\Lambda), \frac{1}{\pi/4 + R_\Lambda\sqrt{2}}\right).$$

As a conclusion, for all $0 \le \alpha \le \alpha_b(\Lambda)$, there exists some $R_0 = R_0(\alpha, \Lambda, b)$ such that (α, R_0) is a solution of (I.35). This means that F is a contraction on $B(0, R_0)$, on which we can apply Banach Theorem. This gives a unique solution to the equation $F(Q, \rho) = (Q, \rho)$ in $B(0, R_0) \subset \mathcal{X}$.

Let us now show that P is indeed a solution to (I.4). In fact ρ is a solution to (I.29) and so it is a solution to (I.28). On the other hand, we have $Q = \chi_{(-\infty;0)}(D_{Q,\rho}) - P^0$, and so equation (I.28) means exactly that $\rho = \rho_Q$. Hence, P is a solution to $P = \chi_{(-\infty;0)}(D_Q)$. Thanks to the proof, we know that P satisfies the assumptions of Theorem I.3, and so P is a BDF-stable vacuum.

To end the proof, let us study the behaviour of $A_b(\Lambda)$ as $\Lambda \to \infty$. Differentiating (I.38) with respect to x, we arrive at

$$\kappa_1(\Lambda)b + f(b) + \frac{bf'(x_\Lambda + b/2)}{2} = (x_\Lambda - b/2)x_\Lambda f''(x_\Lambda + b/2). \tag{I.40}$$

Since $\kappa_1(\Lambda) \to \infty$ as $\Lambda \to \infty$, we deduce that $x_\Lambda + b/2 \to 1$, or equivalently $x_\Lambda \to 1 - b/2$. If we now use the fact that $f'(y) = o\big(f''(y)\big)$ as $y \to 1^-$, we finally obtain $f''(x_\Lambda + b/2) \sim_{\Lambda\to\infty} \frac{b\kappa_1(\Lambda)}{(1-b)(1-b/2)}$ and $f'(x_\Lambda + b/2) = o\big(\kappa_1(\Lambda)\big)$. By (I.38), this implies

$$A_\Lambda \sim_{\Lambda\to\infty} \frac{1-b}{\kappa_1(\Lambda)}.$$

Remark also that $R_\Lambda = x_\Lambda/A_\Lambda \sim \frac{1-b/2}{1-b}\kappa_1(\Lambda)$, which ends the proof since when $2 - \sqrt{2} \le b$,

$$\alpha_b(\Lambda) \sim_{\Lambda\to\infty} \frac{\min\left(1-b, \frac{1-b}{\sqrt{2}(1-b/2)}\right)}{\kappa_1(\Lambda)} = \frac{1-b}{\kappa_1(\Lambda)} \sim_{\Lambda\to\infty} \frac{\sqrt{\pi}(1-b)}{C_R\sqrt{2}\sqrt{\log\Lambda}} = \frac{C(1-b)}{\sqrt{\log\Lambda}}$$

with $C = \frac{\sqrt{\pi}}{C_R\sqrt{2}}$. □

4.3 Proof of Proposition I.8: estimates

In this section, we prove the claimed estimates of Proposition I.8.

Remark first that we have

$$\frac{\alpha B'_\Lambda(k)}{\pi + \alpha B'_\Lambda(k)} \le \frac{\alpha B'_\Lambda(0)}{\pi + \alpha B'_\Lambda(0)} = \frac{\alpha B_\Lambda}{1 + \alpha B_\Lambda} \quad \text{and} \quad \frac{\pi}{\pi + \alpha B'_\Lambda(k)} \le 1$$

for all $k \in \mathbb{R}^3$. Therefore, to estimate the norm $\|F_\rho(Q, \rho)\|_{\mathcal{C}}$, it suffices to estimate the norms of $\rho_{1,0}(Q, \rho)$ and $\rho_n(Q, \rho)$, due to (I.31).

For $(Q, \rho) \in \mathcal{X}$, we introduce the notation $R = \frac{Q(x,y)}{|x-y|} \in \mathcal{R}$, $\rho' = \rho - Zn$ and $\varphi' = (\rho - Zn) * \frac{1}{|\cdot|}$.

We then remark that

$$\|F(Q, \rho)\| \le \frac{\alpha B_\Lambda}{1 + \alpha B_\Lambda} 2\sqrt{\pi}Z\|n\|_{\mathcal{C}} + \alpha\|(Q_1, \rho_{1,0})\| + \sum_{n\ge 2} \alpha^n \|(Q_n, \rho_n)\|$$

and estimate each term separately. A similar argument can be done for $F'(Q, \rho)$.

4.3.1 First order terms

Lemma I.9. *We have the following estimates:*

$$\|Q_{0,1}\|_{\mathcal{Q}} \leq \frac{(\log\Lambda)^{1/2}}{2\pi}\|\varphi'\|_{\mathcal{Y}} = 2(\log\Lambda)^{1/2}\|\rho'\|_{\mathcal{C}},$$

$$\|Q_{1,0}\|_{\mathcal{Q}} \leq \|R_Q\|_{\mathcal{R}} \leq C_R\|Q\|_{\mathcal{Q}}, \qquad \|\rho_{1,0}\|_{\mathcal{C}} \leq \frac{C_R(\log\Lambda)^{1/2}}{4\pi}\|Q\|_{\mathcal{Q}}.$$

Therefore

$$\|(Q_1,\rho_{1,0})\| \leq \kappa_1(\Lambda)\|(Q,\rho')\| \leq \kappa_1(\Lambda)\|(Q,\rho)\| + \kappa_1(\Lambda)2\sqrt{\pi}Z\|n\|_{\mathcal{C}},$$

where

$$\kappa_1(\Lambda) = \max\left(\frac{C_R\sqrt{2}}{\sqrt{\pi}}\sqrt{\log\Lambda}\,,\ \sqrt{2}C_R + \frac{\sqrt{\log\Lambda}}{2^{3/2}\sqrt{\pi}}\right).$$

Proof – Recall that

$$\widehat{Q_{0,1}}(p,q) = \frac{1}{2^{5/2}\pi^{3/2}}\widehat{\varphi'}(p-q)M(p,q), \tag{I.41}$$

where the matrix $M(p,q)$ is defined in (I.26), and whose properties are summarized in the following

Lemma I.10. *Let* $\Lambda^+(p) = \frac{\boldsymbol{\alpha}\cdot p+\beta+E(p)}{2E(p)}$ *and* $\Lambda^-(p) = \frac{-(\boldsymbol{\alpha}\cdot p+\beta)+E(p)}{2E(p)}$ *be the projections matrices in* $\mathbb{C}^4$ *onto the eigenspaces of* D^0 *in Fourier space. We then have*

$$\mathrm{Tr}_{\mathbb{C}^4}\left(M(p,q)\right) = -4\frac{1}{E(p)+E(q)}\mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^+(p)\Lambda^-(q)\right)$$

$$|M(p,q)|^2 = \mathrm{Tr}_{\mathbb{C}^4}\left(M(p,q)M(p,q)^*\right) = 8\frac{1}{(E(p)+E(q))^2}\mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^+(p)\Lambda^-(q)\right)$$

$$\mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^+(p)\Lambda^-(q)\right) = \mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^-(p)\Lambda^+(q)\right) = 1 - \frac{p\cdot q+1}{E(p)E(q)}.$$

Moreover, we have

$$\forall p,q\in\mathbb{R}^3, \quad \mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^+(p)\Lambda^-(q)\right) \leq \min\left(\frac{|p-q|^2}{2E((p+q)/2)^2}\,,\ \frac{1}{2}\right). \tag{I.42}$$

Proof of Lemma I.10 – We only prove (*I*.42). We have $\mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^+(p)\Lambda^-(q)\right) \leq |\Lambda^+(p)|\,|\Lambda^-(q)| = 1/2$, so when $t := \frac{|p-q|^2}{4E((p+q)/2)^2} \geq 1/2$, there is nothing to prove. Now we have

$$\begin{aligned} 1 - \frac{p\cdot q+1}{E(p)E(q)} &= 1 - \frac{l^2-k^2+1}{E(l+k)E(l-k)} \text{ with } l = \frac{p+q}{2},\ k = \frac{p-q}{2} \\ &= 1 - \frac{1-t}{\sqrt{(1+t)^2-4zt}} \end{aligned}$$

where $t = |k|^2/E(l)^2$ and $z = \frac{(l\cdot k)^2}{|k|^2(1+l^2)} \in [0;1)$. When $t\in[0;1/2)$ and $z\in[0;1)$, the expression above is decreasing in z and so we obtain

$$1 - \frac{p\cdot q+1}{E(p)E(q)} \leq 1 - \frac{1-t}{\sqrt{(1+t)^2}} = \frac{2t}{1+t} \leq 2t$$

which ends the proof. □

• Let us now treat $Q_{0,1}$. From (I.41), we obtain

$$|\widehat{Q_{0,1}}(p,q)|^2 = \frac{1}{2^5\pi^3}|\widehat{\varphi'}|^2(p-q)|M(p,q)|^2,$$

and so

$$\begin{aligned}\iint E(p-q)^2E(p+q)|\widehat{Q_{0,1}}(p,q)|^2dp\,dq \\ = \frac{1}{2^5\pi^3}\iint dk\,du E(k)^2E(2u)|\widehat{\varphi'}|^2(k)\chi(|u|\leq\Lambda)|M(u+k/2,u-k/2)|^2 \\ \leq \frac{4}{2^5\pi^3}\left(\int dk|k|^2E(k)^2|\widehat{\varphi'}|^2(k)\right)\left(\int_{|u|\leq\Lambda}\frac{1}{E(2u)E(u)^2}\right)\end{aligned}$$

by Lemma I.10. Now we have

$$\int_{|u|\leq\Lambda}\frac{du}{E(2u)E(u)^2} = 4\pi\left(\frac{1}{2}\mathrm{argsh}(2\Lambda)+\frac{1}{\sqrt{3}}\mathrm{argth}\left(\frac{\sqrt{3}\Lambda}{\sqrt{1+4\sqrt{\Lambda}}}\right)\right)\leq 2\pi\log\Lambda \quad \text{(I.43)}$$

for $\Lambda\geq 3$. So we obtain

$$\|Q_{0,1}\|_{\mathcal{Q}}\leq\frac{(\log\Lambda)^{1/2}}{2\pi}\|\varphi'\|_{\mathcal{Y}} = 2(\log\Lambda)^{1/2}\|\rho'\|_{\mathcal{C}}.$$

• $\rho_{1,0}$ and $Q_{1,0}$.
We have

$$Q_{1,0} = -\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\eta\frac{1}{D^0+i\eta}R\frac{1}{D^0+i\eta}$$

so that

$$\begin{aligned}\widehat{Q_{1,0}}(p,q) &= -(2\pi)^{-1}\int_{-\infty}^{+\infty}d\eta\frac{1}{\boldsymbol{\alpha}\cdot p+\beta+i\eta}\widehat{R}(p,q)\frac{1}{\boldsymbol{\alpha}\cdot q+\beta+i\eta} \\ &= -\frac{1}{2}\frac{1}{E(p)+E(q)}\left(\frac{(\boldsymbol{\alpha}\cdot p+\beta)}{E(p)}\widehat{R}(p,q)\frac{(\boldsymbol{\alpha}\cdot q+\beta)}{E(q)}-\widehat{R}(p,q)\right)\end{aligned}$$

and

$$|\widehat{Q_{1,0}}(p,q)|^2\leq\frac{1}{(E(p)+E(q))^2}|\widehat{R}(p,q)|^2\leq\frac{1}{E(p+q)^2}|\widehat{R}(p,q)|^2,$$

showing that

$$\|Q_{1,0}\|_{\mathcal{Q}}\leq\|R\|_{\mathcal{R}}\leq C_R\|Q\|_{\mathcal{Q}}.$$

Now, we have

$$\begin{aligned}\widehat{\rho_{1,0}}(k) &= \frac{1}{(2\pi)^{3/2}}\int_{\mathbb{R}^3}\mathrm{Tr}_{\mathbb{C}^4}\left(\widehat{Q_{1,0}}\left(l+\frac{k}{2},l-\frac{k}{2}\right)\right)dl \\ &= -\frac{1}{2^{5/2}\pi^{3/2}}\int_{\mathbb{R}^3}\mathrm{Tr}_{\mathbb{C}^4}\left(\widehat{R}\left(l+\frac{k}{2},l-\frac{k}{2}\right)M\left(l+\frac{k}{2},l-\frac{k}{2}\right)\right)\chi(|l|\leq\Lambda)dl,\end{aligned}$$

so we obtain

$$|\widehat{\rho_{1,0}}(k)| \leq \frac{1}{2^{5/2}\pi^{3/2}} \left(\int_{\mathbb{R}^3} E(2l)^{-1} |\widehat{R}(l+k/2, l-k/2)|^2 \, dl \right)^{1/2} \times \\ \times \left(\int_{\mathbb{R}^3} E(2l) |M(l+k/2, l-k/2)|^2 \, dl \right)^{1/2}$$

and finally

$$\int \frac{E(k)^2}{|k|^2} |\widehat{\rho_{1,0}}(k)|^2 \, dk \leq \frac{1}{2^5\pi^3} \|R\|_{\mathcal{R}}^2 \int_{|l|\leq\Lambda} \frac{1}{E(2l)E(l)^2} \, dl \leq \frac{\log\Lambda}{2^4\pi^2} \|R\|_{\mathcal{R}}^2$$

by $(I.43)$ which implies

$$\|\rho_{1,0}\|_{\mathcal{C}} \leq \frac{C_R(\log\Lambda)^{1/2}}{4\pi} \|Q\|_{\mathcal{Q}}. \quad \square$$

4.3.2 Second order terms

To simplify the presentation, we introduce the following notation:

$$S_{p,q} := (4\pi)(2\pi)^{-3/p} \left(\int_{\mathbb{R}^3} \frac{du}{E(u)^q} \right)^{\frac{1}{p}}, \qquad S_p := S_{p,p}, \qquad K_p := \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)^p}.$$

Let us recall the following inequality [Sim79, Theorem 4.1]

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_p} \leq (2\pi)^{-3/p} \|f\|_{L^p(\mathbb{R}^3)} \|g\|_{L^p(\mathbb{R}^3)}, \tag{I.44}$$

which implies

$$\|\frac{1}{|D_0|^a} f\|_{\mathfrak{S}_p} \leq \frac{S_{p,ap}}{4\pi} \|f\|_{L^p(\mathbb{R}^3)}.$$

On the other hand, we shall often use the following trick

$$\begin{aligned} (E(p)^2+\eta^2)(E(q)^2+\eta^2) &= (E(p)E(q))^2 + (E(p)^2+E(q)^2)\eta^2 + \eta^4 \\ &\geq \frac{1}{4}E(p+q)^2 + \frac{1}{2}E(p+q)^2\eta^2 \\ &\geq \frac{1}{4}E(p+q)^2 E(\eta)^2 \end{aligned}$$

by Lemma I.5. This implies

$$\frac{1}{\sqrt{E(p)^2+\eta^2}\sqrt{E(q)^2+\eta^2}} \leq \frac{2}{E(p+q)E(\eta)}. \tag{I.45}$$

Recall now that we have $Q_2 = Q_{2,0} + Q_{1,1} + Q_{0,2}$ with

$$\begin{aligned} Q_{2,0} &= -\frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0+i\eta} R \frac{1}{D^0+i\eta} R \frac{1}{D^0+i\eta} \\ Q_{1,1} &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0+i\eta} R \frac{1}{D^0+i\eta} \varphi' \frac{1}{D^0+i\eta} \\ &\quad + \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0+i\eta} \varphi' \frac{1}{D^0+i\eta} R \frac{1}{D^0+i\eta} \\ Q_{0,2} &= -\frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0+i\eta} \varphi' \frac{1}{D^0+i\eta} \varphi' \frac{1}{D^0+i\eta}. \end{aligned}$$

We shall now treat each term separately.

Lemma I.11. *We have* $\rho_{0,2} = 0$ *and the following estimates:*

$$\|Q_{2,0}\|_{\mathcal{Q}} \leq 2^{5/2} K_{3/2} (C_R)^2 \|Q\|_{\mathcal{Q}}^2, \qquad \|Q_{1,1}\|_{\mathcal{Q}} \leq 4 S_6 C_6 K_{3/2} C_R \|Q\|_{\mathcal{Q}} \|\rho'\|_{\mathcal{C}},$$

$$\|Q_{0,2}\|_{\mathcal{Q}} \leq 2\sqrt{10} S_{6,5} C_M C_6 \|\rho'\|_{\mathcal{C}}^2,$$

$$\|\rho_{2,0}\|_{\mathcal{C}} \leq \frac{S_6 C_6 (C_R)^2}{\pi} \|Q\|_{\mathcal{Q}}^2, \qquad \|\rho_{1,1}\|_{\mathcal{C}} \leq 4 \frac{S_{6,5} C_M C_6 C_R}{\pi} \|\rho'\|_{\mathcal{C}} \|Q\|_{\mathcal{Q}},$$

and so

$$\|(Q_2, \rho_2)\| \leq \kappa_2 \|(Q, \rho')\|^2,$$

with

$$\kappa_2 = C_{Q_2} C_R \sqrt{2} + 2\sqrt{\pi} C_{\rho_2},$$

$$C_{Q_2} = \max\left(2^{3/2} K_{3/2}\,,\, \frac{S_6 C_6 K_{3/2}}{\sqrt{2\pi}}\,,\, \frac{\sqrt{5} S_{6,5} C_M C_6}{\pi\sqrt{2}}\right),\ C_{\rho_2} = \max\left(\frac{S_6 C_6}{2\pi}\,,\, \frac{S_{6,5} C_M C_6}{\pi^{3/2}\sqrt{2}}\right).$$

Proof – **Step 1 : Estimates on the exchange term** Q_2**.**

- $Q_{2,0}$. To estimate $Q_{2,0}$, we write

$$|\widehat{Q_{2,0}}(p,q)| \leq \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \int_{\mathbb{R}^3} dp_1 \frac{|\widehat{R}(p,p_1)|}{\sqrt{E(p)^2+\eta^2}} \frac{|\widehat{R}(p_1,q)|}{\sqrt{E(p_1)^2+\eta^2}} \frac{1}{\sqrt{E(q)^2+\eta^2}}$$

and so by $(I.45)$

$$|\widehat{Q_{2,0}}(p,q)| \leq \frac{2^{3/2}}{2\pi} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)^{3/2} E(p+q)^{1/2}} \int_{\mathbb{R}^3} \frac{|\widehat{R}(p,p_1)|}{E(p+p_1)^{1/2}} \frac{|\widehat{R}(p_1,q)|}{E(p_1+q)^{1/2}}\, dp_1$$

which implies

$$E(p-q)E(p+q)^{1/2} |\widehat{Q_{2,0}}(p,q)| \leq 2^{5/2} K_{3/2} \int_{\mathbb{R}^3} \frac{E(p-p_1)|\widehat{R}(p,p_1)|}{E(p+p_1)^{1/2}} \frac{E(p_1-q)|\widehat{R}(p_1,q)|}{E(p_1+q)^{1/2}}\, dp_1$$

and finally

$$\|Q_{2,0}\|_{\mathcal{Q}} \leq 2^{5/2} K_{3/2} \|R\|_{\mathcal{R}}^2 \leq 2^{5/2} K_{3/2} (C_R)^2 \|Q\|_{\mathcal{Q}}^2.$$

- $Q_{1,1}$. We treat for instance

$$Q'_{1,1} := \frac{1}{(2\pi)^{5/2}} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0 + i\eta} R \frac{1}{D^0 + i\eta} \varphi' \frac{1}{D^0 + i\eta}$$

and use the same method to obtain

$$E(p-q)E(p+q)^{1/2} |\widehat{Q'_{1,1}}(p,q)| \leq \frac{4}{(2\pi)^{5/2}} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)} \times$$
$$\times \int_{\mathbb{R}^3} \frac{E(p-p_1)}{E(p+p_1)^{1/2}} |\widehat{R}(p,p_1)| \frac{E(p_1-q)}{(E(p_1)^2+\eta^2)^{1/4} (E(q)^2+\eta^2)^{1/4}} |\widehat{\varphi'}(p_1-q)|\, dp_1.$$

This means that

$$\|Q'_{1,1}\|_{\mathcal{Q}} \leq \frac{4}{2\pi} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)} \|R' \frac{1}{(|D_0|^2+\eta^2)^{1/4}} f \frac{1}{(|D_0|^2+\eta^2)^{1/4}} \|_{\mathfrak{S}_2}$$

where we have introduced R' and f defined by

$$\widehat{R'}(p,q) := \frac{E(p-q)}{E(p+q)^{1/2}}|\widehat{R}(p,q)|, \quad \widehat{f}(k) := E(k)|\widehat{\varphi'}(k)|.$$

But now

$$\begin{aligned}\|R'\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}f\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}\|_{\mathfrak{S}_2} &\leq \|R'\|_{\mathfrak{S}_2}\|\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}f\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}\|_{\mathfrak{S}_\infty}\\ &\leq \|R\|_{\mathcal{R}}\|\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}f\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}\|_{\mathfrak{S}_6}.\end{aligned}$$

If we now use inequality $(I.44)$, we obtain

$$\begin{aligned}\|\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}f\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}\|_{\mathfrak{S}_6} &\leq \|\frac{1}{(|D_0|^2+\eta^2)^{\frac{1}{4}}}|f|^{1/2}\|^2_{\mathfrak{S}_{12}}\\ &\leq (2\pi)^{-1/2}\left(\int_{\mathbb{R}^3}\frac{du}{1+|u|^2+\eta^2)^3}\right)^{1/6}\|f\|_{L^6}\\ &= \frac{S_6}{4\pi E(\eta)^{1/2}}\|f\|_{L^6}.\end{aligned}$$

Finally since

$$\|f\|_{L^6} \leq C_6\|\nabla f\|_{L^2} = C_6\|\varphi'\|_{\mathcal{Y}}, \tag{I.46}$$

and $\|\varphi'\|_{\mathcal{Y}} = (4\pi)\|\rho'\|_{\mathcal{C}}$, we obtain

$$\|Q'_{1,1}\|_{\mathcal{Q}} \leq \frac{4S_6C_6}{2\pi}\int_{-\infty}^{+\infty}\frac{d\eta}{E(\eta)^{3/2}}\|R\|_{\mathcal{R}}\|\rho'\|_{\mathcal{C}}$$

and

$$\|Q_{1,1}\|_{\mathcal{Q}} \leq 8S_6C_6K_{3/2}C_R\|Q\|_{\mathcal{Q}}\|\rho'\|_{\mathcal{C}}.$$

• $Q_{0,2}$. Unfortunately, the method used above cannot be applied to $Q_{0,2}$. In this case, we have to calculate this term explicitely. We can write

$$Q_{0,2} = \sum_{\varepsilon_1,\varepsilon_2,\varepsilon_3\in\{\pm\}} Q_{0,2}^{\varepsilon_1\varepsilon_2\varepsilon_3}$$

where for instance (by a residuum formula)

$$\widehat{Q_{0,2}^{+++}}(p,q) = \widehat{Q_{0,2}^{---}}(p,q) = 0,$$

$$\widehat{Q_{0,2}^{+--}}(p,q) = (2\pi)^{-3}\int_{\mathbb{R}^3}dp_1\frac{\Lambda^+(p)\widehat{\varphi'}(p-p_1)\Lambda^-(p_1)\widehat{\varphi'}(p_1-q)\Lambda^-(q)}{(E(p)+E(q))(E(p)+E(p_1))},$$

$$\widehat{Q_{0,2}^{+-+}}(p,q) = (2\pi)^{-3}\int_{\mathbb{R}^3}dp_1\frac{\Lambda^+(p)\widehat{\varphi'}(p-p_1)\Lambda^-(p_1)\widehat{\varphi'}(p_1-q)\Lambda^+(q)}{(E(p)+E(p_1))(E(q)+E(p_1))},$$

and similar formulas for the other $Q_{0,2}^{\varepsilon_1\varepsilon_2\varepsilon_3}$. We now treat for instance $\widehat{Q_{0,2}^{+--}}$. Using $(I.20)$, we may obtain

$$E(p-q)E(p+q)^{1/2}|\widehat{Q_{0,2}^{+--}}| \leq 2(2\pi)^{-3}\int_{\mathbb{R}^3} dp_1 \frac{E(p-p_1)}{E(p+p_1)^{2/3}}\times \\ \times \left|\Lambda^+(p)\widehat{\varphi'}(p-p_1)\Lambda^-(p_1)\right| \times \frac{E(p_1-q)|\widehat{\varphi'}(p_1-q)|}{E(p_1)^{1/3}E(q)^{1/2}}dp_1.$$

So, we may write

$$\|Q_{0,2}^{+--}\|_{\mathcal{Q}} \leq 2\|M_f \frac{1}{|D_0|^{1/3}} f \frac{1}{|D_0|^{1/2}}\|_{\mathfrak{S}_2} \leq 2\|M_f\|_{\mathfrak{S}_2}\|\frac{1}{|D_0|^{1/3}} f \frac{1}{|D_0|^{1/2}}\|_{\mathfrak{S}_\infty},$$

where

$$\widehat{M_f}(p,q) := (2\pi)^{-3/2}\frac{\widehat{f}(p-q)\,|\Lambda^+(p)\Lambda^-(q)|}{E(p+q)^{2/3}}. \tag{I.47}$$

Lemma I.12. *When M_f is defined by formula $(I.47)$, then*

$$\|M_f\|_{\mathfrak{S}_2} \leq \frac{C_M}{4\pi}\left(\int_{\mathbb{R}^3}|k|^2|\widehat{f}(k)|^2dk\right)^{1/2}$$

where $C_M := 2\left(\int_0^\infty \frac{t^2\,dt}{E(2t)^{4/3}E(t)^2}\right)^{1/2} \simeq 2.1589$.

Proof – We have $|\Lambda^+(p)\Lambda^-(q)|^2 = \mathrm{Tr}_{\mathbb{C}^4}\left(\Lambda^+(p)\Lambda^-(q)\right)$ and so, by $(I.42)$,

$$\begin{aligned}\iint dp\,dq|\widehat{M_f}(p,q)|^2 &\leq (2\pi)^{-3}\iint dp\,dq\frac{|p-q|^2|\widehat{f}(p-q)|^2}{2E((p+q)/2)^2E(p+q)^{4/3}} \\ &\leq (2\pi)^{-3}\int dk|k|^2|\widehat{f}(k)|^2\int\frac{du}{2E(2u)^{4/3}E(u)^2} \\ &\leq (2\pi)^{-2}\left(\int dk|k|^2|\widehat{f}(k)|^2\right)\int_0^\infty\frac{t^2\,dt}{E(2t)^{4/3}E(t)^2}. \quad \square\end{aligned}$$

Finally, since by $(I.46)$

$$\|\frac{1}{|D_0|^{1/3}} f \frac{1}{|D_0|^{1/2}}\|_{\mathfrak{S}_\infty} \leq \|\frac{1}{|D_0|^{1/3}} f \frac{1}{|D_0|^{1/2}}\|_{\mathfrak{S}_6} \leq \frac{S_{6,5}}{4\pi}\|f\|_{L^6},$$

we obtain

$$\|Q_{0,2}^{+--}\|_{\mathcal{Q}} \leq 2S_{6,5}C_MC_6\|\rho'\|_{\mathcal{C}}^2.$$

This result is immediatly extended to the others terms and since we can prove

$$|Q|^2 = |Q^{++-}+Q^{+--}|^2 + |Q^{-++}+Q^{--+}|^2 + |Q^{+-+}|^2 + |Q^{-+-}|^2,$$

we arrive at

$$\|Q_{0,2}\|_{\mathcal{Q}} \leq 2\sqrt{10}S_{6,5}C_MC_6\|\rho'\|_{\mathcal{C}}^2.$$

Step 2 : Estimates on the density ρ_2. Let us now treat the density ρ_2. The general idea of the proof is to estimate $\langle\rho_2,\zeta\rangle$ in terms of the norm $\|\zeta\|_{\mathcal{C}'}$ by using

$$|\langle\rho,\zeta\rangle| = |\mathrm{Tr}\,(Q\zeta)| = \left|\mathrm{Tr}\left(\widehat{Q\zeta}\right)\right| = \left|\int_{\mathbb{R}^3}\mathrm{Tr}_{\mathbb{C}^4}\left(\widehat{Q\zeta}\right)(p,p)\,dp\right| \leq \int_{\mathbb{R}^3}|\widehat{Q\zeta}(p,p)|dp.$$

This can be done if we know that $Q\zeta \in \mathfrak{S}_1$. But we have

$$\|Q\zeta\|_{\mathfrak{S}_1} = \|Q|D_0|^2 \frac{1}{|D_0|^2}\zeta\|_{\mathfrak{S}_1} \leq \|Q|D_0|^2\|_{\mathfrak{S}_2} \|\frac{1}{|D_0|^2}\zeta\|_{\mathfrak{S}_2} \leq E(\Lambda)^2 \|Q\|_{\mathfrak{S}_2} \frac{S_{2,4}}{4\pi}\|\zeta\|_{L^2},$$

showing that $Q\zeta \in \mathfrak{S}_1$ when $\zeta \in L^2$. So, in what follows, we shall assume that $\zeta \in \mathcal{C}' \cap L^2$ and prove a bound depending only on $\|\zeta\|_{\mathcal{C}'}$. By the density of $\mathcal{C}' \cap L^2$ in $\mathcal{C}'$, this will give us a bound on $\|\rho\|_{\mathcal{C}}$.

Let us remark first that $\rho_{0,2}$ vanishes. Indeed we have

$$\widehat{\rho_{0,2}}(k) = \frac{1}{(2\pi)^{3/2}} \int_{|p|\leq\Lambda} \mathrm{Tr}_{\mathbb{C}^4} \left(\widehat{Q_{0,2}}(p+k/2, p-k/2) \right) dp$$

and

$$\begin{aligned}
\mathrm{Tr}_{\mathbb{C}^4} \left(\widehat{Q_{0,2}}(p,q) \right) &= \frac{1}{(2\pi)^4} \int_{-\infty}^{+\infty} d\eta \int_{\mathbb{R}^3} dp_1 \mathrm{Tr}_{\mathbb{C}^4} \left(\frac{1}{D_0(p)+i\eta} \widehat{\varphi'}(p-p_1) \times \right. \\
&\qquad \left. \times \frac{1}{D_0(p_1)+i\eta} \widehat{\varphi'}(p_1-q) \frac{1}{D_0(q)+i\eta} \right) \\
&= \frac{1}{(2\pi)^4} \int_{-\infty}^{+\infty} d\eta \int_{\mathbb{R}^3} dp_1 \frac{\widehat{\varphi'}(p-p_1)\widehat{\varphi'}(p_1-q)}{\sqrt{E(p)^2+\eta^2}\sqrt{E(p_1)^2+\eta^2}\sqrt{E(q)^2+\eta^2}} \times \\
&\qquad \times \mathrm{Tr}_{\mathbb{C}^4} \Big[(D_0(p)-i\eta))(D_0(p_1)-i\eta)(D_0(q)-i\eta) \Big].
\end{aligned}$$

Now the terms linear in the Dirac matrices are traceless and the remaining terms are odd in η and vanish after integration. This can be easily generalized to $\rho_{0,2k}$ for all k, and is known as *Furry's Theorem* in the physics litterature.

• $\rho_{2,0}$. We use here a method similar to what we have done above. We estimate for some $\zeta \in \mathcal{C}' \cap L^2$ and $Q_\zeta := Q_{2,0}\zeta$

$$\begin{aligned}
|\widehat{Q_\zeta}(p,p)| &\leq (2\pi)^{-5/2} \int_{-\infty}^{+\infty} d\eta \iint \frac{|\widehat{R}(p,p_1)|\,|\widehat{R}(p_1,p_2)|\,|\widehat{\zeta}(p_2-p)|\,dp_1\,dp_2}{\sqrt{E(p)^2+\eta^2}\sqrt{E(p_1)^2+\eta^2}\sqrt{E(p_2)^2+\eta^2}}, \\
&\leq 4(2\pi)^{-5/2} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)} \iint dp_1\,dp_2 \frac{E(p-p_1)|\widehat{R}(p,p_1)|\,E(p_1-p_2)|\widehat{R}(p_1,p_2)|}{E(p+p_1)^{1/2}E(p_1+p_2)^{1/2}} \times \\
&\qquad \times \frac{|\widehat{\zeta}(p_2-p)|}{E(p_2-p)(E(p)^2+\eta^2)^{1/4}(E(p_2)^2+\eta^2)^{1/4}}, \\
&\leq 4(2\pi)^{-5/2} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)} \iint dp_1\,dp_2 \frac{R'(p,p_1)R'(p_1,p_2)\widehat{\zeta'}(p_2-p)}{(E(p)^2+\eta^2)^{1/4}(E(p_2)^2+\eta^2)^{1/4}}
\end{aligned}$$

where $\widehat{R'}(p,q) = \frac{E(p-q)|\widehat{R}(p,q)|}{E(p+q)^{1/2}}$ and $\widehat{\zeta'}(k) = E(k)^{-1}\widehat{\zeta}(k)$. This means that

$$\begin{aligned}
|\langle \rho_{2,0}, \zeta \rangle| &\leq \frac{4}{2\pi} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)} \|R'R' \frac{1}{(|D_0|^2+\eta^2)^{1/4}} \zeta' \frac{1}{(|D_0|^2+\eta^2)^{1/4}}\|_{\mathfrak{S}_1} \\
&\leq \frac{4S_6}{(2\pi)(4\pi)} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)^{3/2}} \|R'\|_{\mathfrak{S}_2}^2 \|\zeta'\|_{L^6} \leq \frac{S_6 C_6 (C_R)^2 K_{3/2}}{\pi} \|Q\|_{\mathcal{Q}}^2 \|\zeta\|_{\mathcal{C}'}
\end{aligned}$$

by $(I.46)$, showing that

$$\|\rho_{2,0}\|_{\mathcal{C}} \leq \frac{S_6 C_6 (C_R)^2}{\pi} K_{3/2} \|Q\|_{\mathcal{Q}}^2.$$

• $\rho_{1,1}$. Unfortunately, as for $Q_{0,2}$, we have to calculate $\rho_{1,1}$ explicitely. Let us start for instance with $\rho_{1,1}^{+--}$, the density associated with one of the two terms of $Q_{1,1}$

$$(2\pi)^{-3/2} \int_{\mathbb{R}^3} dp_1 \frac{\Lambda^+(p)\widehat{R}(p-p_1)\Lambda^-(p_1)\widehat{\varphi'}(p_1-q)\Lambda^-(q)}{(E(p)+E(q))(E(p)+E(p_1))}.$$

We use the same method as above and estimate for some $\zeta \in \mathcal{C}' \cap L^2$ the term

$$\widehat{Q_\zeta}(p,p) = (2\pi)^{-3} \iint dp_1\, dp_2 \frac{\Lambda^+(p)\widehat{R}(p,p_1)\Lambda^-(p_1)\widehat{\varphi'}(p_1-p_2)\Lambda^-(p_2)}{(E(p)+E(p_2))(E(p)+E(p_1))} \widehat{\zeta}(p_2-p),$$

by

$$\begin{aligned}
|\widehat{Q_\zeta}(p,p)| &\leq (2\pi)^{-3} \iint \frac{|\Lambda^+(p)\widehat{R}(p,p_1)\Lambda^-(p_1)|}{E(p+p_1)^{1/2}} \frac{|\widehat{\varphi'}(p_1-p_2)|}{E(p_1)^{1/2}E(p_2)^{1/3}} \times \\
&\qquad \times \frac{|\widehat{\zeta}(p_2-p)| \times |\Lambda^-(p_2)\Lambda^+(p)|}{E(p+p_2)^{2/3}} dp_1\, dp_2, \\
&\leq 2(2\pi)^{-3/2} \iint \frac{E(p-p_1)|\Lambda^+(p)\widehat{R}(p,p_1)\Lambda^-(p_1)|}{E(p+p_1)^{1/2}} \times \\
&\qquad \times \frac{\widehat{f}(p_1-p_2)}{E(p_1)^{1/2}E(p_2)^{1/3}} \widehat{M_{\zeta'}}(p_2,p) dp_1\, dp_2,
\end{aligned}$$

with $\widehat{f}(k) := E(k)|\widehat{\varphi'}(k)|$ and $\widehat{\zeta}'(k) := |\widehat{\zeta}(k)|/E(k)$. Now

$$|\widehat{Q_\zeta}(p,p)| \leq 2(2\pi)^{-3/2} \iint \widehat{R_1}(p,p_1) \frac{\widehat{f}(p_1-p_2)}{E(p_1)^{1/2}E(p_2)^{1/3}} \widehat{M_{\zeta'}}(p_2,p)\, dp_1\, dp_2$$

with

$$\widehat{R_1}(p,p_1) := \frac{E(p-p_1)|\Lambda^+(p)\widehat{R}(p,p_1)\Lambda^-(p_1)|}{E(p+p_1)^{1/2}}.$$

We thus have

$$\begin{aligned}
|\langle \rho_{1,1}^{+--}, \zeta \rangle| &\leq 2\|R_1 \frac{1}{|D_0|^{1/2}} f \frac{1}{|D_0|^{1/3}} M_{\zeta'}\|_{\mathfrak{S}^1} \\
&\leq 2\|\Lambda^+ R \Lambda^-\|_{\mathcal{R}} \|M_{\zeta'}\|_{\mathfrak{S}_2} \|\frac{1}{|D_0|^{1/2}} f \frac{1}{|D_0|^{1/3}}\|_{\mathfrak{S}_\infty} \\
&\leq 2\|\Lambda^+ R \Lambda^-\|_{\mathcal{R}} \|M_{\zeta'}\|_{\mathfrak{S}_2} \|\frac{1}{|D_0|^{1/2}} f \frac{1}{|D_0|^{1/3}}\|_{\mathfrak{S}_6} \\
&\leq \frac{S_{6,5} C_M C_6}{2\pi} \|\Lambda^+ R \Lambda^-\|_{\mathcal{R}} \|\zeta\|_{\mathcal{C}'} \|\rho'\|_{\mathcal{C}},
\end{aligned}$$

and finally

$$\|\rho_{1,1}^{+--}\|_{\mathcal{C}} \leq \frac{S_{6,5} C_M C_6}{2\pi} \|\Lambda^+ R \Lambda^-\|_{\mathcal{R}} \|\rho'\|_{\mathcal{C}}.$$

We now treat $\rho_{1,1}^{+-+}$ and estimate

$$\widehat{Q_\zeta}(p,p) = (2\pi)^{-3} \iint dp_1\, dp_2 \frac{\Lambda^+(p)\widehat{R}(p,p_1)\Lambda^-(p_1)\widehat{\varphi'}(p_1-p_2)\Lambda^+(p_2)}{(E(p)+E(p_1))(E(p_1)+E(p_2))}\widehat{\zeta}(p_2-p),$$

by

$$|\widehat{Q_\zeta}(p,p)| \leq 2(2\pi)^{-3} \iint \frac{E(p-p_1)|\Lambda^+(p)\widehat{R}(p,p_1)\Lambda^-(p_1)|}{E(p+p_1)^{1/2}} \times \\ \times \frac{\widehat{f}(p_1-p_2)|\Lambda^-(p_1)\Lambda^+(p_2)|}{E(p_1+p_2)^{2/3}} \frac{\widehat{\zeta'}(p_2-p)}{E(p_2)^{5/6}} dp_1\, dp_2.$$

Using the same argument as above, we arrive at

$$\|\rho_{1,1}^{+-+}\|_{\mathcal{C}} \leq \frac{S_{6,5}C_M C_6}{2\pi}\|\Lambda^+ R\Lambda^-\|_{\mathcal{R}}\|\rho'\|_{\mathcal{C}}.$$

To treat $\rho_{1,1}^{++-}$, we remark that

$$\frac{1}{(E(p)+E(p_2))(E(p_1)+E(p_2))} \leq \frac{1}{(E(p)+E(p_1))(E(p_1)+E(p_2))} \\ + \frac{1}{(E(p)+E(p_1))(E(p)+E(p_2))}$$

and use the same estimates as above to get

$$\|\rho_{1,1}^{++-}\|_{\mathcal{C}} \leq \frac{S_{6,5}C_M C_6}{\pi}\|\Lambda^+ R\Lambda^-\|_{\mathcal{R}}\|\rho'\|_{\mathcal{C}}.$$

Finally, since $\sum_{\varepsilon_1,\varepsilon_2\in\{\pm\}} \|\Lambda^{\varepsilon_1} R\Lambda^{\varepsilon_2}\|_{\mathcal{R}}^2 = \|R\|_{\mathcal{R}}^2$, we end up with

$$\|\rho_{1,1}\|_{\mathcal{C}} \leq 2\frac{S_{6,5}C_M C_6}{\pi}\|\rho'\|_{\mathcal{C}}\left(\sum_{\varepsilon_1,\varepsilon_2\in\{\pm\}}\|\Lambda^{\varepsilon_1}R\Lambda^{\varepsilon_2}\|_{\mathcal{R}}\right) \leq 4\frac{S_{6,5}C_M C_6 C_R}{\pi}\|\rho'\|_{\mathcal{C}}\|Q\|_{\mathcal{Q}}.$$

4.3.3 The general n^{th} order case

Now that we have explained how the proof works for the second order, let us estimate the general n^{th} order term.

Lemma I.13. *We have the following estimates*

$$\forall n \geq 3, \qquad \|Q_n\|_{\mathcal{Q}} \leq nK_{\frac{n}{2}}C_Q\|(Q,\rho')\|^n, \quad \textit{with} \quad C_Q = \sqrt{2}\left(\frac{S_6C_6}{2\sqrt{\pi}}\right)^3,$$

$$\forall n \geq 5, \qquad \|\rho_n\|_{\mathcal{C}} \leq nK_{\frac{n+1}{2}}C_\rho\|(Q,\rho')\|^n, \quad \textit{with} \quad C_\rho = \frac{S_6C_6}{4\pi}\left(\frac{S_6C_6}{2\sqrt{\pi}}\right)^5,$$

$$\|\rho_4\|_{\mathcal{C}} \leq C_{\rho_4}\|(Q,\rho')\|^4, \quad \textit{with} \quad C_{\rho_4} := \frac{K_2S_6C_6}{\pi}\left(\frac{S_6C_6}{2\sqrt{\pi}}\right)^2.$$

Therefore,

$$\forall n \geq 4, \qquad \|(Q_n,\rho_n)\| \leq \kappa_n\|(Q,\rho')\|^n$$

with

$$\kappa_4 = 4C_RK_2C_Q\sqrt{2} + 2\sqrt{\pi}C_{\rho_4}, \qquad \kappa_n = nC_RK_{\frac{n}{2}}C_Q\sqrt{2} + 2nK_{\frac{n+1}{2}}C_\rho\sqrt{\pi}.$$

Remark that it can be proved that $K_n \sim_{n\to\infty} \frac{C}{\sqrt{n}}$, which gives the claimed behaviour for κ_n as $n \to \infty$.

Proof – **Step 1 : Estimates on the exchange term Q_n.**

- *$Q_{k,l}$ with $k \geq 1$ and $k + l = n \geq 3$.* Recall that

$$Q_{k,l} = \frac{(-1)^{l+1}}{2\pi} \sum_{I\cup J=\{1,\dots,n\},\ |I|=k,\ |J|=l} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0 + i\eta} \prod_{j=1}^{n} \left(R_j \frac{1}{D^0 + i\eta} \right),$$

where $R_j = R$ if $j \in I$ and $R_j = \varphi'$ if $j \in J$. For the sake of simplicity, we treat only

$$Q'_{k,l} = \frac{1}{2\pi} \int_{-\infty}^{+\infty} d\eta \frac{1}{D^0 + i\eta} \left(R \frac{1}{D^0 + i\eta} \right)^k \left(\varphi' \frac{1}{D^0 + i\eta} \right)^l.$$

We have

$$\begin{aligned}
|\widehat{Q'_{k,l}}(p,q)| \leq \frac{1}{(2\pi)^{1+\frac{3l}{2}}} \int_{-\infty}^{+\infty} d\eta \int\cdots\int \frac{1}{(E(p)^2+\eta^2)^{1/2}} |\widehat{R}(p,p_1)| \times \\
\times \frac{1}{(E(p_1)^2+\eta^2)^{1/4}} \prod_{j=1}^{k-1} \left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}} |\widehat{R}(p_j,p_{j+1})| \frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}} \right) \times \\
\times \prod_{j=k}^{n-2} \left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}} |\widehat{\varphi'}(p_j - p_{j+1})| \frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}} \right) \times \\
\times \frac{1}{(E(p_{n-1})^2+\eta^2)^{1/4}} |\widehat{\varphi'}(p_{n-1} - q)| \frac{1}{(E(q)^2+\eta^2)^{1/2}} \, dp_1 \cdots dp_{n-1}
\end{aligned}$$

so by $(I.45)$,

$$\begin{aligned}
E(p-q)E(p+q)^{1/2} |\widehat{Q'_{k,l}}(p,q)| \leq \frac{2^{\frac{k+1}{2}} E(p-q)}{(2\pi)^{1+\frac{3l}{2}}} \int_{-\infty}^{+\infty} \frac{d\eta}{E(\eta)^{\frac{k+1}{2}}} \int\cdots\int \frac{|\widehat{R}(p,p_1)|}{E(p+p_1)^{1/2}} \times \\
\times \prod_{j=1}^{k-1} \frac{|\widehat{R}(p_j,p_{j+1})|}{E(p_j+p_{j+1})^{1/2}} \prod_{j=k}^{n-2} \left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}} |\widehat{\varphi'}(p_j - p_{j+1})| \frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}} \right) \times \\
\times \frac{1}{(E(p_{n-1})^2+\eta^2)^{1/4}} |\widehat{\varphi'}(p_{n-1} - q)| \frac{1}{(E(q)^2+\eta^2)^{1/4}} \, dp_1 \cdots dp_{n-1}.
\end{aligned}$$

Now if we use the easy generalization of $(I.18)$,

$$E(p-q) \leq E(p-p_1) + E(p_1-p_2) + \cdots + E(p_{n-2}-p_{n-1}) + E(p_{n-1}-q),$$

we obtain by a similar argument as before

$$\begin{aligned}
\|Q'_{k,l}\|_{\mathcal{Q}} \leq \frac{2^{\frac{k+1}{2}}}{2\pi} \Bigg[k(C_R)^k (4\pi C_\infty)^l \left(\int_{-\infty}^{\infty} \frac{d\eta}{E(\eta)^{l+\frac{k+1}{2}}} \right) \\
+ l(C_R)^k (4\pi C_\infty)^{l-1} S_6 C_6 \left(\int_{-\infty}^{\infty} \frac{d\eta}{E(\eta)^{l+\frac{k}{2}}} \right) \Bigg] \|Q\|_{\mathcal{Q}}^k \|\rho'\|_{\mathcal{C}}^l.
\end{aligned}$$

To obtain this result, we have estimated each term containing a φ' by using

$$\|\frac{1}{(|D_0|^2+E(\eta)^2)^{1/4}}\varphi'\frac{1}{(|D_0|^2+E(\eta)^2)^{1/4}}\|_{\mathfrak{S}_\infty} \leq \frac{1}{E(\eta)}\|\varphi'\|_{L^\infty} \leq \frac{C_\infty}{E(\eta)}\|\varphi'\|_{\mathcal{Y}}$$

and when $E(p_j-p_{j+1})$ appears in front of a $\widehat{\varphi'}(p_j-p_{j+1})$ (i.e. when $j\geq k$), by using

$$\|\frac{1}{(|D_0|^2+E(\eta)^2)^{1/4}}f\frac{1}{(|D_0|^2+E(\eta)^2)^{1/4}}\|_{\mathfrak{S}_\infty}$$
$$\leq \|\frac{1}{(|D_0|^2+E(\eta)^2)^{1/4}}f\frac{1}{(|D_0|^2+E(\eta)^2)^{1/4}}\|_{\mathfrak{S}_6} \leq \frac{S_6}{E(\eta)^{1/2}4\pi}\|f\|_{L^6}.$$

So we have

$$\begin{aligned}\|Q'_{k,l}\|_{\mathcal{Q}} &\leq 2^{\frac{k+1}{2}}\left[k(C_R)^k(4\pi C_\infty)^l K_{l+\frac{k+1}{2}} + l(C_R)^k(4\pi C_\infty)^{l-1}S_6C_6K_{l+\frac{k}{2}}\right]\|Q\|_{\mathcal{Q}}^k\|\rho'\|_{\mathcal{C}}^l \\ &\leq 2^{\frac{k+1}{2}}n(C_R)^k(4\pi C_\infty)^l K_{l+\frac{k}{2}}\max\left(1,\frac{S_6C_6}{4\pi C_\infty}\right)\|Q\|_{\mathcal{Q}}^k\|\rho'\|_{\mathcal{C}}^l\end{aligned}$$

which implies

$$\|Q_{k,l}\|_{\mathcal{Q}} \leq \tbinom{n}{k}\,2^{\frac{k+1}{2}}n(C_R)^k(4\pi C_\infty)^l K_{l+\frac{k}{2}}\max\left(1,\frac{S_6C_6}{4\pi C_\infty}\right)\|Q\|_{\mathcal{Q}}^k\|\rho'\|_{\mathcal{C}}^l.$$

- $Q_{0,n}$ *with* $n\geq 3$. Recall that

$$Q_{0,n} = \frac{(-1)^{n+1}}{2\pi}\int_{-\infty}^{+\infty}d\eta\frac{1}{D^0+i\eta}\left(\varphi'\frac{1}{D^0+i\eta}\right)^n$$

so that

$$E(p-q)E(p+q)^{1/2}|\widehat{Q_{0,n}}(p,q)| \leq \frac{E(p-q)}{(2\pi)^{1+\frac{3n}{2}}}\int_{-\infty}^{+\infty}\frac{d\eta}{E(\eta)^{\frac{1}{2}}}\int\cdots\int dp_1\cdots dp_{n-1}\times$$
$$\times\frac{|\widehat{\varphi'}(p-p_1)|}{(E(p)^2+\eta^2)^{1/4}(E(p_1)^2+\eta^2)^{1/4}}\prod_{j=1}^{n-2}\left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}}|\widehat{\varphi'}(p_j-p_{j+1})|\frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}}\right)\times$$
$$\times\frac{1}{(E(p_{n-1})^2+\eta^2)^{1/4}}|\widehat{\varphi'}(p_{n-1}-q)|\frac{1}{(E(q)^2+\eta^2)^{1/4}}.$$

We now use $(I.46)$ to bound for some f_1, f_2 and f_3

$$\|\prod_{j=1}^{3}\left(\frac{1}{(|D_0|^2+\eta^2)^{1/4}}f_j\frac{1}{(|D_0|^2+\eta^2)^{1/4}}\right)\|_{\mathfrak{S}^2} \leq \frac{(S_6)^3}{E(\eta)^{3/2}(4\pi)^3}\prod_{j=1}^{3}\|f_j\|_{L^6} \qquad \text{(I.48)}$$

to obtain

$$\|Q_{0,n}\|_{\mathcal{Q}} \leq n\frac{\sqrt{2}}{2\pi}\left(\int_{-\infty}^{\infty}\frac{d\eta}{E(\eta)^{2+(n-3)}}\right)(S_6C_6)^3(4\pi C_\infty)^{n-3}\|\rho'\|_{\mathcal{C}}^n$$

or

$$\|Q_{0,n}\|_{\mathcal{Q}} \leq nK_{n-1}\sqrt{2}(S_6C_6)^3(4\pi C_\infty)^{n-3}\|\rho'\|_{\mathcal{C}}^n.$$

Finally, we can write for instance (recall that $C_\infty = 1/(2\sqrt{\pi})$)

$$\|Q_n\|_{\mathcal{Q}} \leq nK_{n/2}C_Q\left(C_R\sqrt{2}\|Q\|_{\mathcal{Q}} + 2\sqrt{\pi}\|\rho'\|_{\mathcal{C}}\right)^n \tag{I.49}$$

where

$$C_Q = \sqrt{2}\max\left(1, \frac{S_6C_6}{4\pi C_\infty}, \left(\frac{S_6C_6}{4\pi C_\infty}\right)^3\right) = \sqrt{2}\frac{(S_6C_6)^3}{8\pi^{3/2}},$$

and since $K_{n/2} \geq K_{n-1}$ when $n \geq 2$.

Step 2 : Estimates on the density ρ_n.

• *$\rho_{k,l}$ with $k \geq 2$ and $n \geq 3$.* As before we treat for instance the density $\rho'_{k,l}$ of the $Q'_{k,l}$ where the k R's are on the left and the l φ''s are on the right. For some fixed $\zeta \in \mathcal{C}' \cap L^2$, we introduce $Q_\zeta := Q'_{k,l}\zeta$. We thus estimate

$$\begin{aligned}
|\widehat{Q_\zeta}(p,p)| \leq \frac{1}{(2\pi)^{1+\frac{3(l+1)}{2}}}\int_{-\infty}^{+\infty} d\eta \int\cdots\int \frac{1}{(E(p)^2+\eta^2)^{1/4}}|\widehat{R}(p,p_1)|\times \\
\times \frac{1}{(E(p_1)^2+\eta^2)^{1/4}}\prod_{j=1}^{k-1}\left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}}|\widehat{R}(p_j,p_{j+1})|\frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}}\right)\times \\
\times \prod_{j=k}^{n-1}\left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}}|\widehat{\varphi'}(p_j-p_{j+1})|\frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}}\right)\times \\
\times \frac{1}{(E(p_n)^2+\eta^2)^{1/4}}|\widehat{\zeta}(p_n-p)|\frac{1}{(E(p)^2+\eta^2)^{1/4}}\,dp_1\cdots dp_n.
\end{aligned}$$

We now use as before

$$|\widehat{\zeta}(p_n-p)| \leq \frac{|\widehat{\zeta}(p_n-p)|}{E(p_n-p)}\big(E(p-p_1)+E(p_1-p_2)+\cdots+E(p_{n-1}-p_n)\big)$$

to obtain

$$|\langle\rho'_{k,l},\zeta\rangle| \leq n\frac{2^{k/2}S_6C_6}{4\pi}(C_R)^k(4\pi C_\infty)^l\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d\eta}{E(\eta)^{l+\frac{k+1}{2}}}\right)\|Q\|_{\mathcal{Q}}^k\|\rho'\|_{\mathcal{C}}^l\|\zeta\|_{\mathcal{C}'}$$

and so

$$\|\rho_{k,l}\|_{\mathcal{C}} \leq n\binom{n}{k}\frac{S_6C_6}{4\pi}(C_R\sqrt{2})^k(4\pi C_\infty)^l K_{l+\frac{k+1}{2}}\|Q\|_{\mathcal{Q}}^k\|\rho'\|_{\mathcal{C}}^l.$$

• *$\rho_{1,l}$ with $l \geq 2$.* We may treat for instance with the same notation as before

$$\begin{aligned}
|\widehat{Q_\zeta}(p,p)| \leq \frac{1}{(2\pi)^{1+\frac{3n}{2}}}\int_{-\infty}^{+\infty} d\eta \int\cdots\int \frac{1}{(E(p)^2+\eta^2)^{1/4}}|\widehat{R}(p,p_1)|\times \\
\times \frac{1}{(E(p_1)^2+\eta^2)^{1/4}}\prod_{j=1}^{n-1}\left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}}|\widehat{\varphi'}(p_j-p_{j+1})|\frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}}\right)\times \\
\times \frac{1}{(E(p_n)^2+\eta^2)^{1/4}}|\widehat{\zeta}(p_n-p)|\frac{1}{(E(p)^2+\eta^2)^{1/4}}\,dp_1\cdots dp_n.
\end{aligned}$$

We now use $(I.48)$ and obtain

$$|\langle \rho'_{1,l}, \zeta\rangle| \leq n\frac{2^{1/2}(S_6C_6)^3}{4\pi} C_R(4\pi C_\infty)^{l-2}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d\eta}{E(\eta)^{l-2+\frac{4}{2}}}\right)\|Q\|_{\mathcal{Q}}\|\rho'\|_{\mathcal{C}}^l\|\zeta\|_{\mathcal{C}'}$$

and so

$$\|\rho_{1,l}\|_{\mathcal{C}} \leq n\binom{n}{n-1}\frac{2^{1/2}(S_6C_6)^3K_l}{4\pi} C_R(4\pi C_\infty)^{l-2}\|Q\|_{\mathcal{Q}}\|\rho'\|_{\mathcal{C}}^l.$$

- $\rho_{0,l}$ *with* $l \geq 5$. We want to estimate

$$\begin{aligned}|\widehat{Q_\zeta}(p,p)| \leq \frac{1}{(2\pi)^{1+\frac{3(n+1)}{2}}}\int_{-\infty}^{+\infty}d\eta\int\cdots\int\frac{1}{(E(p)^2+\eta^2)^{1/4}}|\widehat{\varphi'}(p-p_1)|\times\\ \times\frac{1}{(E(p_1)^2+\eta^2)^{1/4}}\prod_{j=1}^{n-1}\left(\frac{1}{(E(p_j)^2+\eta^2)^{1/4}}|\widehat{\varphi'}(p_j-p_{j+1})|\frac{1}{(E(p_{j+1})^2+\eta^2)^{1/4}}\right)\times\\ \times\frac{1}{(E(p_n)^2+\eta^2)^{1/4}}|\widehat{\zeta}(p_n-p)|\frac{1}{(E(p)^2+\eta^2)^{1/4}}\,dp_1\cdots dp_n.\end{aligned}$$

Since there are at least 6 functions, we may use $(I.48)$ twice and obtain

$$|\langle \rho'_{0,l}, \zeta\rangle| \leq n\frac{(S_6C_6)^6}{4\pi}(4\pi C_\infty)^{l-5}\left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d\eta}{E(\eta)^{l-2}}\right)\|\rho'\|_{\mathcal{C}}^n\|\zeta\|_{\mathcal{C}'}$$

and so

$$\|\rho_{0,l}\|_{\mathcal{C}} \leq n\frac{(S_6C_6)^6}{4\pi}(4\pi C_\infty)^{l-5}K_{l-2}\|\rho'\|_{\mathcal{C}}^n.$$

Now since $K_{n-2} \leq K_{(n+1)/2}$ when $n \geq 5$, we obtain

$$\|\rho_n\|_{\mathcal{C}} \leq nK_{\frac{n+1}{2}}C_\rho\left(C_R\sqrt{2}\|Q\|_{\mathcal{Q}} + 2\sqrt{\pi}\|\rho'\|_{\mathcal{C}}\right)^n \tag{I.50}$$

with

$$C_\rho := \frac{S_6C_6}{4\pi}\max\left(1, \left(\frac{C_6S_6}{2\sqrt{\pi}}\right)^2, \left(\frac{C_6S_6}{2\sqrt{\pi}}\right)^5\right) = \frac{S_6C_6}{4\pi}\left(\frac{C_6S_6}{2\sqrt{\pi}}\right)^5.$$

For ρ_4, we notice that $\rho_{0,4} = 0$ for the same reason as $\rho_{0,2}$, and that $K_2 \geq K_{5/2}$. Therefore we obtain

$$\|\rho_4\|_{\mathcal{C}} \leq C_{\rho_4}\left(C_R\sqrt{2}\|Q\|_{\mathcal{Q}} + 2\sqrt{\pi}\|\rho'\|_{\mathcal{C}}\right)^4 \tag{I.51}$$

with

$$C_{\rho_4} := \frac{4K_2S_6C_6}{4\pi}\max\left(1, \left(\frac{C_6S_6}{2\sqrt{\pi}}\right)^2\right) = \frac{S_6C_6K_2}{\pi}\left(\frac{C_6S_6}{2\sqrt{\pi}}\right)^2. \quad \square$$

4.3.4 The third order density ρ_3

Lemma I.14. *We have*

$$\|\rho_3\|_{\mathcal{C}} \leq C_{\rho_3}\left(C_R\sqrt{2}\|Q\|_{\mathcal{Q}} + 2\sqrt{\pi}\|\rho'\|_{\mathcal{C}}\right)^3$$

and therefore

$$\|(Q_3, \rho_3)\| \leq \kappa_3\|(Q, \rho')\|^3$$

with

$$\kappa_3 = 3C_RK_{3/2}C_Q\sqrt{2} + 2\sqrt{\pi}C_{\rho_3}, \qquad C_{\rho_3} = \frac{15C_MS_6(S_{6,4})^2(C_6)^4}{\pi(4\pi C_\infty)^3}.$$

Proof – Notice that thanks to the previous proof, we already have some estimates on $\rho_{3,0}$, $\rho_{2,1}$ and $\rho_{1,2}$. It remains to study $\rho_{0,3}$. As before and as in [HS03], we have to compute $\rho_{0,3}$ explicitely by a residuum formula. We thus write

$$\rho_{0,3} = \sum_{\varepsilon_1,\dots,\varepsilon_4\in\{\pm\}} \rho_{0,3}^{\varepsilon_1\varepsilon_2\varepsilon_3\varepsilon_4}$$

with an obvious definition.

• Let us treat first $\rho_{0,3}^{+---}$. We thus fix some $\zeta \in \mathcal{C}' \cap L^2$ and estimate the term

$$\widehat{Q_\zeta}(p,p) = (2\pi)^{-6} \iiint dp_1\, dp_2\, dp_3 \frac{\Lambda^+(p)\widehat{\varphi'}(p-p_1)\Lambda^-(p_1)}{E(p)+E(p_1)} \times \\ \times \frac{\widehat{\varphi'}(p_1-p_2)\Lambda^-(p_2)\widehat{\varphi'}(p_2-p_3)\Lambda^-(p_3)}{(E(p)+E(p_2))(E(p)+E(p_3))} \widehat{\zeta}(p_3-p),$$

by

$$|\widehat{Q_\zeta}(p,p)| \leq (2\pi)^{-6} \iiint dp_1\, dp_2\, dp_3 \frac{\left|\Lambda^+(p)\widehat{\varphi'}(p-p_1)\Lambda^-(p_1)\right|}{E(p+p_1)^{2/3}} \times \\ \times \frac{\left|\widehat{\varphi'}(p_1-p_2)\right| \left|\widehat{\varphi'}(p_2-p_3)\right|}{E(p_1)^{1/3}E(p_2)E(p)} \left|\widehat{\zeta}(p_3-p)\right|.$$

So if we follow the method used above, we obtain

$$\|\rho_{0,3}^{+---}\|_{\mathcal{C}} \leq \frac{3C_M S_6 (S_{6,4})^2 (C_6)^4}{4\pi} \|\rho'\|_{\mathcal{C}}^3.$$

Now, it is easily seen that $\rho_{0,3}^{-+++}$, $\rho_{0,3}^{--+-}$, $\rho_{0,3}^{+++-}$, $\rho_{0,3}^{-+--}$, $\rho_{0,3}^{+-++}$, $\rho_{0,3}^{--+-}$ and $\rho_{0,3}^{++-+}$ can be treated by exactly the same method.

• Let us now treat for instance $\rho_{0,3}^{++--}$. Thanks to the residuum formula, we have to study

$$\widehat{Q_\zeta}(p,p) = (2\pi)^{-6} \iiint dp_1\, dp_2\, dp_3 \Lambda^+(p)\widehat{\varphi'}(p-p_1)\Lambda^+(p_1)\widehat{\varphi'}(p_1-p_2)\Lambda^-(p_2) \\ \widehat{\varphi'}(p_2-p_3)\Lambda^-(p_3)\widehat{\zeta}(p_3-p) \times \left(\frac{1}{(E(p)+E(p_2))(E(p_1)+E(p_2))(E(p_1)+E(p_3))} + \right. \\ \left. \frac{1}{(E(p)+E(p_2))(E(p)+E(p_3))(E(p_1)+E(p_3))} \right).$$

If we now use the same method as above for each of the two terms of this sum, we arrive at

$$\|\rho_{0,3}^{++--}\|_{\mathcal{C}} \leq 2\frac{3C_M S_6 (S_{6,4})^2 (C_6)^4}{4\pi} \|\rho'\|_{\mathcal{C}}^3.$$

This is easily generalized to the study of $\rho_{0,3}^{--++}$, $\rho_{0,3}^{+-+-}$, $\rho_{0,3}^{-+-+}$, $\rho_{0,3}^{+--+}$ and $\rho_{0,3}^{-++-}$.

Summing now all these terms, we obtain

$$\|\rho_{0,3}\|_{\mathcal{C}} \leq 20\frac{3C_M S_6 (S_{6,4})^2 (C_6)^4}{4\pi} \|\rho'\|_{\mathcal{C}}^3 = \frac{15C_M S_6 (S_{6,4})^2 (C_6)^4}{\pi} \|\rho'\|_{\mathcal{C}}^3$$

and

$$\|\rho_3\|_{\mathcal{C}} \leq C_{\rho_3} \left(C_R \sqrt{2} \|Q\|_{\mathcal{Q}} + 2\sqrt{\pi} \|\rho'\|_{\mathcal{C}} \right)^3, \tag{I.52}$$

with

$$\begin{aligned} C_{\rho_3} &= 3\frac{S_6 C_6}{4\pi} \max \left(K_{3+1/2}\,,\, K_3\,,\, K_2 \left(\frac{S_6 C_6}{4\pi C_\infty} \right)^2, \frac{20 C_M (S_{6,4})^2 (C_6)^3}{(4\pi C_\infty)^3} \right) \\ &= \frac{15 C_M S_6 (S_{6,4})^2 (C_6)^4}{\pi (4\pi C_\infty)^3}. \quad \square \end{aligned}$$

Appendix: Derivation of the BDF energy

In this section, we recall some basics about the second-quantization in no-photon QED and explain how the energy $\mathcal{E}^{\mathrm{BDF}}$ is derived from this theory, as a mean-field approximation. We mainly follow the method of Chaix-Iracane [CI89, Cha90], but with the notation of [Tha92, BBHS99, HS98]. See also [HS03] for more details concerning the polarization of the vacuum.

Since this derivation is somewhat formal, we may work in $\mathcal{H} = L^2(\mathbb{R}^3, \mathbb{C}^4)$. To simplify the presentation, we introduce $P^0_- := P^0$ and $P^0_+ := 1 - P^0$.

Free particles, Fock space, free vacuum

We first introduce $\mathcal{F}^{(1)}_+ := P^0_+\mathcal{H}$ and $\mathcal{F}^{(1)}_- := CP^0_-\mathcal{H}$ which are called respectively the free electron and the free positron state subspace. C is the charge-conjugation operator defined by $C\psi = i\beta\alpha_2\overline{\psi}$. We define $\mathcal{F}^{(0)}_+ = \mathcal{F}^{(0)}_- = \mathbb{C}$ and $\mathcal{F}^{(n)}_+ = \bigwedge^n_{k=1}\mathcal{F}^{(1)}_+$, $\mathcal{F}^{(m)}_- = \bigwedge^m_{k=1}\mathcal{F}^{(1)}_-$ for $n, m \geq 1$. The space of n free electrons and m free positrons is then defined by $\mathcal{F}^{(n,m)} = \mathcal{F}^{(n)}_+ \otimes \mathcal{F}^{(m)}_-$ and the associated Fock space is

$$\mathcal{F} := \bigoplus_{n,m=0}^{\infty} \mathcal{F}^{(n,m)}. \tag{I.53}$$

For any $f \in \mathcal{H}$, we define the free electron annihilation and creation operators $a_0(f)$ and $a_0^*(f)$. The free electron annihilation operator $a_0(f)$ fulfills $a_0(f)(\mathcal{F}^{(n,m)}) \subset \mathcal{F}^{(n-1,m)}$ for $n \geq 1$ and is defined on $\mathcal{F}^{(n,m)}$ by

$$a_0(f)(\psi)(x_1, ..., x_{n-1}; y_1, ..., y_m) = \sqrt{n+1}\int_{\mathbb{R}^3} \overline{P^0_+(f)} \cdot \psi(x, x_1, ..., x_{n-1}; y_1, ..., y_m)\, dx,$$

and $a_0(f)(\mathcal{F}^{(0,m)}) = 0$. The free electron creation operator $a_0^*(f)$ fulfills $a_0^*(f)(\mathcal{F}^{(n,m)}) \subset \mathcal{F}^{(n+1,m)}$ for $n \geq 0$ and is defined on tensor products of $\mathcal{F}^{(n,m)}$ by

$$a_0^*(f)(\psi_+ \otimes \psi_-) = \Big(P^0_+(f) \wedge \psi_+\Big) \otimes \psi_-$$

when $\psi_+ \in \mathcal{F}^{(n)}_+$ and $\psi_+ \in \mathcal{F}^{(m)}_-$, or more explicitely for $\psi \in \mathcal{F}^{(n,m)}$

$$a_0^*(f)(\psi)(x_1, ..., x_{n+1}; y_1, ..., y_m) = \frac{1}{\sqrt{n}}\sum_{j=1}^{n}(-1)^{j+1}P^0_+(f)(x_j)\psi(x_1, ..., x_{j-1}, x_{j+1}, ..., x_n; y_1, ..., y_m).$$

Correspondingly, the free positron annihilation and creation operators $b_0(f)$ and $b_0^*(f)$ are defined by

$$b_0(f)(\psi)(x_1, ..., x_n; y_1, ..., y_{m-1}) = (-1)^n\sqrt{m+1}\int_{\mathbb{R}^3} \overline{CP^0_-(f)}\cdot\psi(x_1, ..., x_n; y, y_1, ..., y_{m-1})\, dx,$$

$$b_0(f)(\mathcal{F}^{(n,0)}) = 0, \qquad b_0^*(f)(\psi_+ \otimes \psi_-) = (-1)^n\psi_+ \otimes \Big(CP^0_-(f) \wedge \psi_-\Big)$$

when $\psi_+ \in \mathcal{F}^{(n)}_+$ and $\psi_+ \in \mathcal{F}^{(m)}_-$. As before, $b_0(f)(\mathcal{F}^{(n,m)}) \subset \mathcal{F}^{(n,m-1)}$ and $b_0^*(f)(\mathcal{F}^{(n,m)}) \subset \mathcal{F}^{(n,m+1)}$. Remark that a_0 and b_0^* are anti-linear, while a_0^* and b_0 are linear. They fulfill

the Canonical Anti-commutation Relations

$$\{a_0(f), a_0(g)\} = \{a_0^*(f), a_0^*(g)\} = \{a_0(f), b_0(g)\} \\ = \{a_0^*(f), b_0^*(g)\} = \{a_0^*(f), b_0(g)\} = \{a_0(f), b_0^*(g)\} = 0, \quad \text{(I.54)}$$

$$\{a_0(f), a_0^*(g)\} = \langle f, P_+^0 g\rangle, \qquad \{b_0^*(f), b_0(g)\} = \langle f, P_-^0 g\rangle, \quad \text{(I.55)}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual scalar product of $L^2(\mathbb{R}^3, \mathbb{C}^4)$.

For the *free vacuum state* Ω_0, a unit vector spanning $\mathcal{F}^{(0,0)} = \mathbb{C}$, we obtain

$$a_0(f)\Omega_0 = 0 \text{ and } b_0(f)\Omega_0 = 0, \quad \text{(I.56)}$$

a property that caracterizes uniquely Ω_0 up to a phase factor.

Let us now define the *field operator* $\Psi(f)$ on the Fock space $\mathcal{F}$ by

$$\Psi(f) = a_0(f) + b_0^*(f).$$

In terms of $\Psi(f)$, the CAR become

$$\{\Psi(f), \Psi(g)\} = \{\Psi^*(f), \Psi^*(g)\} = 0, \qquad \{\Psi(f), \Psi^*(g)\} = \langle f, g\rangle 1,$$

for all $(f, g) \in \mathcal{H}^2$.

Remark that we have by definition

$$\Psi(P_+^0 f) = a_0(f), \qquad \Psi^*(P_-^0 f) = b_0(f) \quad \text{(I.57)}$$

and $\Psi(P_+^0 f)\Omega_0 = \Psi^*(P_-^0 f)\Omega_0 = 0$, which also caracterizes the free vacuum Ω_0 up to a phase factor.

Dressed particles and vacuum

In this description, the free electrons and positrons are defined with respect to the projector P^0, or equivalently the splitting $\mathcal{H} = \mathcal{H}_-^0 \oplus \mathcal{H}_+^0$. We want now to change this definition and introduce the *dressed electrons and positrons*. To this end, we fix a new projector P on $\mathcal{H}$, use again the notation $P_- := P$ and $P_+ := 1 - P$, and introduce the dressed particle annihilation operators

$$a_P(f) := \Psi(P_+ f), \qquad b_P(f) := \Psi^*(P_- f), \quad \text{(I.58)}$$

by using $(I.57)$. Similar formula can be given for the dressed particle creation operators a_P^* and b_P^*. These dressed operators satisfy the same anti-commutation relations as for the free operators $(I.54)$ $(I.55)$. We also introduce the dressed electrons and positrons state subspaces

$$\mathcal{H}_+^P = (1 - P)\mathcal{H}, \qquad \mathcal{H}_-^P = P\mathcal{H}.$$

Now, the main question is to know if there exists a *dressed vacuum* Ω_P in the Fock space $\mathcal{F}$. This state has to be a solution to the analogue of $(I.56)$

$$a_P(f)\Omega_P = 0 \text{ and } b_P(f)\Omega_P = 0 \quad \text{(I.59)}$$

for all $f \in \mathcal{H}$. The answer is given by the celebrated Shale-Stinespring Theorem (see Thaller's book [Tha92, Theorem 10.6])

Theorem I.5 (Shale-Stinespring Theorem). *There exists a dressed vacuum Ω_P in the Fock space $\mathcal{F}$ satisfying $(I.59)$ if and only if $P - P^0$ is a Hilbert-Schmidt operator. In this case, Ω_P is unique up to a phase factor.*

This result explains why we assumed in the previous section that $P - P^0 \in \mathfrak{S}_2(\mathcal{H}_\Lambda)$.

Second-quantized Hamiltonian

In the physics literature, the creation and annihilation operators are defined differently. For instance, instead of $a_0^*(f)$ which creates an electron in the state $P_+^0 f$, the operator $a_0^*(x)$ which creates an electron at x is formally used, where $a_0^*(x) = \sum_{i=1}^\infty a_0^*(f_i) f_i(x)$, $(f_i)_{i\geq 1}$ being an orthonormal basis of $\mathcal{H}_+^0$. The operators $a_P(x)$ and $b_P(x)$ are defined similarly.

We shall now use this formalism and follow mainly [HS03]. Formally, the CAR $(I.54, I.55)$ are equivalent to

$$\begin{aligned}\{a_0(x), a_0(y)\} = \{a_0^*(x), a_0^*(y)\} = \{a_0(x), b_0(y)\} \\ = \{a_0^*(x), b_0^*(y)\} = \{a_0^*(x), b_0(y)\} = \{a_0(x), b_0^*(y)\} = 0, \quad \text{(I.60)}\end{aligned}$$

$$\{a_0(x), a_0^*(y)\} = [1 - P^0](x, y),\ \{b_0^*(x), b_0(y)\} = P^0(x, y) \quad \text{(I.61)}$$

where $[1 - P^0](x, y)$, $P^0(x, y)$ are the integral kernels of the projectors $1 - P$ and P.

We now start with writing down the formal unregularized no-photon Hamiltonian

$$\mathbb{H}_{\text{ur}} = \int dx\, \Psi^*(x) D^{\alpha\varphi} \Psi(x) + \frac{\alpha}{2} \int dx \int dy \frac{\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y)}{|\mathbf{x} - \mathbf{y}|}, \quad \text{(I.62)}$$

which acts on the Fock space $\mathcal{F}$. As explained for instance in [Tha92, HS98], the free vacuum may not belong to the domain of this formally defined operator. Therefore, the formal energy of the free vacuum is substracted from $(I.62)$ by a procedure which is called "normal ordering", denoted by double dots $: - :_{P^0}$. In each product of annihilation and creation operators, the a_0^* and b_0^* are moved to the left as if they anticommute with the a_0 and b_0. For instance

$$: \Psi^*(x)\Psi(y) :_{P^0} = a_0^*(x)a_0(y) + a_0^*(y)b_0(x) + b_0(x)a_0(y) - b_0^*(y)b_0(x) = \Psi^*(x)\Psi(y) - P^0(x, y). \quad \text{(I.63)}$$

As a first step we thus regularize $\mathbb{H}_{\text{ur}}$ as done in Chaix and Iracane [CI89, Sections 3.5 and 4.1], namely we normal order with respect to the free Projector P^0,

$$\mathbb{H} = \int dx\, :\Psi^*(x) D^{\alpha\varphi} \Psi(x):_{P^0} + \frac{\alpha}{2} \int dx \int dy \frac{:\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y):_{P^0}}{|\mathbf{x} - \mathbf{y}|}. \quad \text{(I.64)}$$

This kind of regularization appears implicitly in any QED textbook. It corresponds to the subtraction of the energy of the free Dirac sea, and the interaction energy with the free Dirac sea, which is usually referred to as mass renormalization. In fact in [HS03, Section 3] these subtractions were justified by the first two guiding principles (denoted as **W1** and **W2**) formulated and justified by Weisskopf [Wei36].

The definition of the normal ordering depends on the annihilation and creation operators and it thus depends on the projector P which is used. Now we want to express $\mathbb{H}$ in terms of $: - :_P$ for some P. Using (I.63) we obtain the reordering relations

$$: \Psi^*(x)\Psi(y) :_{P^0} = : \Psi^*(x)\Psi(y) :_P + Q(x, y), \quad \text{(I.65)}$$

where $Q = P - P^0$, and

$$\begin{aligned}: \Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y) :_{P^0} = &: \Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y) :_P + 2 : \Psi^*(x)\Psi(x) :_P \operatorname{Tr}_{\mathbb{C}^4}(Q(y, y)) \\ &- 2 : \Psi^*(x)\Psi(y) :_P Q(x, y) + \operatorname{Tr}_{\mathbb{C}^4}(Q(x, x)) \operatorname{Tr}_{\mathbb{C}^4}(Q(y, y)) - |Q(x, y)|^2.\end{aligned}$$

Therefore we can rewrite $\mathbb{H}$ with respect to an arbitrary dressed vacuum P,

$$\mathbb{H} = \int :\Psi^*(x)D^{\alpha\varphi}\Psi(x):_P dx + \frac{\alpha}{2}\iint \frac{:\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y):_P}{|x-y|}dx\,dy$$
$$+\alpha\iint \frac{:\Psi^*(x)\Psi(x):_P \mathrm{Tr}_{\mathbb{C}^4}(Q(y,y))}{|x-y|}dx\,dy - \alpha\iint \frac{:\Psi^*(x)\Psi(y):_P Q(x,y)}{|x-y|}dx\,dy$$
$$+\mathrm{Tr}(D^\varphi Q) + \frac{\alpha}{2}\iint \frac{\mathrm{Tr}_{\mathbb{C}^4}(Q(x,x))\mathrm{Tr}_{\mathbb{C}^4}(Q(y,y))}{|x-y|}dx\,dy - \frac{\alpha}{2}\iint \frac{|Q(x,y)|^2}{|x-y|}dx\,dy. \quad (I.66)$$

The last line represents the energy of the dressed vacuum P measured with respect to P^0, whereas in the second line the vacuum polarization potentials appear.

Restriction to Bogoliubov-Dirac-Fock states

A general strategy would now be to fix a charge q and to look at the expectation value of the operator $\mathbb{H}$ in an electron-positron state ψ with total charge q (electrons and positrons are defined with respect to P) and then minimizing over all admissible states ψ and over P.

For a slight simplification, we will follow [CI89] and restrict ourselves to Bogoliubov-Dirac-Fock type states. In this approximation method, a dressed vacuum P is first chosen such that $P - P^0 \in \mathfrak{S}_2(\mathcal{H})$. Then a BDF state is simply a Slater determinant made with n electrons and m positrons defined with respect to the dressed vacuum P ($n, m \geq 0$ are not fixed in this theory). This is a state of $\mathcal{F}$ which takes the form

$$\psi = a_P^*(f_1)\cdots a_P^*(f_n)\, b_P^*(g_1)\cdots b_P^*(g_m)\,\Omega_P,$$

where $(f_1, ..., f_n) \in (\mathcal{H}_+^P)^n$ and $(g_1, ..., g_m) \in (\mathcal{H}_-^P)^m$ are such that $\langle f_i, f_j\rangle = \delta_{ij}$, $\langle g_i, g_j\rangle = \delta_{ij}$, and Ω_P is the dressed vacuum in $\mathcal{F}$ obtained by Theorem I.5. The *one-body density matrix* of such a state is the operator γ_ψ defined by its kernel

$$\gamma_\psi(x,y) = \sum_{i=1}^n f_i(x)\overline{f_i(y)}^T - \sum_{j=1}^m g_j(x)\overline{g_j(y)}^T$$

often written as

$$\gamma_\psi = \sum_{i=1}^n |f_i\rangle\langle f_i| - \sum_{j=1}^m |g_j\rangle\langle g_j|,$$

and it satisfies

$$-P \leq \gamma_\psi \leq 1 - P.$$

The associated *density of charge* is the function $\rho_\psi(x) = \mathrm{Tr}_{\mathbb{C}^4}(\gamma(x,x))$ or

$$\rho_\psi(x) = \sum_{i=1}^n |f_i(x)|^2 - \sum_{j=1}^m |g_j(x)|^2.$$

The energy $\langle\psi|\mathbb{H}|\psi\rangle$ of such a BDF-state $\psi \in \mathcal{F}$ is easily computed with formula $(I.66)$. As usually in this type of approximation, it only depends on the one-body density

matrix γ_ψ (and its spin-summed diagonal ρ_ψ) and it is shown to be (see [BBHS99] for details and [CI89, Formula 4.8])

$$\begin{aligned}\mathcal{E}^{\mathrm{BDF}}(P,\gamma_\psi) &:= \langle\psi|\mathbb{H}|\psi\rangle = \operatorname{tr}\left(\left(D^{\alpha\varphi}+\alpha\rho_Q * \frac{1}{|\cdot|}-\alpha\frac{Q(x,y)}{|x-y|}\right)\gamma_\psi\right)\\ &\quad +\frac{\alpha}{2}\iint\frac{\rho_\psi(x)\rho_\psi(y)}{|x-y|}dx\,dy-\frac{\alpha}{2}\iint\frac{|\gamma_\psi(x,y)|^2}{|x-y|}dx\,dy+\mathcal{E}(Q)\\ &= \mathcal{E}(Q+\gamma_\psi).\end{aligned}$$

Bibliography

[ASS94] J. Avron, R. Seiler, and B. Simon. The index of a pair of projections. *J. Funct. Anal.*, 120:220–237, 1994.

[BBHS99] V. Bach, J.M. Barbaroux, B. Helffer, and H. Siedentop. On the stability of the relativistic electron-positron field. *Commun. Math. Phys.*, 201:445–460, 1999.

[BLS94] V. Bach, E.H. Lieb, and J.P. Solovej. Generalized Hartree-Fock theory and the Hubbard model. *J. Statist. Phys.*, 76(1-2):3–89, 1994.

[BRS01] R. Brummelhuis, N. Röhrl, and H. Siedentop. Stability of the relativistic electron-positron field of atoms in Hartree-Fock approximation: Heavy elements. *Doc. Math.*, 6:1–9, 2001.

[Cha90] P. Chaix. *Une Méthode de Champ Moyen Relativiste et Application à l'Etude du Vide de l'Electrodynamique Quantique*. PhD thesis, 1990.

[CI89] P. Chaix and D. Iracane. From quantum electrodynamics to mean field theory: I. the Bogoliubov-Dirac-Fock formalism. *J. Phys. B.*, 22(23):3791–3814, 1989.

[CIL89] P. Chaix, D. Iracane, and P.L. Lions. From quantum electrodynamics to mean field theory: II. Variational stability of the vacuum of quantum electrodynamics in the mean-field approximation. *J. Phys. B.*, 22(23):3815–3828, 1989.

[Dir34] P.A.M. Dirac. Discussion of the infinite distribution of electrons in the theory of the positron. *Proc. Camb. Philos. Soc.*, 30:150–163, 1934.

[Dys49a] F.J. Dyson. The radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.*, 75(2):486–502, 1949.

[Dys49b] F.J. Dyson. The S matrix in Quantum Electrodynamics. *Phys. Rev.*, 75(11):1736–1755, 1949.

[ES99] M.J. Esteban and E. Séré. Solutions of the Dirac-Fock equations for atoms and molecules. *Comm. Math. Phys.*, 203:499–530, 1999.

[FO34] W.H. Furry and J.R. Oppenheimer. On the theory of the electron and positive. *Phys. Rev., II. Ser.*, 45:245–262, 1934.

[FW49] J.D. French and V.F. Weisskopf. The electromagnetic shift of energy levels. *Phys. Rev., II. Ser.*, 75:1240–1248, 1949.

[GMR85] W. Greiner, B. Müller, and J. Rafelski. *Quantum Electrodynamics of Strong Fields*. Texts and Monographs in Physics. Springer-Verlag, first edition, 1985.

[Hai03] C. Hainzl. On the vacuum polarization density caused by an external field. 2003.

[Hei34] W. Heisenberg. Bemerkungen zur Diracschen Theorie des Positrons. *Z. Phys.*, 90:209–231, 1934.

[HRS00] D. Hundertmark, N. Röhrl, and H. Siedentop. The sharp bound on the stability of the relativistic electron-positron field in Hartree-Fock approximation. *Comm. Math. Phys.*, 211(3):629–642, 2000.

[HS98] B. Helffer and H. Siedentop. Form perturbation of the second quantized Dirac field. *Mathematical Physics Electronic Journal*, 4, 1998.

[HS03] C. Hainzl and H. Siedentop. Non-perturbative mass and charge renormalization in relativistic no-photon quantum electrodynamics. *Comm. Math. Phys.*, 243:241–260, 2003.

[IZ80] C. Itzykson and J.-B. Zuber. *Quantum Field Theory*. McGraw-Hill, New York, 1980.

[KS77] M. Klaus and G. Scharf. The regular external field problem in quantum electrodynamics. *Helv. Phys. Acta*, 50:779–802, 1977.

[LS00] E.H. Lieb and H. Siedentop. Renormalization of the regularized relativistic electron-positron field. *Comm. Math. Phys.*, 213(3):673–683, 2000.

[Mil94] P.W. Milonni. *The Quantum Vacuum: An Introduction to Quantum Electrodynamics*. Academic Press, Inc., Boston, 1994.

[Mit81] M.H. Mittleman. Theory of relativistic effects on atoms: Configuration-space hamiltonian. *Phys. Rev. A*, 24(3):1167–1175, 1981.

[MPS98] P. J. Mohr, G. Plunien, and G. Soff. QED corrections in heavy atoms. *Phys. Rep.*, 293:227–369, 1998.

[PR36] W. Pauli and M.E. Rose. Remarks on the polarization effects in the positron theory. *Phys. Rev II*, 49:462–465, 1936.

[Sch49] J. Schwinger. Quantum electrodynamics. II. Vacuum polarization and self-energy. *Phys. Rev., II. Ser.*, 75:651–679, 1949.

[Ser35] R. Serber. Linear modifications in the Maxwell field equations. *Phys. Rev., II. Ser.*, 48:49–54, 1935.

[Sim79] B. Simon. *Trace Ideals and their Applications*, volume 35 of *London Mathematical Society Lecture Notes Series*. Cambridge University Press, 1979.

[Tha92] B. Thaller. *The Dirac Equation*. Springer Verlag, 1992.

[Ueh35] E.A. Uehling. Polarization effects in the positron theory. *Phys. Rev., II. Ser.*, 48:55–63, 1935.

[Wei36] V. Weisskopf. Über die Elektrodynamik des Vakuums auf Grund der Quantentheorie des Elektrons. *Math.-Fys. Medd., Danske Vid. Selsk.*, 16(6):1–39, 1936.

[Wei96] S. Weinberg. *The quantum theory of fields*, volume I. Cambridge University Press, 1996.

[Wol73] J.H. Wolkowisky. Existence of solutions of the Hartree equations for n electrons. an application of the Schauder-Tychonoff theorem. *Indiana Univ. Math. J.*, 22:551–568, 1972-73.

[a] CEREMADE, CNRS UMR 7534,
Université Paris IX Dauphine,
Place du Maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France.
hainzl, lewin, sere@ceremade.dauphine.fr

[b] Laboratoire de Mathématiques,
Université Paris-Sud,
Bât. 425,
91 405 Orsay Cedex, France.

C.II

Self-Consistent Solution for the Polarized Vacuum in a No-Photon QED Model

Ce chapitre reprend le texte intégral d'un article écrit en collaboration avec Christian Hainzl et Éric Séré. Il est destiné à un public de physiciens.

Résumé

Dans cet article, en suivant des idées de Chaix-Iracane [CI89], nous déduisons des premiers principes de l'électrodynamique quantique négligeant les photons, une fonctionnelle qui permet la description rigoureuse de la polarisation du vide. Nous présentons également un nouvel outil permettant de définir la charge du vide. Enfin, nous proposons une renormalisation de charge qui s'applique simultanément à tous les ordres du développement par rapport à la constante de structure fine α.

Self-Consistent Solution for the Polarized Vacuum in a No-Photon QED Model[1]

Christian Hainzl[a,b], Mathieu Lewin[a] & Éric Séré[a]

Abstract

Starting from a first principle QED Hamiltonian, neglecting photons, we derive a functional – following ideas of Chaix-Iracane [CI89] – which allows us to rigorously describe the polarized vacuum in a self-consistent way. Thereby we develop novel means to characterize charge sectors. Additionally we perform a charge renormalization scheme applying simultaneously to all orders of the fine structure constant α.

1 Introduction

Despite the incredible predictive power of Quantum Electrodynamics (QED) its description in terms of perturbation theory restricts its range of applicability. In fact a mathematical consistent formulation is still unknown. We want to make a tiny step in that direction. Following ideas from Chaix and Iracane [CI89] we rigorously treat a model, which we derive from no-photon QED, describing ground states of atomic systems. Commonly speaking we deal with divergences arising from vacuum polarization (VP) effects.

By means of a Bogoliubov transform applied to a standard regularized Hamiltonian, describing interacting particle-antiparticle systems, we derive a well defined Hamiltonian. We impose an ultraviolet momentum cutoff Λ which is a necessity when dealing with VP-effects. Our only restriction on Λ is its finiteness. Although VP plays a minor role in the Lamb-shift calculations of low-Z hydrogen-like atoms, it is important for high-Z atoms [MPS98, Sha02] and even plays a major role for muonic atoms [PY59, GRS60] and for heavy ion collision [RMG81].

As our main issue we describe a stable polarized vacuum *self-consistently* in a *non-perturbative* way. Such a solution minimizes an appropriate energy functional and differs slightly from the Furry picture which can be obtained after the first iteration of a fixed point algorithm starting with the projector on the negative energy spectrum of the free Dirac operator. Additionally we introduce the notion of *supertraceclass* operators which turns out to be an appropriate tool to describe different *charge sectors*. We will discuss a simplified model in more detail, neglecting the exchange energy. For the corresponding

[1]The authors are thankful to Robert Seiringer and Vladimir M. Shabaev for valuable comments. They acknowledge support through the European Union's IHP network Analysis & Quantum HPRN-CT-2002-00277. E.S. acknowledges support from the Institut Universitaire de France

self-consistent solution of the VP-density we will perform a fully – to any order – consistent charge renormalization scheme.

As an application our procedure might be useful for the Lamb-shift calculation of atomic bound states: usually, see, e.g., [MPS98, Sha02], one works in the Furry picture, i.e. with the (generalized) eigenstates $|\psi_m\rangle$, solving $[D^0 - \alpha\varphi]|\psi_m\rangle = \varepsilon_m|\psi_m\rangle$, to evaluate corrections due to Vacuum Polarization, where D^0 denotes the free Dirac operator $D^0 = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta$. The "eigenstates" $|\psi_m\rangle$ satisfy the relation

$$\sum_{\varepsilon_m<0} |\psi_m\rangle\langle\psi_m| = P^\varphi L^2(\mathbb{R}^3)\otimes\mathbb{C}^4, \quad \sum_{\varepsilon_m\geq 0} |\psi_m\rangle\langle\psi_m| = [1-P^\varphi]L^2(\mathbb{R}^3)\otimes\mathbb{C}^4, \tag{II.1}$$

where the sum should rather be understood as an integral. $P^\varphi = \chi_{(-\infty,0)}(D^\varphi)$ is the projector on the negative spectral space of $D^\varphi = D^0 - \alpha\varphi$ therefore the right hand terms in (II.1) are mathematically well defined. A different frame concerning the choice of the electron-positron subspaces should rather be corresponding to an *effective* Hamiltonian $D^0 - \alpha\varphi + V_{\text{eff}}$, i.e. $[D^0 - \alpha\varphi + V_{\text{eff}}]|\psi_m\rangle = \varepsilon_m|\psi_m\rangle$ already including VP effects. In the following we describe a procedure to derive such an effective potential V_{eff} in a *self-consistent* way. It known in the literature, see, e.g., [BS72], that an effective potential can be derived by means of Feynman diagrams. However, we emphasize that our approach is mathematically rigorous, non-perturbative and works for any charge Z of the external potential.

Throughout the paper $\chi_{(-\infty,0)}(H)$ denotes the projector on the negative spectral space of H. In other words it consists of the eigenspace generated by the (generalized) eigenvectors of H with negative spectrum. (In the physical literature $\chi_{(-\infty,0)}(H)$ is often denoted as $\Lambda^-(H)$.)

We use relativistic units $\hbar = c = 1$, set the particle mass equal to one and $\alpha = e^2/(4\pi)$. We emphasize that in the first part e represents the *bare* charge of the fermion. We assume the presence of an external field $\varphi(x) = Zn * 1/|\cdot|$ describing one or more extended nuclei with overall charge density $Zn(x)$, $\int n = 1$.

In the following we will use the shorthands $D^0 = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta$, $D^\varphi = D^0 - \alpha\varphi$ and $\Psi(x)$ denotes the field consisting of a particle annihilation and an anti-particle creation operator, which will be made more explicit below. Our starting point is the formal Hamiltonian

$$\mathbb{H}_{\text{formal}} = \int dx\, \Psi^*(x) D^\varphi \Psi(x) + \frac{\alpha}{2}\int dx \int dy \frac{\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y)}{|\mathbf{x}-\mathbf{y}|}, \tag{II.2}$$

which acts on the Fock space. We use the notation $x = (\mathbf{x},\sigma) \in \mathbb{R}^3\times\{1,2,3,4\}$, σ denoting the spin index. $\int dx$ denotes integration over $\mathbb{R}^3$ and sum over the indices. The momentum cutoff Λ is implemented in the Hilbert space $\mathfrak{H}_\Lambda$ which builds up the corresponding Fock space. Namely

$$\mathfrak{H}_\Lambda = \{f \in L^2(\mathbb{R}^3)\otimes\mathbb{C}^4 | \operatorname{supp}\hat{f} \subset B(0,\Lambda)\},$$

that is the space $\mathfrak{H}_\Lambda$ of spin valued functions whose Fourier transform has support inside a ball with radius Λ. Equation (II.2) contains several divergences. We simply regularize by normal ordering with respect to the projector, P^0, on the negative spectral space of D^0, $P^0 = \chi_{(-\infty,0)}(D^0)$:

$$\mathbb{H} = \int dx\, :\Psi^*(x) D^\varphi \Psi(x):_{P^0} + \frac{\alpha}{2}\int dx\int dy \frac{:\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y):_{P^0}}{|\mathbf{x}-\mathbf{y}|}. \tag{II.3}$$

This corresponds to subtracting the energy of the free Dirac sea P^0 and the interaction energy with P^0. Indeed, the latter subtraction represents a generalized *mass* renormalization.

The field operator $\Psi(x)$ with respect to the free Dirac sea P^0 is given by $\Psi(x) = a_0(x) + b_0^*(x)$, satisfying the anti-commutation relation

$$\{a_0(x), a_0^*(y)\} = [1 - P^0](x, y), \ \{b_0^*(x), b_0(y)\} = P^0(x, y) \tag{II.4}$$

where $[1 - P^0](x, y)$, $P^0(x, y)$ are the integral kernels of the projectors $1 - P^0$ and P^0. All other possible pairs anticommute.

In the following we want to perform a general Bogoliubov transformation, see [CI89], and rewrite $\mathbb{H}$ with respect to any dressed vacuum (projector) P, $P - P^0 \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$. $\mathfrak{S}_2$ is the set of all operators whose square has a summable diagonal, i.e. $\mathfrak{S}_2(\mathfrak{H}_\Lambda) = \{A \in B(\mathfrak{H}_\Lambda) | \operatorname{tr} A^*A < \infty\}$ ($B(\mathfrak{H}_\Lambda)$ denotes the set of bounded operators). By definition of normal ordering

$$\begin{aligned} :\Psi^*(x)\Psi(y):_{P^0} &= a_0^*(x)a_0(y) + a_0^*(y)b_0(x) + b_0(x)a_0(y) - b_0^*(y)b_0(x) \\ &= \Psi^*(x)\Psi(y) - P^0(x, y). \end{aligned} \tag{II.5}$$

Representing now (II.5) by means of a general P, thereby changing from one representation to the other by an appropriate Bogoliubov transformation, one derives

$$:\Psi^*(x)\Psi(y):_{P^0} =: \Psi^*(x)\Psi(y):_P + Q(x, y), \tag{II.6}$$

with $Q = P - P^0$. This is a direct consequence of the fact that, with respect to P, we can rewrite the field Ψ as $\Psi(x) = a_P(x) + b_P^*(x)$, where the particle-antiparticle operators a_P, b_P fulfill the relations (II.4) with P^0 being replaced by P. After straightforward calculations one sees that

$$\begin{aligned} :\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y):_{P^0} &= :\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y):_P \\ &\quad + 2:\Psi^*(x)\Psi(x):_P \operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{y}, \mathbf{y})) - 2:\Psi^*(x)\Psi(y):_P Q(x, y) \\ &\quad + \operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{x}, \mathbf{x}))\operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{y}, \mathbf{y})) - |Q(x, y)|^2. \end{aligned}$$

Here and in the following $|Q|^2 = Q^*Q$. Therefore we can rewrite (II.3) with respect to an arbitrary dressed P as

$$\begin{aligned} \mathbb{H} = &\int :\Psi^*(x)D^\varphi\Psi(x):_P dx + \frac{\alpha}{2}\iint \frac{:\Psi^*(x)\Psi(x)\Psi^*(y)\Psi(y):_P}{|\mathbf{x} - \mathbf{y}|} dx\, dy \\ &+ \alpha \iint \frac{:\Psi^*(x)\Psi(x):_P \operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{y}, \mathbf{y}))}{|\mathbf{x} - \mathbf{y}|} d\mathbf{x}\, d\mathbf{y} - \alpha \iint \frac{:\Psi^*(x)\Psi(y):_P Q(x, y)}{|\mathbf{x} - \mathbf{y}|} dx\, dy \\ &+ \operatorname{tr}(D^0 Q) - \int d\mathbf{x}\varphi(\mathbf{x})\operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{x}, \mathbf{x})) + \frac{\alpha}{2}\iint \frac{\operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{x}, \mathbf{x}))\operatorname{Tr}_{\mathbb{C}^4}(Q(\mathbf{y}, \mathbf{y}))}{|\mathbf{x} - \mathbf{y}|} d\mathbf{x}\, d\mathbf{y} \\ &- \frac{\alpha}{2}\iint \frac{|Q(x, y)|^2}{|\mathbf{x} - \mathbf{y}|} dx\, dy. \end{aligned} \tag{II.7}$$

Note, due to the Shale-Stinespring Theorem [Tha92, Thm 10.6], the requirement $P - P^0 \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$ is necessary and sufficient to guarantee that the Bogoliubov transformation is implementable in the Fock space, which means that there exists a vacuum vector Ω_P in the Fock space, such that $a_P\Omega_P = b_P\Omega_P = 0$. We do not specify this implementation in

more detail, since we are primarily interested in the vacuum energy (see, e.g., [HLS04a, Appendix]). Evaluating the vacuum expectation value $\langle\Omega_P|\mathbb{H}|\Omega_P\rangle$ only the last two lines of (II.7) remain representing the energy of the dressed vacuum P measured with respect to P^0. In the Furry picture, P is fixed to be the projector $P^\varphi = \chi_{(-\infty,0)}(D^\varphi)$, and the energy of $P^\varphi - P^0$ is usually neglected. The second line of (II.7), in the Furry picture, represents the interaction with the VP-potentials, more precisely, the first term is the interaction with the VP-density $\rho^\varphi(\mathbf{x}) = \mathrm{Tr}_{\mathbb{C}^4}\big([P^\varphi - P^0](\mathbf{x},\mathbf{x})\big)$. This density diverges logarithmically in the cutoff Λ. This becomes finite by throwing away the physically unimportant divergent part of ρ^φ, a procedure usually referred to as charge renormalization. We refer to [HS03] for an intense study of the Furry picture. Let us remark that $P^\varphi - P^0 \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$ was shown by Klaus and Scharf [KS77].

In our case of a general P we do not *a priori* renormalize: this would destroy the naturalness of the Hamiltonian. Namely, such a renormalization procedure would depend on the representation P, therefore the Hamiltonian (II.7) would be no longer unitary equivalent with respect to different representations. Instead we give a fully consistent renormalization procedure later on for a slightly simplified model. That means we present a renormalization of the vacuum polarization density which applies simultaneously to any order of α and is in the spirit of the usual charge renormalization procedure [IZ80, Equ. (7-18)].

In the following we want to concentrate on the vacuum energy, which ceases to be a constant if one minimizes over P. Such a minimizer, if it exists, we call a stable vacuum.

2 Existence of a stable polarized vacuum

To define our energy functional properly we have to overcome the difficulty that the variables $Q = P - P^0$, D^0Q are not necessarily traceclass, this means $\mathrm{tr}\,|Q|$, $|Q| = \sqrt{Q^*Q}$, can be infinite. Indeed this is the case for $P^\varphi - P^0$. For any fixed Λ, $\rho^\varphi(x)$ is a well defined function (see [HLS04a, Equ. (6)]), which diverges for each x if Λ tends to ∞. Nevertheless even for fixed Λ the operator $P^\varphi - P^0$ is *not* traceclass.

For that reason we introduce the notion of supertraceclass operators (see, [Tha92, Section 5.7]). We say $A \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$ is supertraceclass with respect to P^0, i.e. A in $\mathfrak{S}_1^{P^0}(\mathfrak{H}_\Lambda)$, if the operators $A_{++} := (1-P^0)A(1-P^0)$ and $A_{--} := P^0AP^0$ are traceclass ($\in \mathfrak{S}_1(\mathfrak{H}_\Lambda)$), $\mathrm{tr}\,|A_{++}|, \mathrm{tr}\,|A_{--}| < \infty$, and we define

$$\mathrm{str}_{P^0} A = \mathrm{tr}\, A_{++} + \mathrm{tr}\, A_{--}. \tag{II.8}$$

Notice, if A is even traceclass then $\mathrm{str}_{P^0} A = \mathrm{tr}\, A$. In fact we show in [HLS04a] that any difference of two projectors $Q = P - P^0 \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$ (consequently also D^0Q, since due to the momentum cutoff Λ, D^0 is a bounded operator) is automatically in $\mathfrak{S}_1^{P^0}(\mathfrak{H}_\Lambda)$, moreover, $\mathrm{str}_{P^0} Q = \mathrm{str}_P Q$ is an integer which can be interpreted as the charge of Q (see below).

Calculating now the vacuum expectation value $\langle\Omega_P|\mathbb{H}|\Omega_P\rangle$ we see that the terms corresponding to the first two lines in (II.7) vanish and it rests the energy of the vacuum which we denote as $\mathcal{E}(Q) = \langle\Omega_P|\mathbb{H}|\Omega_P\rangle$. Defining the kinetic energy properly by means of the notion of supertraceclass we write $\mathcal{E}$ as

$$\mathcal{E}(Q) = \mathrm{str}_{P^0}(D^0Q) - \alpha\int \rho_Q\varphi + \frac{\alpha}{2}D(\rho_Q,\rho_Q) - \frac{\alpha}{2}\iint \frac{|Q(x,y)|^2}{|\mathbf{x}-\mathbf{y}|}dx\,dy, \tag{II.9}$$

$\rho_Q(\mathbf{x}) = \mathrm{Tr}_{\mathbb{C}^4}(Q(\mathbf{x},\mathbf{x}))$, $D(f,g) = \int \frac{f(\mathbf{x})g(\mathbf{y})}{|\mathbf{x}-\mathbf{y}|} d\mathbf{x}\, d\mathbf{y}$, acting on the extended, convex, set

$$\mathfrak{S}_\Lambda = \{Q \,|\, 0 \le Q + P^0 \le 1,\, Q \in \mathfrak{S}_1^{P^0}(\mathfrak{H}_\Lambda),\, \rho_Q \in \mathcal{C}\}, \tag{II.10}$$

with $\mathcal{C} = \{f \in \mathfrak{H}_\Lambda |\, D(f,f) < \infty\}$. We remark that due to our cutoff the density $\rho_Q(\mathbf{x})$ is well defined [HLS04a, Equ. (6)], and that for mathematical reasons we extended the physical states to their convex hull. As our main result we obtain that, for any Λ, $\mathcal{E}$ has a minimizer on $\mathcal{S}_\Lambda$, therefore there exists a stable vacuum.

Theorem II.1. *Let $0 \le \alpha \le 4/\pi$, $n \in \mathcal{C}$. Then there exists a minimizer Q of $\mathcal{E}$ in $\mathcal{S}_\Lambda$. Furthermore $P = Q + P^0$ is a projector satisfying the self-consistent equation*

$$P = \chi_{(-\infty,0)}\Big(D^0 - \alpha\varphi + \alpha\rho_Q * \frac{1}{|\cdot|} - \alpha\frac{Q(x,y)}{|\mathbf{x}-\mathbf{y}|}\Big). \tag{II.11}$$

The proof of the theorem is complicated and will be given elsewhere [HLS04b].

Equation (II.11) corresponds to Dirac's picture that the "correct" vacuum P should be the projector on the negative spectrum of an effective one-body Hamiltonian.

Remark, that in the case without external potential, $\varphi = 0$, the free projector P^0 solves (II.11) and minimizes $\mathcal{E}$ [CIL89, BBHS99].

Numerically the self-consistent solution of (II.11) could be evaluated by a fixed point algorithm, starting with P^0, which converges to the solution (II.11) as shown in [HLS04a]. Notice thereby $P = P^\varphi + \mathcal{O}(\alpha^2)$. In fact in [HLS04a] we prove (II.11) under reasonable restrictions, $\alpha\sqrt{\log\Lambda} \le 0.3$, using the Banach fixed point theorem. This proof is much more constructive than the direct proof in [HLS04b], and it might better reflect the smoothness properties corresponding to the solution of (II.11). In a different model [LS00] the Banach fixed point theorem was already used to derive a self-consistent projector.

Quite generally, as presented in [HLS04a] – following [CI89] – the Hamiltonian $\mathbb{H}$ can be applied on Hartree-Fock type states

$$|\psi\rangle = a_P^*(f_1)\dots a_P^*(f_l) b_P^*(g_1)\dots b_P^*(g_m)\Omega_P, \tag{II.12}$$

denoted as Bogoliubov-Dirac-Fock (BDF) states (see [CI89]), where $f_i \in [1-P]\mathfrak{H}_\Lambda$, for $1 \le i \le l$, $g_j \in P\mathfrak{H}_\Lambda$, for $1 \le j \le m$. Evaluating the expectation value $\langle\psi|\mathbb{H}|\psi\rangle$ one realizes that this only depends on the one particle density matrix γ_ψ associated to $|\psi\rangle$,

$$\gamma_\psi = \sum_{i=1}^{l} |f_i\rangle\langle f_i| - \sum_{j=1}^{m} |g_j\rangle\langle g_j|,$$

which fulfills $-P \le \gamma_\psi \le 1 - P$. More precisely

$$\langle\psi|\mathbb{H}|\psi\rangle = \mathcal{E}(Q + \gamma_\psi),$$

with $Q = P - P^0$. Consequently the real physical object is the sum of the vacuum density and the particle density, $Q + \gamma_\psi$, satisfying $-P^0 \le Q + \gamma_\psi \le 1 - P^0$ showing that $Q + \gamma_\psi$ is an element of $\mathcal{S}_\Lambda$. Quite generally, all admissible states, $Q + \gamma_\psi$, are in the set $\mathcal{S}_\Lambda$. This explains why we call the solution of (II.11) a *stable* vacuum: it is the state of lowest energy among all admissible states.

3 The charge of the vacuum

By means of the next theorem we emphasize the usefulness of the notion of *supertrace*. We have seen that the physical objects in $\mathcal{S}_\Lambda$ consists of elements $Q+\gamma$ where γ describes an appropriate particle-antiparticle state whose trace has an integer value and Q denote a vacuum state whose supertrace is an integer. Consequently the supertrace of the whole object is an integer which shows that the supertrace is an appropriate tool to describe charge sectors in $\mathcal{S}_\Lambda$.

Theorem II.2. *For any orthogonal projector P, such that $P-P^0 \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$, $\mathrm{str}_{P^0}[P-P^0]$ is an integer. If additionally P satisfies* (II.11) *and $0 \notin \sigma(D^\lambda)$, the spectrum of D^λ, for all $\lambda \in [0,\alpha]$, where $D^\lambda = D^0 - \lambda\varphi + \lambda\rho_Q * \frac{1}{|\cdot|} - \lambda\frac{Q(x,y)}{|\mathbf{x}-\mathbf{y}|}$, then $\mathrm{str}_{P^0}[P-P^0]=0$.*

In words, we assume in Theorem II.2 that no eigenvalue crosses 0 when deforming D^λ from 0 to α. We remark that under the assumptions made in [HLS04a, Theorem 4] we are sure that such crossings of eigenvalues do not happen.

Proof – Using the proof of [ASS94, Theorem 4.1] we obtain (see [HLS04a, Lemma 1]) for any orthogonal projector P, $P-P^0 \in \mathfrak{S}_2(\mathfrak{H}_\Lambda)$, that $\mathrm{str}_{P^0}[P-P^0] = \mathrm{tr}[P-P^0]^{2m+1}$, for all $m \geq 1$, which is automatically an integer according to Avron, Seiler, and Simon [ASS94]. Applying now [Hai03, Theorem 2] to the solution of (II.11) under the assumptions mentioned in the theorem we arrive at $\mathrm{str}_{P^0}[P-P^0]=0$. □

This Theorem tells us that if the overall charge of the nuclei is not too big the vacuum stays neutral, cf. [GMR85, Hai03].

Notice that $\mathrm{str}_{P^0} Q \neq \int \rho_Q$, since even for fixed Λ, Q is in general not traceclass.

4 Reduced vacuum energy functional

Recall up to now the charge was kept to be the bare one. In other words we did not renormalize the density ρ_Q. Next we want to derive a renormalization consistent to any order. This distinguishes us from usual renormalization procedures where only the first order (in α) of the VP density is renormalized. For that reason we simplify our vacuum energy by neglecting the exchange term,

$$\mathcal{E}_{\mathrm{red}}(Q) = \mathrm{str}_{P^0}(D^0 Q) - \alpha \int \rho_Q\varphi + \frac{\alpha}{2} D(\rho_Q,\rho_Q). \tag{II.13}$$

From a physical point of view this is quite natural, since the exchange term is usually treated together with a term describing the interaction with the photon field to form the standard electron *self-energy* that is a subject of the mass renormalization (see the remark after (II.3)).

Theorem II.3. *For any positive Λ there exists a minimizer $Q \in \mathcal{S}_\Lambda$ for $\mathcal{E}_{\mathrm{red}}$ satisfying*

$$Q = \chi_{(-\infty,0)}\Big(D^0 - \alpha\varphi + \alpha\rho_Q * \frac{1}{|\cdot|}\Big) - P^0. \tag{II.14}$$

The proof is on the one hand a direct consequence of the proof of Theorem II.1, on the other hand much simpler. Namely the fact that the functional is convex and strongly continuous in Q immediately implies that it is weakly lower semicontinuous on $\mathcal{S}_\Lambda$. Since it is trivially bounded from below one sees quite easily that the minimum is attained.

To see that the minimum satisfies the equation (II.14) one proceeds similar to [HLS04a, Theorem 3].

In order to perform our renormalization scheme we expand (II.14) in powers of α. Using the resolvent representation [Kat66, Section VI, Lemma 5.6] and recalling $\rho_Q(\mathbf{x}) = \mathrm{Tr}_{\mathbb{C}^4}(Q(\mathbf{x},\mathbf{x}))$ we derive from (II.14) the self-consistent equation for the VP-density

$$\rho_Q = -\frac{1}{2\pi}\int_{-\infty}^{\infty} d\eta \ \mathrm{tr}\left[\frac{1}{D^0 - \alpha\varphi + \alpha\rho_Q * \frac{1}{|\cdot|} + i\eta} - \frac{1}{D^0 + i\eta}\right], \tag{II.15}$$

which is guaranteed to exist as a well defined object in our set $\mathcal{C}$. Applying the resolvent equation

$$\frac{1}{A-\alpha B} - \frac{1}{A} = \alpha\frac{1}{A}B\frac{1}{A} + \alpha^2\frac{1}{A}B\frac{1}{A}B\frac{1}{A} + \alpha^3\frac{1}{A}B\frac{1}{A}B\frac{1}{A}B\frac{1}{A-\alpha B}$$

using Furry's Theorem [Fur37], telling us that the corresponding α^2-term with two potentials vanish, we obtain

$$\rho_Q = \alpha F_1(\rho_Q - Zn) + F_3(\alpha\rho_Q - \alpha Zn) \tag{II.16}$$

with

$$F_3(\alpha\rho) = \int_{-\infty}^{\infty} d\eta \,\mathrm{Tr}_{\mathbb{C}^4}\left[\frac{1}{D^0+i\eta}\left((\alpha\rho) * \frac{1}{|\cdot|}\right)\frac{1}{D^0+i\eta}\left((\alpha\rho) * \frac{1}{|\cdot|}\right)\times \right.$$
$$\left. \times\frac{1}{D^0+i\eta}\left((\alpha\rho) * \frac{1}{|\cdot|}\right)\frac{1}{D^0-\alpha\rho+i\eta}\right].$$

As realized first by Dirac [Dir34b, Dir34a] and Heisenberg [Hei34], cf. also [FO34], the term $F_1(\rho)$ plays a particular role since it is logarithmically ultraviolet divergent. Following, e.g., Pauli-Rose [PR36], one evaluates in Fourier representation

$$\widehat{F_1}[\rho](k) = -\hat{\rho}(k)B_\Lambda(k),$$

with

$$B_\Lambda(k) = -\frac{1}{\pi^2|k|^2}\int_{|l|\le\Lambda}\frac{(l+k/2)\cdot(l-k/2)+1-E(l+k/2)E(l-k/2)}{E(l+k/2)E(l-k/2)(E(l+k/2)+E(l-k/2))}dl, \tag{II.17}$$

$E(p) = \sqrt{1+p^2}$, which can be decomposed into [PR36, Equ. (5)-(9)] $C_\Lambda(k) = B_\Lambda - B_\Lambda(k)$, with

$$B_\Lambda = B_\Lambda(0) = \frac{1}{\pi}\int_0^{\frac{\Lambda}{E(\Lambda)}}\frac{z^2-z^4/3}{1-z^2}\,dz = \frac{2}{3\pi}\log(\Lambda) - \frac{5}{9\pi} + \frac{2}{3\pi}\log 2 + O(1/\Lambda^2). \tag{II.18}$$

Denote $\rho = \rho_Q - Zn$ the total density, then (II.16) reads in terms of ρ

$$\hat{\rho} + Z\hat{n} = -\alpha B_\Lambda\hat{\rho} - \alpha C_\Lambda(k)\hat{\rho} + \widehat{F}_3(\alpha\rho), \tag{II.19}$$

or equivalently

$$\alpha\hat{\rho} = -\frac{\alpha}{1+\alpha B_\Lambda}Z\hat{n} - \frac{\alpha}{1+\alpha B_\Lambda}C_\Lambda(k)\alpha\hat{\rho} + \frac{\alpha}{1+\alpha B_\Lambda}\widehat{F}_3(\alpha\rho). \tag{II.20}$$

To perform our renormalization scheme we fix as physical (renormalized) objects $\alpha_{\rm r}\rho_{\rm r} = \alpha\rho$, with $\alpha_{\rm r} = \frac{\alpha}{1+\alpha B_\Lambda}$, cf. [IZ80, Equ. (7-18)]. It holds

$$\lim_{\Lambda\to\infty} C_\Lambda(k) = C(k) = -\frac{1}{2\pi}\int_0^1 dx(1-x^2)\log[1+k^2(1-x^2)/4], \tag{II.21}$$

which was first calculated by Serber and Uehling [Ser35, Ueh35]. Therefore we can rewrite the self-consistent equation (II.19) as

$$\alpha_{\rm r}\hat\rho_{\rm r} = -\alpha_{\rm r}Z\hat n - \alpha_{\rm r}^2 C_\Lambda(k)\hat\rho_{\rm r} + \alpha_{\rm r}\widehat{F}_3(\alpha_{\rm r}\rho_{\rm r}), \tag{II.22}$$

independently of the bare α. This equation uniquely defines the VP density only depending on the physical observable $\alpha_{\rm r}$, which is what we understand under consistent to any order. The $\alpha_{\rm r}$ represents the dressed coupling constant, which is observed in experiment and whose value is approximately 1/137. Notice that from formula $\alpha_{\rm r} = \frac{\alpha}{1+\alpha B_\Lambda}$, it follows that necessarily $\alpha_{\rm r}B_\Lambda < 1$ and $\alpha_{\rm r}B_\Lambda \to 1$ as $\Lambda\to\infty$.

Returning to the *effective* Hamiltonian $D^0 - \alpha\varphi + \alpha\rho_Q$ and inserting (II.22), i.e. expressing in terms of the physical objects, we obtain

$$D^0 - \alpha_{\rm r}\rho_{\rm r} * \frac{1}{|\cdot|} = D^0 - \alpha_{\rm r}Zn * \frac{1}{|\cdot|} + V_{\rm eff}, \tag{II.23}$$

with

$$V_{\rm eff} = \frac{2}{\pi^3}\mathcal{F}^{-1}\left[\frac{\alpha_{\rm r}^2 C_\Lambda(k)\hat\rho_{\rm r}(k) + \alpha_{\rm r}\hat F_3(\alpha_{\rm r}\rho_{\rm r})}{k^2}\right](x)$$

the effective self-consistent potential, where $\mathcal{F}^{-1}$ denotes the inverse Fourier transform. Notice, this equation is valid for any strength of the external potential. However, expanding $\rho_{\rm r}$ in $\alpha_{\rm r}$, which only makes sense if $\alpha Z < 1$, one obtains to lowest order in $\alpha_{\rm r}$

$$\begin{aligned} V_{\rm eff} &= \alpha_{\rm r}^2 Z\frac{2}{\pi^3}\mathcal{F}^{-1}\left[\frac{C_\Lambda(k)\hat n(k)}{k^2}\right](x) \\ &\simeq \frac{\alpha_{\rm r}^2 Z}{3\pi}\int_1^\infty dt(t^2-1)^{1/2}\left[\frac{2}{t^2}+\frac{1}{t^4}\right]\int dx' e^{-2|x-x'|t}\frac{n(x')}{|x-x'|}, \end{aligned} \tag{II.24}$$

the Uehling potential [BR82]. Concerning a point like particle this potential was first written down in closed form by Schwinger [Sch49]. The next term in $V_{\rm eff}$ is of order $\mathcal{O}\left(\alpha_{\rm r}(\alpha_{\rm r}Z)^3\right)$. In principle all higher order corrections can be evaluated explicitly, which is not the task of our paper, and might play a role in Lamb-shift calculations of large-Z or muonic atoms. Finally we note that the convergence of the term in the right hand side of (II.19), in the case of the VP-density in the Furry picture, i.e. $\alpha^3\widehat{F}_3(\alpha n)$, was shown in various papers. The most clarifying proof with respect to spurious third order contributions can probably be found in [Sof88] (for earlier proofs, in particular corresponding to muonic atoms, we refer to the references in [Sof88]). However the fact that this term, $\alpha^3\widehat{F}_3(\alpha n)$, additionally gives rise to a well defined self-adjoint operator was recently proven in [HS03].

Bibliography

[ASS94] J. Avron, R. Seiler, and B. Simon. The index of a pair of projections. *J. Funct. Anal.*, 120:220–237, 1994.

[BBHS99] V. Bach, J.M. Barbaroux, B. Helffer, and H. Siedentop. On the stability of the relativistic electron-positron field. *Commun. Math. Phys.*, 201:445–460, 1999.

[BR82] E. Borie and G.A. Rinker. *Rev. Mod. Phys.*, 54:67, 1982.

[BS72] M.A. Braun and T.N. Sibirkina. *Teor. Mat. Fiz.*, 12(1):48, 1972. English translation in: Theor. Math. Phys. v. 12, N1.

[CI89] P. Chaix and D. Iracane. From quantum electrodynamics to mean field theory: I. the Bogoliubov-Dirac-Fock formalism. *J. Phys. B.*, 22(23):3791–3814, 1989.

[CIL89] P. Chaix, D. Iracane, and P.L. Lions. From quantum electrodynamics to mean field theory: II. Variational stability of the vacuum of quantum electrodynamics in the mean-field approximation. *J. Phys. B.*, 22(23):3815–3828, 1989.

[Dir34a] P.A.M. Dirac. Discussion of the infinite distribution of electrons in the theory of the positron. *Proc. Camb. Philos. Soc.*, 30:150–163, 1934.

[Dir34b] P.A.M. Dirac. Théorie du positron. *Solvay report*, XXV:203–212, 1934.

[FO34] W.H. Furry and J.R. Oppenheimer. On the theory of the electron and positive. *Phys. Rev., II. Ser.*, 45:245–262, 1934.

[Fur37] W.H. Furry. A symmetry theorem in the positron theory. *Phys. Rev.*, 51:125, 1937.

[GMR85] W. Greiner, B. Müller, and J. Rafelski. *Quantum Electrodynamics of Strong Fields*. Texts and Monographs in Physics. Springer-Verlag, first edition, 1985.

[GRS60] R. Glauber, W. Rarita, and P. Schwed. Vacuum polarization effects on energy levels in μ-mesonic atoms. *Phys. Rev.*, 120(2):609–613, 1960.

[Hai03] C. Hainzl. On the vacuum polarization density caused by an external field. 2003.

[Hei34] W. Heisenberg. Bemerkungen zur Diracschen Theorie des Positrons. *Z. Phys.*, 90:209–231, 1934.

[HLS04a] C. Hainzl, M. Lewin, and E. Séré. Existence of a stable polarized vacuum in the Bogoliubov-Dirac-Fock approximation. 2004.

[HLS04b] C. Hainzl, M. Lewin, and E. Séré. in preparation. 2004.

[HS03] C. Hainzl and H. Siedentop. Non-perturbative mass and charge renormalization in relativistic no-photon quantum electrodynamics. *Comm. Math. Phys.*, 243:241–260, 2003.

[IZ80] C. Itzykson and J.-B. Zuber. *Quantum Field Theory*. McGraw-Hill, New York, 1980.

[Kat66] T. Kato. *Perturbation Theory for Linear Operators*, volume 132 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, Berlin, 1966.

[KS77] M. Klaus and G. Scharf. The regular external field problem in quantum electrodynamics. *Helv. Phys. Acta*, 50:779–802, 1977.

[LS00] E.H. Lieb and H. Siedentop. Renormalization of the regularized relativistic electron-positron field. *Comm. Math. Phys.*, 213(3):673–683, 2000.

[MPS98] P. J. Mohr, G. Plunien, and G. Soff. QED corrections in heavy atoms. *Phys. Rep.*, 293:227–369, 1998.

[PR36] W. Pauli and M.E. Rose. Remarks on the polarization effects in the positron theory. *Phys. Rev II*, 49:462–465, 1936.

[PY59] A. Peterman and Y. Yamaguchi. Corrections to the 3d-2p transitions in μ-mesonic phosphorus and the mass of the muon. *Phys. Rev. Letters*, 2(8):359–361, 1959.

[RMG81] J. Reinhardt, B. Müller, and W. Greiner. Theory of positron production in heavy-ion collision. *Phys. Rev. A*, 24(1):103–128, 1981.

[Sch49] J. Schwinger. Quantum electrodynamics. II. Vacuum polarization and self-energy. *Phys. Rev., II. Ser.*, 75:651–679, 1949.

[Ser35] R. Serber. Linear modifications in the Maxwell field equations. *Phys. Rev., II. Ser.*, 48:49–54, 1935.

[Sha02] V.M. Shabaev. Two-time Green's function method in quantum electrodynamics of high-Z few-electron atoms. *Phys. Rep.*, 356:119–228, 2002.

[Sof88] G. Soff. *Phys. Rev. A*, 38:5066, 2088.

[Tha92] B. Thaller. *The Dirac Equation*. Springer Verlag, 1992.

[Ueh35] E.A. Uehling. Polarization effects in the positron theory. *Phys. Rev., II. Ser.*, 48:55–63, 1935.

[a] CEREMADE, CNRS UMR 7534,
Université Paris IX Dauphine,
Place du Maréchal de Lattre de Tassigny,
75775 Paris Cedex 16, France.
hainzl, lewin, sere@ceremade.dauphine.fr

[b] Laboratoire de Mathématiques,
Université Paris-Sud,
Bât. 425,
91 405 Orsay Cedex, France.

Bibliographie générale

[1] *Exponential decay of two-body eigenfunctions: a review*, Proceedings of the symposium on mathematical physics and quantum field theory (Berkeley, CA, 1999), vol. 4, Electron. J. Differ. Equ. Conf., 1999.

[2] S. Agmon, *Lectures on exponential decay of solutions of second-order elliptic equations*, Princeton University Press, 1982.

[3] A. Ambrosetti and P.H. Rabinowitz, *Dual variation methods in critical point theory and applications*, J. Functional Analysis **14** (1973), 349–381.

[4] T. Ando, *Properties of fermions density matrices*, Rev. Mod. Phys. **35** (1963), no. 3, 690–702.

[5] N. Ashcroft and N. Mermin, *Solid state physics*, Saunders, Philadelphia, 1976.

[6] P.W. Atkins and R.S. Friedman, *Molecular quantum mechanics. third edition*, Oxford University Press, 1997.

[7] J. Avron, R. Seiler, and B. Simon, *The index of a pair of projections*, J. Funct. Anal. **120** (1994), 220–237.

[8] V. Bach, J.M. Barbaroux, B. Helffer, and H. Siedentop, *On the stability of the relativistic electron-positron field*, Commun. Math. Phys. **201** (1999), 445–460.

[9] V. Bach, E.H. Lieb, and J.P. Solovej, *Generalized Hartree-Fock theory and the Hubbard model*, J. Statist. Phys. **76** (1994), no. 1-2, 3–89.

[10] A. Bahri, *Critical points at infinity in some variational problems*, Pitman Research Notes in Mathematics Series, vol. 182, Longman Scientific and Technical Ed., 1989.

[11] J.M. Barbaroux, W. Farkas, B. Helffer, and H. Siedentop, *On the Hartree-Fock equations of the electron-positron field*, preprint arXiv : math-ph/0404009, 2004.

[12] M. Sh. Birman, *The discrete spectrum in gaps of the perturbed periodic Schrödinger operator. I. Regular perturbations*, Boundary value problems, Schrödinger operators, deformation quantization, Math. Top., vol. 8, Akademie Verlag, Berlin, 1995, pp. 334–352.

[13] J.D. Bjorken and S.D. Drell, *Relativistic quantum mechanics*, McGraw-Hill, 1964.

[14] E. Borie and G.A. Rinker, Rev. Mod. Phys. **54** (1982), 67.

[15] M. Born and R. Oppenheimer, *Quantum theory of molecules*, Ann. Physics **84** (1927), 457–484.

[16] J. Borwein and D. Preiss, *A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions*, Trans. AMS **303** (1987), 517–527.

[17] M.A. Braun and T.N. Sibirkina, Teor. Mat. Fiz. **12** (1972), no. 1, 48, English translation in: Theor. Math. Phys. v. 12, N1.

[18] R. Brummelhuis, N. Röhrl, and H. Siedentop, *Stability of the relativistic electron-positron field of atoms in Hartree-Fock approximation: Heavy elements*, Doc. Math. **6** (2001), 1–9.

[19] E. Cancès, *Simulation moléculaire et effets d'environnement. Une perspective mathématique et numérique*, Ph.D. thesis, École Nationale des Ponts et Chausées, 1998.

[20] E. Cancès and C. Le Bris, *On the convergence of SCF algorithms for the Hartree-Fock equations*, M2AN Math. Model. Numer. Anal. **34** (2000), no. 4, 749–774.

[21] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris, and Y. Maday, *Computational quantum chemistry: a primer*, Handbook of Numerical Analysis, vol. X, pp. 3–270, Elsevier, 2003.

[22] I. Catto and P.-L. Lions, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part I.*, Comm. Part. Diff. Eq. **17** (1992), no. 7-8, 1051–1110.

[23] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part II.*, Comm. Part. Diff. Eq. **18** (1993), no. 1-2, 305–354.

[24] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories. Part III.*, Comm. Part. Diff. Eq. **18** (1993), no. 3-4, 381–429.

[25] ———, *Binding of atoms and stability of molecules in Hartree and Thomas-fermi type theories. Part IV.*, Comm. Part. Diff. Eq. **18** (1993), no. 7-8, 1149–1159.

[26] P. Chaix, *Une méthode de champ moyen relativiste et application à l'etude du vide de l'electrodynamique quantique*, Ph.D. thesis, 1990.

[27] P. Chaix and D. Iracane, *From quantum electrodynamics to mean field theory: I. The Bogoliubov-Dirac-Fock formalism.*, J. Phys. B. **22** (1989), no. 23, 3791–3814.

[28] P. Chaix, D. Iracane, and P.L. Lions, *From quantum electrodynamics to mean field theory: II. Variational stability of the vacuum of quantum electrodynamics in the mean-field approximation.*, J. Phys. B. **22** (1989), no. 23, 3815–3828.

[29] L.M. Cheung, S.T. Elbert, and K. Ruedenberg, *MCSCF optimization through combined use of natural orbital and the Brillouin-Levy-Berthier theorem*, Int. J. Quantum Chem. **16** (1979), 1069–1101.

[30] S.S-L. Chiu, J.W. McDouall, and I.H. Hiller, *Prediction of whole reaction paths for large molecular systems*, J. Chem. Soc. Faraday Trans. **90** (1994), no. 12, 1575–1579.

[31] Y.S. Choi and P.J. McKenna, *A mountain pass method for the numerical solution of semilinear elliptic problems*, Nonlinear Anal. **20** (1993), no. 4, 417–437.

[32] Y.S. Choi, P.J. McKenna, and M. Romano, *A mountain pass method for the numerical solution of semilinear wave equations*, Numer. Math. **64** (1993), no. 4, 487–509.

[33] A.J. Coleman, *Structure of fermions density matrices*, Rev. Mod. Phys. **35** (1963), no. 3, 668–689.

[34] A.J. Coleman and V.I. Yukalov, *Reduced density matrices: Coulson's challenge*, Springer Verlag, 2000.

[35] J.-M. Combes and R. Seiler, *Spectral properties of atomic and molecular systems*, Quantum dynamics of molecules (Proc. NATO Adv. Study Inst., Univ. Cambridge, Cambridge, 1979), NATO Adv. Study Inst. Ser., Ser. B: Physics, vol. 57, Plenum, New York, 1980, pp. 435–482.

[36] J.M. Combes, *On the Born-Oppenheimer approximation*, International Symposium on Mathematical Problems in Theoretical Physics (Kyoto Univ., Kyoto, 1975), Springer, Berlin, 1975, pp. 467–471. Lecture Notes in Phys., 39.

[37] O. Costin and A. Soffer, *Resonance theory for Schrödinger operators*, Commun. Math. Phys. **224** (2001), 133–152.

[38] M. Dimassi, *Développements asymptotiques pour des perturbations fortes de l'opérateur de Schrödinger périodique*, Ann. Inst. H. Poincaré Phys. Théor. **61** (1994), no. 2, 189–204.

[39] P.A.M. Dirac, *Theory of electrons and positrons*, Nobel lecture delivered at Stockholm, 1933.

[40] ———, *Discussion of the infinite distribution of electrons in the theory of the positron*, Proc. Camb. Philos. Soc. **30** (1934), 150–163.

[41] ———, *Théorie du positron*, Solvay report **XXV** (1934), 203–212.

[42] R.M. Dreizler and E.K.U. Gross, *Density functional theory*, Springer Verlag, 1990.

[43] F.J. Dyson, *The radiation theories of Tomonaga, Schwinger, and Feynman*, Phys. Rev. **75** (1949), no. 2, 486–502.

[44] ———, *The S matrix in Quantum Electrodynamics*, Phys. Rev. **75** (1949), no. 11, 1736–1755.

[45] R.H.A. Eade and M.A. Robb, *Direct minimization in MCSCF theory. The quasi-Newton method.*, Chem. Phys. Letters **83** (1981), no. 2, 362–368.

[46] R. Elber and M. Karplus, *A method for determining reaction paths in large molecules: application to myoglobin*, Chem. Phys. Letters **139** (1987), no. 5, 375–380.

[47] M.J. Esteban and E. Séré, *Solutions of the Dirac-Fock equations for atoms and molecules*, Commun. Math. Phys. **203** (1999), 499–530.

[48] ———, *Nonrelativistic limit of the Dirac-Fock equations*, Ann. Henri Poincaré **2** (2001), no. 5, 941–961.

[49] ———, *An overview on linear and nonlinear Dirac equations*, Disc. and Cont. Dynam. Systems **8** (2002), no. 2, 381–397.

[50] ———, *Dirac-Fock models for atoms and molecules and related topics*, Proc. ICMP, 2003.

[51] G. Fang and N. Ghoussoub, *Morse-type information on Palais-Smale sequences obtained by min-max principles*, Manuscripta Math. **75** (1992), 81–95.

[52] V. Fock, *Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems*, Z. Phys. **61** (1930), 126–148.

[53] J.B. Foresman and A. Frisch, *Exploring chemistry with electronic structure methods*, Gaussian Inc, 1996, Chapter 3 and references therein.

[54] S. Fournais, M. and T. Hoffmann-Ostenhof, and T. Østergaard Sørensen, *Analyticity of the density of electronic wavefunctions*, Preprint, 2002.

[55] ———, *The electron density is smooth away from the nuclei*, Commun. Math. Phys. **228** (2002), 401–415.

[56] ———, *On the regularity of the density of electronic wavefunctions*, Mathematical results in quantum mechanics (Taxco, 2001), Contemp. Math., vol. 307, 2002, pp. 143–148.

[57] J.D. French and V.F. Weisskopf, *The electromagnetic shift of energy levels*, Phys. Rev., II. Ser. **75** (1949), 1240–1248.

[58] G. Friesecke, *The multiconfiguration equations for atoms and molecules: charge quantization and existence of solutions*, Arch. Rat. Mech. Analysis **169** (2003), 35–71.

[59] ———, *On the infinitude of nonzero eigenvalues of the single-electron density matrix for atoms and molecules*, Proc. Roy. Soc. London Ser. A **459** (2003), 47–52.

[60] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, B.B. Stefanov, G. liu, A. Liashenko, P. Piskorz, I. Kpmaromi, G. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P.M.W. Gill, B.G. Johnson, W. Chen, M.W. Wong, J.L. Andres, M. Head-Gordon, E.S. Replogle, and J.A. Pople, *Gaussian 98 (Revision A.7)*, Gaussian Inc., Pittsburgh PA, 1998.

[61] W.H. Furry, *A symmetry theorem in the positron theory*, Phys. Rev. **51** (1937), 125.

[62] W.H. Furry and J.R. Oppenheimer, *On the theory of the electron and positive*, Phys. Rev., II. Ser. **45** (1934), 245–262.

[63] N. Ghoussoub, *Duality and perturbation methods in critical point theory*, Cambridge University Press, 1993.

[64] R. Glauber, W. Rarita, and P. Schwed, *Vacuum polarization effects on energy levels in μ-mesonic atoms*, Phys. Rev. **120** (1960), no. 2, 609–613.

[65] C. Gomez, C. Bunks, J.P. Chancelier, F. Delebecque, M. Goursat, R. Nikoukhah, and S. Steer, *Engineering and scientific computing with scilab*, Birkauser, 1999.

[66] W. Greiner, B. Müller, and J. Rafelski, *Quantum electrodynamics of strong fields*, first ed., Texts and Monographs in Physics, Springer-Verlag, 1985.

[67] C. Hainzl, *On the vacuum polarization density caused by an external field*, 2003.

[68] C. Hainzl, M. Lewin, and E. Séré, *Existence of a stable polarized vacuum in the Bogoliubov-Dirac-Fock approximation*, Preprint ArXiv: `math-ph/0403005`, 2004.

[69] ———, *Self-consistent solution for the polarized vacuum in a no-photon QED model*, Preprint ArXiv: `physics/0404047`, 2004.

[70] C. Hainzl and H. Siedentop, *Non-perturbative mass and charge renormalization in relativistic no-photon quantum electrodynamics*, Comm. Math. Phys. **243** (2003), 241–260.

[71] D. Hartree, *The wave-mechanics of an atom with a non-coulomb central field. Part I. Theory and methods.*, Proc. Camb. Phil. Soc. **24** (1928), 89–312.

[72] M. Head-Gordon, *Quantum chemistry and molecular processes*, J. Phys. Chem. **100** (1996), 13213–13225.

[73] W.J. Hehre, L. Radom, P.v.R. Schleyer, and J.A. Pople, *Ab initio molecular orbital theory*, Wiley, 1986.

[74] W. Heisenberg, *Bemerkungen zur Diracschen Theorie des Positrons*, Z. Phys. **90** (1934), 209–231.

[75] B. Helffer and H. Siedentop, *Form perturbation of the second quantized Dirac field*, Mathematical Physics Electronic Journal **4** (1998).

[76] G. Henkelman, G. Jóhannesson, and H. Jónsson, *Methods for finding saddle points and minimum energy paths*, Ed. S. D. Schwartz (Kluwer Academic Publishers), 2000.

[77] G. Henkelman and H. Jónsson, *A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives*, J. Chem. Physics **111** (1999), no. 15, 7010–7022.

[78] ———, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points*, J. Chem. Physics **113** (2000), no. 22, 9978–9985.

[79] G. Henkelman, H. Jónsson, and B.P. Uberuaga, *A climbing image nudged elastic band method for finding saddle points and minimum energy paths*, J. Chem. Physics **113** (2000), no. 22, 9901–9904.

[80] P.D. Hislop and I.M Sigal, *Introduction to spectral theory with application to Schrödinger operators*, Springer (New York), 1996.

[81] M. and T. Hoffmann-Ostenhof, *Schrödinger inequalities and asymptotic behavior of the electron density of atoms and molecules*, Phys. Rev. A **16** (1977), no. 5, 1782–1785.

[82] M. and T. Hoffmann-Ostenhof and T. Østergaard Sørensen, *Electron wavefunctions and densities for atoms*, Ann. Henri Poincaré **2** (2001), no. 1, 77–100.

[83] M. and T. Hoffmann-Ostenhof and H. Stremnitzer, *Local properties of Coulombic wave functions*, Commun. Math. Phys. **163** (1994), 185–215.

[84] D. Hundertmark, N. Röhrl, and H. Siedentop, *The sharp bound on the stability of the relativistic electron-positron field in Hartree-Fock approximation*, Comm. Math. Phys. **211** (2000), no. 3, 629–642.

[85] W. Hunziker, *On the spectra of Schrödinger multiparticle Hamiltonians*, Helv. Phys. Acta **39** (1966), 451–462.

[86] W. Hunziker and I.M. Sigal, *The quantum N-body problem*, J. Math. Phys **41** (2000), no. 6, 3348–3509.

[87] P. Indelicato, *Projection operators in Multiconfiguration Dirac-Fock calculations. Application to the ground state of heliumlike ions*, Phys. Rev. A **51** (1995), no. 2, 1132–1145.

[88] C. Itzykson and J.-B. Zuber, *Quantum field theory*, McGraw-Hill, New York, 1980.

[89] P.G. Jasien and R. Shepard, *A general polyatomic potential energy surface fitting method*, Int. J. Quantum Chemistry **22** (1988), 183–198.

[90] K. Kato, *New idea for proof of analyticity of solutions to analytic non linear equations*, SUT Journal of Math. **32** (1996), no. 2, 157–161.

[91] T. Kato, *Fundamental properties of hamiltonian operators of Schrödinger type*, Trans. Am. Math. Soc **70** (1951), 195–221.

[92] ———, *On the eigenfunctions of many particle systems in quantum mechanics*, Commun. Pure and Appl. Math. **10** (1957), 151–171.

[93] ———, *Perturbation theory for linear operators*, second ed., Springer, 1995.

[94] M. Klaus and G. Scharf, *The regular external field problem in quantum electrodynamics*, Helv. Phys. Acta **50** (1977), 779–802.

[95] M. Klein, A. Martinez, R. Seiler, and X. P. Wang, *On the Born-Oppenheimer expansion for polyatomic molecules*, Comm. Math. Phys. **143** (1992), no. 3, 607–639.

[96] F. Klopp, *Impuretés dans une structure périodique*, Journées "Équations aux Dérivées Partielles" (Saint Jean de Monts, 1990), École Polytech., Palaiseau, 1990, pp. Exp. No. XX, 12.

[97] ———, *Étude semi-classique d'une perturbation d'un opérateur de Schrödinger périodique*, Ann. Inst. H. Poincaré Phys. Théor. **55** (1991), no. 1, 459–509.

[98] W. Kohn and L.J. Sham, *Self-consistent equations including exchange and correlation effects*, Phys. Rev. **140** (1965), A1133–A1138.

[99] L.D. Landau and E.H. Lifchitz, *Quantum mechanics*, Pergamon Press, 1977.

[100] C. Le Bris, *Quelques problèmes mathématiques en chimie quantique moléculaire*, Ph.D. thesis, École Polytechnique, 1993.

[101] ———, *A general approach for multiconfiguration methods in quantum molecular chemistry*, Ann. Inst. H. Poincaré Anal. Non linéaire **11** (1994), no. 6, 441–484.

[102] ———, *Mathématiques appliquées et chimie quantique*, Matapli - Bulletin de la Société de Mathématiques Appliquées et Industrielles **58** (1999), 25–36.

[103] J.F. Léon, *Excited states for Coulomb systems in the Hartree-Fock approximation*, Comm. Math. Phys. **120** (1988), no. 2, 261–268.

[104] M. Lewin, *The multiconfiguration methods in quantum chemistry: Palais-Smale condition and existence of minimizers*, C. R. Acad. Sci. Paris, Ser. I **334** (2002), 299–304.

[105] ———, *A mountain pass for reacting molecules*, Ann. Henri Poincaré (2004).

[106] ———, *Solutions of the multiconfiguration equations in quantum chemistry*, Arch. Rat. Mech. Anal. **171** (2004), no. 1, 83–114.

[107] A.S. Lewis and M.L. Overton, *Eigenvalue optimization*, Acta Num. **5** (1996), 149–190.

[108] E.H. Lieb, *Bound on the maximum negative ionization of atoms and molecules*, Phys. Rev. A. **29** (1984), no. 6, 3018–3028.

[109] ———, *The stability of matter : from atoms to stars*, Bull. Amer. Math. Soc. **22** (1990), 1–49.

[110] E.H. Lieb and H. Siedentop, *Renormalization of the regularized relativistic electron-positron field*, Comm. Math. Phys. **213** (2000), no. 3, 673–683.

[111] E.H. Lieb, H. Siedentop, and J.P. Solovej, *Stability and instability of relativistic electrons in classical electromagnetic fields*, J. Stat. Phys. **89** (1997), no. 1–2, 37–59.

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

[113] ———, *The Thomas-Fermi theory of atoms, molecules and solids*, Adv. Math. **23** (1977), 22–116.

[114] E.H. Lieb and W.E. Thirring, *Universal nature of Van Der Waals forces for Coulomb systems*, Phys. Rev. A **34** (1986), 40–46.

[115] P.L. Lions, *The concentration-compactness principle in the calculus of variations. The locally compact case, part I*, Ann. Inst. Henri Poincaré **1** (1984), no. 2, 109–149.

[116] ———, *Solutions of Hartree-Fock equations for Coulomb systems*, Commun. Math. Phys. **109** (1987), 33–87.

[117] 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** (1955), no. 6, 1474–1489.

[118] ———, *Quantum theory of many-particle systems. II. Study of the ordinary Hartree-Fock approximation*, Phys. Rev. **97** (1955), no. 6, 1490–1508.

[119] ———, *Quantum theory of many-particle systems. III. Extension of the Hartree-Fock scheme to include degenerate systems and correlation effects*, Phys. Rev. **97** (1955), no. 6, 1509–1520.

[120] ———, *Correlation problem in many-electron quantum mechanics. I. Review of different approaches and discussion of some current ideas*, Adv. Chem. Phys. **2** (1959), 207–322.

[121] M.A. Robb M. Frisch, I.N. Ragazos and H.B. Schlegel, *An evaluation of three direct mcscf procedures*, Chem. Phys. Letters **189** (1992), no. 6, 524–528.

[122] M. Mekli and I.M. Sigal, *A time-dependent theory of quantum resonances*, Commun. Math. Phys. **201** (1999), 549–576.

[123] P.W. Milonni, *The quantum vacuum: An introduction to quantum electrodynamics*, Academic Press, Inc., Boston, 1994.

[124] M.H. Mittleman, *Theory of relativistic effects on atoms: Configuration-space hamiltonian*, Phys. Rev. A **24** (1981), no. 3, 1167–1175.

[125] P. J. Mohr, G. Plunien, and G. Soff, *QED corrections in heavy atoms*, Phys. Rep. **293** (1998), 227–369.

[126] C. B. Morrey Jr, *On the analyticity of the solutions of analytic non-linear elliptic systems of partial differential equations*, Amer. J. Math. **80** (1958), 198–218.

[127] A. Outassourt, *Comportement semi-classique pour l'opérateur de Schrödinger à potentiel périodique*, J. Funct. Anal. **72** (1987), no. 1, 65–93.

[128] M.L. Overton, *On minimizing the maximum eigenvalue of a symmetric matrix*, SIAM J. Matrix Anal. Appl. **9** (1988), 256–268.

[129] ———, *Large scale optimization of eigenvalues*, SIAM J. Optim. **2** (1992), 88–120.

[130] E. Paturel, *Solutions of the Dirac-Fock equations without projector*, Ann. Henri Poincaré **1** (2000), no. 6, 1123–1157.

[131] W. Pauli and M.E. Rose, *Remarks on the polarization effects in the positron theory*, Phys. Rev II **49** (1936), 462–465.

[132] A. Peterman and Y. Yamaguchi, *Corrections to the 3d-2p transitions in μ-mesonic phosphorus and the mass of the muon*, Phys. Rev. Letters **2** (1959), no. 8, 359–361.

[133] W. Quapp and D. Heidrich, *Analysis of the concept of minimum energy path on the potential energy surface of chemically reacting systems*, Theoret. Chim. Acta **66** (1984), 245–260.

[134] W. Quapp, H. Hirsch, O. Imig, and D. Heidrich, *Searching for saddle points of potential energy surfaces by following a reduced gradient*, J. Comput. Chem. **19** (1998), no. 9, 1087–1100.

[135] P.H. Rabinowitz, *Minimax methods in critical point theory with applications to differential equations*, Number 65 in Conference Board of the Mathematical Sciences, Regional Conference series in Mathematics, American Mathematical Society, 1986.

[136] M. Reed and B. Simon, *Methods of modern mathematical physics*, vol. I – Functionnal Analysis, Academic Press, 1972.

[137] ———, *Methods of modern mathematical physics*, vol. IV – Analysis of Operators, Academic Press, 1978.

[138] J. Reinhardt, B. Müller, and W. Greiner, *Theory of positron production in heavy-ion collision*, Phys. Rev. A **24** (1981), no. 1, 103–128.

[139] B. O. Roos, *The Complete Active Space Self-Consistent Field method and its applications in electronic structure calculation*, Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys. **69** (1987), 399–446.

[140] H. B. Schlegel, *Ab inition methods in quantum chemistry*, vol. 1, pp. 249–286, Lawley Ed, 1987, and references cited therein.

[141] H.B. Schlegel, Adv. Chem. Phys. **67** (1987), 249.

[142] E. Schrödinger, *Quantisierung als eigenwertproblem (vierte mitteilung)*, Ann. Phys. (Leipzig) **81** (1926), 220–250.

[143] J. Schwinger, *Quantum electrodynamics. II. Vacuum polarization and self-energy*, Phys. Rev., II. Ser. **75** (1949), 651–679.

[144] R. Serber, *Linear modifications in the Maxwell field equations*, Phys. Rev., II. Ser. **48** (1935), 49–54.

[145] V.M. Shabaev, *Two-time Green's function method in quantum electrodynamics of high-Z few-electron atoms*, Phys. Rep. **356** (2002), 119–228.

[146] R. Shepard, *The multiconfiguration self-consistent field method*, Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys. **69** (1987), 63–200.

[147] I.M. Sigal, *Geometric methods in the quantum many-body problem. Non existence of very negative ions*, Commun. Math. Phys. **85** (1982), 309–324.

[148] B. Simon, *Geometric methods in multiparticle quantum systems*, Commun. Math. Phys. **55** (1977), 259–274.

[149] ———, *Trace ideals and their applications*, London Mathematical Society Lecture Notes Series, vol. 35, Cambridge University Press, 1979.

[150] J.C. Slater, *A note on Hartree's method*, Phys. Rev. **35** (1930), 210–211.

[151] G. Soff, Phys. Rev. A **38** (2088), 5066.

[152] A. Soffer and M.I. Weinstein, *Time dependent resonance theory*, GAFA, Geom. Funct. Anal. **8** (1998), 1086–1128.

[153] P. Strange, *Relativistic quantum mechanics*, Cambridge University Press, 1998.

[154] J. Sucher, *Foundations of the relativistic theory of many-particle atoms*, Phys. Rev. A **22** (1980), no. 2, 348–362.

[155] ———, *Relativistic many-electrons Hamiltonians*, Phys. Scrypta **36** (1987), 271–281.

[156] E. Teller, Rev. Mod. Phys. **34** (1962), 267.

[157] B. Thaller, *The Dirac equation*, Springer Verlag, 1992.

[158] W. Thirring, *Quantum mathematical physics*, vol. Atoms, Molecules and Large Systems, Springer, Second Edition 2002.

[159] E.A. Uehling, *Polarization effects in the positron theory*, Phys. Rev., II. Ser. **48** (1935), 55–63.

[160] C. Van Winter, *Theory of finite systems of particles. I. The Green function*, Mat.-Fys. Skr. Danske Vid. Selsk. **2** (1964), no. 8.

[161] M.J. Vilkas, Y. Ishikawa, and K. Koc, *Quadratically convergent multiconfiguration Dirac-Fock and multireference relativistic configuration-interaction calculations for many-electrons systems*, Phys. Rev. E **58** (1998), no. 4, 5096–5110.

[162] J. von Neumann, *Mathematishe grundlagen der quantenmechanik*, Springer Verlag (Berlin), 1932.

[163] E. Weinan, R. Weiqing, and E. Vanden-Eijnden, *String method for the study of rare events*, Phys. Rev. B **66** (2002), 052301.

[164] S. Weinberg, *The quantum theory of fields*, vol. I, Cambridge University Press, 1996.

[165] V. Weisskopf, *Über die Elektrodynamik des Vakuums auf Grund der Quantentheorie des Elektrons*, Math.-Fys. Medd., Danske Vid. Selsk. **16** (1936), no. 6, 1–39.

[166] H.-J. Werner, *Matrix-formulated direct Multiconfiguration Self-Consistent Field and Multiconfiguration Reference Configuration-Interaction methods*, Ab Initio Methods in Quantum Chemistry - II in Adv. Chem. Phys. **69** (1987), 1–62.

[167] H.-J. Werner and W. Meyer, *A quadratically convergent multiconfiguration-self-consistent field method with simultaneous optimization of orbitals and CI coefficients*, J. Chem. Phys. **73** (1980), no. 5, 2342–2356.

[168] J.H. Wolkowisky, *Existence of solutions of the Hartree equations for n electrons. An application of the Schauder-Tychonoff theorem*, Indiana Univ. Math. J. **22** (1972-73), 551–568.

[169] E. Zeidler, *Applied functional analysis, applications to mathematical physics*, Springer (New York), 1995.

[170] G. M. Zhislin, *Discussion of the spectrum of Schrödinger operators for systems of many particles. (in Russian)*, Trudy Moskovskogo matematiceskogo obscestva **9** (1960), 81–120.

[171] G.M. Zhislin, *On the finiteness of the discrete spectrum of the energy operator of negative atomic and molecular ions*, Teoret. Mat. Fiz. **21** (1971), 332–341.