HAL
open science

Loi de van der Waals-London pour les systèmes d'atomes et de molécules relativistes

Michael Hartig

▶ **To cite this version:**

Michael Hartig. Loi de van der Waals-London pour les systèmes d'atomes et de molécules relativistes. Physique mathématique [math-ph]. Université de Toulon, 2019. Français. NNT : 2019TOUL0009 . tel-02494777

HAL Id: tel-02494777

https://theses.hal.science/tel-02494777

Submitted on 29 Feb 2020

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.

ÉCOLE DOCTORALE 548 – Mer et Sciences

Centre de Physique Théorique, Université de Toulon

Loi de van der Waals-London pour les systèmes d'atomes et de molécules relativistes

présentée par :

Michael HARTIG

soutenue le **14 juin 2019**

pour obtenir le grade de Docteur en

Mathématiques

Loi de van der Waals-London pour les systèmes d'atomes et de molécules relativistes

THÈSE dirigée par :

M. BARBAROUX Jean-Marie - Professeur des Universités, Université de Toulon

JURY :

M. FAUPIN Jérémy - Professeur des Universités, Université de Lorraine – Rapporteur

M. ANAPOLITANOS Ioannis - Professeur, Karlsruher Institut für Technologie – Rapporteur

M. BRUNEAU Vincent – Professeur des Universités, Université de Bordeaux – Examinateur

Mme. GAITAN Patricia – Professeur des Universités, Aix Marseille Université– Examinatrice

M. HUNDERTMARK Dirk – Professeur, Karlsruher Institut für Technologie – Examinateur

Mme. PANATI Annalisa – Maître de conférences, Université de Toulon- Examinatrice

VAN DER WAALS-LONDON INTERACTION OF ATOMS WITH PSEUDO-RELATIVISTIC KINETIC ENERGY

MICHAEL C. HARTIG

Abstract. We consider a multiatomic system where the nuclei are assumed to be point charges at fixed positions. Particles interact via Coulomb potential and electrons have pseudo-relativistic kinetic energy. We prove the van der Waals-London law, which states that the interaction energy between neutral atoms decays as the sixth power of the distance $|D|$ between the atoms. We rigorously compute all the terms in the binding energy up to the order $|D|^{-9}$ with error term of order $\mathcal{O}(|D|^{-10})$. As intermediate steps we prove exponential decay of eigenfunctions of multiparticle Schrödinger operators with permutation symmetry imposed by the Pauli principle, new estimates of the localization error.

Contents

1. Structure of the thesis 2
2. Mécanique quantique 3
2.1. Observables 3
2.2. Évolution temporelle 4
2.3. Spectre d'un opérateur 4
2.4. États liés et états de diffusion 5
2.5. L'approximation du noyau fixe 6
2.6. Caractère auto-adjoint du Hamiltonien 6
2.7. Stabilité de la matière 7
3. Systèmes multi-particules 8
3.1. Caractérisation du spectre essentiel 10
3.2. Partition de l'espace d'état, ou localisation 12
3.3. Les valeurs propres discrètes 13
3.4. Énergie au-dessous du spectre essentiel 15
3.5. Comportement asymptotique pour les états liés 16
3.6. L'opérateur d'énergie relativiste 16
3.7. Considérations de symétries 18
4. La force de dispersion de Van der Waals-London 19
5. Modèle, notations et résultat principal 21
5.1. Molécules diatomiques 22
6. Quantum mechanics 25
6.1. Observables 25
6.2. Time evolution 26
6.3. Spectrum of an operator 26
6.4. Bound States and Decaying States 27
6.5. Fixed nucleus approximation 28
6.6. Self-adjointness of the Hamiltonian 28
6.7. Stability of matter 29
7. Multiparticle systems 30
7.1. Characterization of the essential spectrum 32
7.2. Partition of the state space, or localization 34

7.3. The discrete eigenvalues 35
7.4. Energy below the essential spectrum 36
7.5. Asymptotic behaviour of bound states 38
7.6. Relativistic kinetic energy operator 38
7.7. Symmetry considerations 40
8. The Van der Waals - London dispersion 41
9. The model, notations and main results 42
9.1. Diatomic molecules 44
9.2. Extension to M-atomic molecules 48
10. Exponential decay of eigenfunctions 50
11. Localization error estimate 53
11.1. Multiparticle localization error estimate 55
12. Proof of Theorem 9.2 56
12.1. Estimate from below 56
12.2. Estimate from above 60
13. Proof of Theorem 9.4 64
Appendix A. The HVZ theorem 66
Appendix B. Existence of a ground state for atoms and positive ions 70
Appendix C. Proof of Lemma 10.5 72
Appendix D. Intercluster interaction in diatomic molecules 78
Appendix E. Orthogonality relations 83
Appendix F. Remark on actions of the permutation group 86
Appendix G. Spin, Symmetries and Representation Theory 86
G.1. Quantum numbers 86
G.2. Symmetries of the physical system 87
G.3. Representation theory 89
G.4. Projection onto irreducible representations 93
G.5. Young diagrams 94
Appendix H. Prospect: generalization and application of the method 97
Appendix I. The Brown-Ravenhall operator 98
I.1. Introduction to the model 98
I.2. Localization error estimate modified 100
I.3. Estimate from below 105
I.4. Estimate from above 108
References 110

1. Structure of the thesis

This thesis starts with a french introduction, that will be repeated in english. We give a general and non-rigorous introduction to quantum mechanics and multiparticle spectral analysis in french in Chapter 2 and 3. We proceed to motivate and present the main result in Chapter 4 and 5.

The corresponding english version of this introduction is given in Chapter 6 and 7. In Chapter 8 we present the physical problem that is discussed in the core part of this thesis together with a brief overview on the literature of mathematical physics on the topic. The results presented in Chapter 9 to Appendix F are based upon a collaboration with Jean-Marie Barbaroux (Aix-Marseille Univ et Université de Toulon), Dirk Hundertmark (KIT, Karlsruhe), and Semjon Vugalter (KIT, Karlsruhe). In Appendix G we discuss some representation theoretical results that are related to the Pauli explusion priciple and questions of symmetry. We discuss in Appendix H possible generalizations to the method that was derived in Chapter 9ff

and application thereof to other models. We elaborate an application of the method to the Brown-Ravenhall operator in Appendix I and give an outline of the proof adapted to this model. The results obtained in Appendix I are in large based on the works by Morozov [30, 31] and Morozov Vugalter [32].

2. Mécanique quantique

Dans cette courte introduction à la mécanique quantique nous suivons essentiellement le livre de Gustafson et Sigal [17] sur les éléments fondamentaux de la mécanique quantique en nous attardant plus particulièrement sur le cas de systèmes à une particule.

Nous donnerons un aperçu des résultats les plus importants pour le cas de systèmes à plusieurs particules qui permettra de bâtir un cadre adapté à la compréhension des résultats démontrés dans cette thèse. Compte tenu de la complexité du sujet qu'est la mécanique quantique, nous ne ferons qu'écorcher le sujet en nous concentrant sur les aspects étudiés ici. Pour une introduction plus détaillée, le lecteur pourra consulter par exemple [17].

En mécanique quantique, comme le montre l'expérience des fentes de Young, l'électron et les autres particules élémentaires ont un comportement d'ondes et ne possèdent pas de position déterministe, contrairement au cas de la mécanique classique dans laquelle un objet possède, au moins à une certaine échelle, une position et une impulsion exacte. Dans le cadre classique on peut exprimer des quantités mesurables, comme par exemple l'énergie cinétique, l'énergie potentielle ou des quantités similaires, comme des fonctions qui dépendent de propriétés déterministes telle la position la masse ou l'impulsion.

En mécanique quantique, la position d'une particule est donnée en terme de distribution de probabilité.

L'*état* d'une particule est décrit par une fonction de la position et du temps, à valeur complexe, $\psi(x,t)$ avec $x \in \mathbb{R}^3$ et $t \in \mathbb{R}$. On normalise cette fonction comme suit

$$\int_{\mathbb{R}^3} |\psi(x,t)|^2 dx = 1 \tag{2.1}$$

et on interprète $|\psi(\cdot,t)|^2$ comme la densité de probabilité de la position de la particule au temps t. On notera que dans certains cas il suffit d'avoir $\psi(\cdot,t) \in L^2(\mathbb{R}^3)$, qui implique que la fonction est normalisable au sens de (2.1). Par ailleurs, pour justifier le nom de *fonction d'onde*, cette fonction devra être solution d'une équation des ondes. Dans notre cas, il s'agit de l'équation de Schrödinger suivante:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{2.2}$$

avec condition initiale

$$\psi\big|_{t=0} = \psi_0 \tag{2.3}$$

où

$$H = T + V \tag{2.4}$$

est le *Hamiltonien de Schrödinger* ou *opérateur d'énergie* agissant sur l'espace de Hilbert $\mathcal{H}$. Le premier terme dans l'équation (2.4) est l'opérateur d'énergie cinétique et le terme V, opérateur de multiplication par la fonction $V(x)$, est le potentiel.

2.1. Observables.

Les propriétés mesurables expérimentalement, appelées *observables*, doivent être évaluées de manière non déterministe. Cela signifie que chaque valeur peut être observée avec une certaine probabilité et jamais avec une certitude absolue. Cela implique que ces mesures classiques sont exprimées à l'aide de variables aléatoires comme en théorie des probabilités. Dans l'approche de Schrödinger, les observables sont des opérateurs qui agissent sur les fonctions dans $L^2(\mathbb{R}^3)$. De

plus, on requiert le caractère *auto-adjoint* de ces opérateurs; avec cette propriété, on s'assure entre autres que les observables ont un spectre de valeurs réelles.

2.2. **Évolution temporelle.** Dans cette section, on aborde la notion de caractère auto-adjoint, en particulier pour l'opérateur énergie du système qui correspond à l'observable donnée par l'opérateur de Schrödinger H. Si on ne pouvait pas assurer le caractère auto-adjoint de H, l'interprétation physique des résultats serait, au mieux, discutable. On aborde ici les implications du caractère auto-adjoint. On verra plus tard dans le Chapter 2.6 pourquoi les opérateurs de la forme (2.4) que nous étudierons sont auto-adjoints. Sans entrer dans les détails, et de façon mathématiquement non rigoureuse, nous allons motiver ici certains des résultats fondamentaux de la mécanique quantique. La solution canonique d'une équation différentielle du type

$$\frac{d}{dt}f(t) = ikf(t) \tag{2.5}$$

pour un certain $k \in \mathbb{R}$, est $f(t) = e^{ikt}f(0)$. Nous aimerions utiliser une approche similaire si on remplace k par un opérateur H dans l'équation différentielle (2.2). Pour un opérateur borné B sur l'espace de Hilbert $\mathcal{H}$, on peut définir l'exponentielle e^{itB} à l'aide de la série convergente

$$e^{itB} := \sum_{n=0}^{\infty} \frac{(it)^n B^n}{n!}. \tag{2.6}$$

Pour un opérateur *non borné* A, comme c'est le cas pour notre Hamiltonien H, on ne peut pas utiliser une telle série. On peut cependant définir l'opérateur e^{itA} par le calcul fonctionnel à l'aide du théorème spectral, voir par exemple [40, Chapter VII], pour lequel le caractère auto-adjoint de A est requis.

Un second théorème clef en mécanique quantique qui repose sur le caractère auto-adjoint de H est le théorème de Stone ([40, Theorem VIII.8]). Ce théorème montre une correspondance bi-univoque entre opérateurs auto-adjoints H et un *groupe unitaire à un paramètre* $U(t)$. Pour tout temps $t \in \mathbb{R}$, l'opérateur $U(t)$ est unitaire, fortement continu, et pour un état $\psi_0 \in L^2(\mathbb{R}^3)$ on peut dériver $U(t)\psi_0$ par rapport à t, avec $\partial_t U(t)\psi = iH\psi$. Ceci justifie l'interprétation physique d'évolution temporelle avec toutes les propriétés attendues d'un opérateur de la forme e^{itH}. Par abus de notation, on écrira $U(t) = e^{itH}$ et pour un état $\psi_0 \in L^2(\mathbb{R}^3)$ on appellera $\psi(x,t) = U(t)\psi_0(x)$ l'évolution au temps t de l'état ψ_0.

2.3. **Spectre d'un opérateur.** Le spectre d'un opérateur est défini par

$$\sigma(A) = \{\lambda \in \mathbb{C} \big| A - \lambda \text{ n'est pas inversible (n'a pas d'inverse borné)}\}. \tag{2.7}$$

On notera que pour un opérateur auto-adjoint A, le spectre est réel ($\sigma(A) \subset \mathbb{R}$). On dira que λ est une *valeur propre isolée* de A si λ est une valeur propre de A pour laquelle il existe $\epsilon > 0$ tel que $(\lambda - \epsilon, \lambda + \epsilon) \cap \sigma(A) = \{\lambda\}$. Typiquement, dans l'analyse spectrale des Hamiltoniens des systèmes atomiques ou moléculaires, on s'intéresse à deux types de valeurs dans le spectre. On définit d'abord le *spectre discret* comme

$$\sigma_{\text{disc}}(A) := \big\{\lambda \in \sigma(A) \big| \lambda \text{ is an isolated eigenvalue with finite multiplicity}\}. \tag{2.8}$$

Et on définit le *spectre essentiel* comme

$$\sigma_{\text{ess}}(A) := \sigma(A) \setminus \sigma_{\text{disc}}(A). \tag{2.9}$$

Il existe d'autre façons de décomposer le spectre. La décomposition en spectre discret et spectre essentiel est utilisée dans le cas des Hamiltoniens atomiques et moléculaires pour la stabilité du spectre essentiel pour une certaine classe de perturbations.

On va se concentrer dans cette partie introductive sur le cas d'opérateurs de Schrödinger H avec opérateur énergie cinétique $-\Delta$ et potentiel de Coulomb V.

Les valeurs propres $E \in \sigma_{\text{disc}}(H)$ correspondent aux niveaux d'énergie qui sont réalisés par une fonction $\psi \in \mathcal{D}(H)$, où $\mathcal{D}(H)$ est le domaine de l'opérateur H. Ces niveaux d'énergie sont visualisés par le spectre de raies d'absorption d'un atome. Ces raies d'absorption correspondent aux différences d'énergies entre l'énergie de l'état fondamental et celles des états excités d'un atome.

2.4. **États liés et états de diffusion.** Dans ce paragraphe, nous motivons la classification du spectre en spectre discret et spectre essentiel. Les définitions suivantes sont extraites de [17, §6.2], assorties de quelques remarques supplémentaires.

Regardons les solutions de l'équation de Schrödinger

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{2.10}$$

de condition initiale

$$\psi\big|_{t=0} = \psi_0. \tag{2.11}$$

Supposons que $\psi_0 \in \{\text{espace engendré par les fonctions propres de } H\}$. Alors une solution de (2.10) est donnée par $\psi(x,t) = U(t)\psi_0(x)$, pour le groupe unitaire à un paramètre $U(t)$ décrit dans le paragraphe 2.2, formellement identifié à l'exponentielle $e^{-\frac{i}{\hbar}tH}$. Alors par [39, Theorem XI.115] pour tout $\epsilon > 0$ il existe un R tel que

$$\inf_t \int_{|x|<R} |\psi(x,t)|^2 dx > 1 - \epsilon. \tag{2.12}$$

Un état ψ qui vérifie (2.12) est appelé un état lié, puisqu'il reste essentiellement localisé dans une même région de l'espace pour tout temps.

On note que pour une fontion propre ϕ_0 de H correspondant à l'énergie E on a

$$\phi(x,t) = e^{-\frac{i}{\hbar}tH}\phi_0(x) = e^{-\frac{i}{\hbar}tE}\phi_0(x) \tag{2.13}$$

et donc la distribution de probabilité $|\phi(\cdot,t)|^2$ est constante dans le temps ce qui implique aussi

$$\int_{|x|\geq R} |\phi(x,t)|^2 dx = \int_{|x|\geq R} |\phi_0(x)|^2 dx \to 0 \tag{2.14}$$

quand $R \to \infty$.

D'autre part, si $\psi_0 \in \{\text{espace engendré par les fonctions propres de } H\}^\perp$ l'évolution temporelle $\psi = U(t)\psi_0$ sortira de toute boule de rayon R. Plus précisément

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \int_{|x|\leq R} |\psi(x,t)|^2 dx dt = 0. \tag{2.15}$$

De tels états sont appelés *états de diffusion* puisqu'ils décrivent une particule qui se déplace à l'infini.

La discussion ci-dessus, basée sur les théorèmes de Wiener et RAGE (voir [39, p.340ff]) est associée à la décomposition du spectre en spectre ponctuel et spectre continu. Pour les modèles d'atomes et de molécules, lorsqu'on est interessé par l'état fondamental ou par les états excités, comme c'est le cas dans la présente thèse, la décomposition naturelle est celle en spectre discret et essentiel, comme discuté dans le paragraphe 2.3. Dans ce cas, toute fonction propre associé à une valeur dans le spectre discret est dans le spectre ponctuel et donc vérifie (2.12) et est un état lié. En fait tout état d'énergie dans le spectre discret satisfait la propriété plus forte (2.13).

Remark 2.1. *Pour un Hamiltonien quelconque il est possible d'observer des valeurs propres plongées dans le spectre essentiel. Ce n'est pas l'objet de cette thèse d'étudier ce type de valeurs propres et d'en donner une interprétation physique.*

2.5. **L'approximation du noyau fixe.** La masse d'un noyau comparée à la masse d'un électron est très importante. Dans de nombreux cas, il est alors possible de simplifier l'étude d'un système en supposant les noyaux fixes, i.e. la position des noyaux est traitée comme un paramètre, puisque l'énergie cinétique des noyaux est négligeable. La discussion qui suit est une discussion qui n'est pas mathématiquement rigoureuse sur les motivations de considérer les noyaux comme fixes.

Dans un système où deux corps sont en rotation l'un autour de l'autre, comme c'est le cas par exemple dans l'atome d'hydrogène classique, l'impulsion des deux corps doit avoir la même valeur. Classiquement, l'énergie cinétique d'une particule est donné par $E_{\text{kin}} = \frac{p^2}{m}$, où m est la masse de la particule et p est son impulsion. Dans le cas d'un atome d'hydrogène, les impulsions de l'électron et du noyau (constitué d'un proton) ont la même valeur. Donc, la relation entre les énergies cinétiques est

$$E_{\text{kin}}^{(pr)} m_{pr} = E_{\text{kin}}^{(e)} m_e \tag{2.16}$$

où les indices pr et e sont utilisés pour désigner respectivement le proton et l'électron. Comme $m_{pr} \approx 2{\cdot}10^3 m_e$, on a les relations inverses pour les énergies cinétiques. Pour des atomes plus importants, cette différence entre les énergies cinétiques du noyau et de l'électron sont encore plus marquées. Ceci pour deux raisons: tout d'abord, le noyau d'un atome plus gros contient aussi des neutrons qui rendent le rapport des masses encore plus important; en fait, le tableau des éléments montre qu'approximativement, les 20 premiers éléments contiennent à peu près autant de neutrons que de protons, alors que pour les éléments suivants il y a plus de neutrons que de protons. La seconde raison est que l'impulsion d'un noyau dans un atome croit au plus linéairement avec le nombre d'électrons car la croissance linéaire correspondrait au cas où tous les électrons seraient groupés en un même point. En général, l'impulsion du noyau croit donc moins vite que linéairement car les électrons ont tendance à s'étaler, et les effets que les électrons ont sur le noyau peuvent donc s'annihiler partiellement entre eux.

Puisque l'on a considéré le noyau comme fixe, la masse du noyau est donc essentiellement sans effet sur le système. Par un changement d'échelle du système on peut s'abstraire du préfacteur sur l'énergie cinétique de l'électron. Soit $-e$ la charge d'un électron. Le Hamitonien d'un système à un électron de position $x \in \mathbb{R}^3$ interagissant via le potentiel de Coulomb avec un noyau de charge eZ dont la position est fixée à l'origine est donné par

$$H_1^Z := -\Delta + \frac{e^2 Z}{|x|}. \tag{2.17}$$

Pour $Z = 1$, l'opérateur dans (2.17) décrit l'atome d'hydrogène. Pour les autres valeurs entières de Z, un tel système est dit *hydrogénoïde* ou bien *atome hydrogénoïde*. L'indice sur H dans (2.17) indique le nombre d'électrons du système, qui est dans cette discussion fixé pour l'instant à un seul.

L'approximation de noyau fixe deviendra encore plus pertinente pour des grands systèmes, et tout spécialement dans le cas de multiples noyaux, qui seront discutés plus tard dans la Section 3.

2.6. **Caractère auto-adjoint du Hamiltonien.** Pour avoir un système en mécanique quantique qui est bien défini, on doit s'assurer que l'opérateur H est auto-adjoint. Nous n'entrerons pas dans le détail de la preuve du caractère auto-adjoint pour les Hamiltoniens considérés, car les preuves sont un peu techniques et bien connues des spécialistes. Nous rappelons cependant le fameux théorème de Kato-Rellich, dont la formulation choisie est celle dans [37, Theorem X.12] auquel le lecteur peut se référer pour une preuve. Le théorème de Kato-Rellich permet d'obtenir le caractère auto-adjoint de Hamiltoniens H dans de nombreux cas.

Definition 2.2. *Soient A et B deux opérateurs de domaine dense sur un espace de Hilbert $\mathcal{H}$. Supposons que*

- $\mathcal{D}(B) \supset \mathcal{D}(A)$
- *Pour deux valeurs réelles a and b et pour tout $\varphi \in \mathcal{D}(A)$,*

$$\|B\varphi\| \leq a\|A\varphi\| + b\|\varphi\|. \tag{2.18}$$

Alors B est dit A-borné. L'infimum des valeurs a est appelé borne relative de B par rapport à A.

Theorem 2.3 (Kato-Rellich theorem)**.** *Supposons que A est auto-adjoint et B est symétrique, et supposons de plus que B est A-borné, de borne relative $a < 1$. Alors $A + B$ est auto-adjoint sur $\mathcal{D}(A)$ et essentiellement auto-adjoint sur tout coeur de A.*

Ce théorème garantit que si on ajoute à un opérateur auto-adjoint A une perturbation B "pas trop grande" par rapport à A, alors $A+B$ reste auto-adjoint. Pour le cas de notre opérateur H_1^Z à une particule on a que l'opérateur $-\Delta$ sur $L^2(\mathbb{R}^3)$ est auto-adjoint, et le caractère auto-adjoint de H_1^Z est obtenu par le corollaire suivant du théorème de Kato-Rellich:

Corollary 2.4. *Soit $V \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ un potentiel réel. Alors $-\Delta + V$ est auto-adjoint.*

Ce corollaire garantit le caractère auto-adjoint pour une grande variété de potentiels physiques intéressants. Nous n'aborderons pas l'interprétation physique concernant le caractère auto-adjoint. Nous allons plutôt discuter d'une question très concrète liée à ce caractère auto-adjoint: la question de la stabilité de la matière. Plus particulièrement: "Qu'est-ce qui empêche un électron de s'effondrer sur le noyau"?

2.7. **Stabilité de la matière.** Nous avons discuté ci-dessus deux scénarios possibles pour l'évolution temporelle d'un état quantique pour un système à un électron: le cas d'un état lié où l'électron reste pour tout temps essentiellement localisé dans une région compacte, et les états diffusifs dans lequel un électron s'échappe à l'infini. Nous n'avons pas abordé le cas d'un électron qui s'approcherait infiniment près du noyau. Un tel scenario est appelé *Instabilité de premier type.*

Definition 2.5. *Un système décrit par le Hamiltonien quantique H_1^Z agissant sur $L^2(\mathbb{R}^3)$ a la propriété de stabilité de premier ordre si*

$$\inf \sigma(H_1^Z) > -\infty. \tag{2.19}$$

Alors que l'opérateur énergie cinétique est un opérateur positif, le potentiel de Coulomb peu a priori conduire à des valeurs négatives, aussi grandes que l'on veut en valeur absolue pour certains états. Une violation de (2.19) semble donc théoriquement possible pour un état dans lequel l'électron serait localisé arbitrairement près du noyau avec probabilité non nulle.

Cependant, l'opérateur énergie cinétique contrôle la singularité du potentiel de Coulomb. Dans [25, Part III], la preuve de la stabilité de premier type pour un opérateur H_1^Z tel que défini par (2.17) utilise une inégalité de Sobolev

$$\|\nabla\psi\|^2 \geq 3\left(\frac{\pi}{2}\right)^{\frac{3}{4}} \|\varrho(\psi)\|_3 \tag{2.20}$$

où $\varrho(\psi) = |\psi|^2$ est la densité de probabilité de présence de l'électron et $\|\cdot\|_3$ désigne la norme dans $L^3(\mathbb{R}^3)$. De plus on a

$$\langle\psi, V\psi\rangle = -e^2 Z \int_{\mathbb{R}^3} \frac{\varrho(\psi)(x)}{|x|} dx. \tag{2.21}$$

Ces deux inégalités impliquent

$$\langle \psi, H_1^Z \psi \rangle \geq \inf_{\substack{\varrho \geq 0 \\ \|\varrho\|_1 = 1}} \left\{ 3 \left(\frac{\pi}{2} \right)^{\frac{3}{4}} \|\varrho\|_3 - e^2 Z \int_{\mathbb{R}^3} \frac{\varrho(x)}{|x|} dx \right\} \tag{2.22}$$

dans laquelle on autorise toute fonction positive ρ telle que $\|\varrho\|_1 = 1$. La minimisation du membre de droite de (2.22) implique

$$\langle \psi, H_1^Z \psi \rangle \geq -\frac{e^4 Z^2}{3}. \tag{2.23}$$

Malheureusement, cette preuve n'apporte pas un éclairage complètement convaincant du fait que l'électron ne s'effondre pas sur le noyau. En s'inspirant de l'explication heuristique faite dans [25], examinons la relation entre l'énergie cinétique et l'énergie de Coulomb pour une famille de fonctions gaussiennes

$$\varphi_k(x) := C_k e^{-\frac{|x|^2}{k^2}}. \tag{2.24}$$

La fonction φ_k est principalement localisée dans une bande de largeur k autour de l'origine. Pour l'énergie cinétique, on a

$$\langle \varphi_k, -\Delta \varphi_k \rangle \sim \frac{1}{k^2}$$

et pour l'énergie de Coulomb

$$\langle \varphi_k, V \varphi_k \rangle \sim -\frac{e^2 Z}{k}.$$

Ainsi, pour que l'électron puisse se rapprocher de l'origine et donne une énergie de Coulomb de l'ordre de $-\frac{e^2 Z}{k}$, le prix à payer en énergie cinétique est de l'ordre de $\frac{1}{k^2}$. Ce dernier terme est clairement dominant par rapport à la partie Coulombienne pour les petites valeurs de k. Ceci explique heuristiquement pour un Hamiltonien de type H_1^Z que l'électron ne s'effondre pas sur le noyau.

3. Systèmes multi-particules

Jusqu'ici, nous n'avons observé que des fonctions d'onde décrivant le comportement d'une seule particule. Nous abordons maintenant le cas de systèmes à plusieurs particules. Dans cette thèse, nous modéliserons les noyaux comme des particules de position fixe, négligeant leur énergie cinétique dans le calcul de l'énergie totale du système. Par contre, nous ne négligerons pas les interactions électrostatiques entre elles, et avec les autres particules. Nous allons introduire un système permettant de décrire les atomes et les molécules. Soient M noyaux positionnés respectivement en $X_1, \cdots, X_M \in \mathbb{R}^3$ et de charge $eZ_1, \cdots eZ_M$, qui interagissent via des potentiels de Coulomb avec N électrons de positions respectives $x_1, \cdots, x_N \in \mathbb{R}^3$ ayant pour masse m_e et charge $-e$. On note par la lettre calligraphique $\mathcal{Z} = (Z_1, \cdots, Z_M)$ le vecteur décrivant les nombres atomiques des M noyaux. L'opérateur de Schrödinger pour un tel système, sur $L^2(\mathbb{R}^{3N})$, est alors donné par

$$H_N^{\mathcal{Z}} = \sum_{j=1}^{N} \frac{p_j^2}{2} - \sum_{j=1}^{N} \sum_{m=1}^{M} \frac{e^2 Z_m}{|x_j - X_m|} + \sum_{1 \leq i < j \leq N} \frac{e^2}{|x_i - x_j|} \tag{3.1}$$

où $p_j = -i\hbar \nabla_{x_j}$ est l'impulsion du j-ème électron. On distingue bien ici les notations H_N^Z and $H_N^{\mathcal{Z}}$. Dans le premier cas, on fait référence à un système de N électrons qui interagissent avec un seul noyau, alors que dans le second, il s'agit d'un système de N électrons qui interagissent avec M noyaux.

Pour une fonction $\psi \in L^2(\mathbb{R}^{3N})$ la densité de probabilité de présence des N électrons respectivement aux points $x_1, \cdots, x_N$ est donnée par $|\psi|^2$.

La première question naturelle pour un tel opérateur est de savoir s'il est auto-adjoint. Dans le cas d'un système à un électron, le théorème de Kato-Rellich et ses applications ont suffit à conclure que pour des potentiels V tels que $V = L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ (ce qui inclut les potentiels de Coulomb), l'opérateur $-\Delta + V$ est auto-adjoint. Une généralisation canonique pour des potentiels pour N électrons $V = L^2(\mathbb{R}^{3N}) + L^\infty(\mathbb{R}^{3N})$ n'est possible que si on néglige les interactions électron-électron. Pour un opérateur du type (3.1) dans lequel les interactions électron-électron sont prises en compte, on utilise le théorème KLMN, voir [37, Theorem X.17]. Heuristiquement le caractère relativement borné du potentiel requis dans le théorème de Kato-Rellich peut être remplacé par la condition plus faible de borne relative au sens des formes, ce qui permet d'étendre l'ensemble des potentiels que l'on peut considérer à $V = R + L^\infty(\mathbb{R}^3)$, où R est l'ensemble des *potentiels de Rollnik*, défini dans [37, p. 170] de la façon suivante:

Definition 3.1. *Une fonction mesurable V sur $\mathbb{R}^3$ est appelée potentiel de Rollnik si*

$$\|V\|_R^2 := \int_{\mathbb{R}^3} \frac{|V(x)||V(y)|}{|x-y|^2} dxdy < \infty. \tag{3.2}$$

On notera que les interactions électron-électron sont des potentiels de Rollnik. Le théorème KLMN s'applique donc et implique que l'opérateur défini par (3.1) est essentiellement auto-adjoint, i.e. avec une extension auto-adjointe unique.

On peut aussi étendre la propriété de stabilité de premier type pour un système à N électrons en utilisant la propriété de borne relative au sens des formes pour le potentiel de Coulomb à plusieurs particules.

Dans le cas général, la construction d'une solution d'une équation de Schrödinger pour plusieurs particules n'a jamais été obtenue. Cependant l'analyse du spectre de l'opérateur énergie donne des informations très intéressantes sur le système. Certaines de ces propriétés ont été discutées dans les Chapitres 2.3 et 2.4. Le spectre essentiel, et plus particulièrement le bas du spectre essentiel peut par exemple nous donner une information sur la stabilité du système. Comme nous l'avons vu dans (2.13), un état orthogonal à l'ensemble des états propres correspond à un état diffusif, ce qui signifie qu'au moins un électron du système s'échappe à l'infini. On notera qu'il peut exister des fonctions propres avec des valeurs propres plongées dans le spectre essentiel. Savoir si de tels états correspondent à des états stables ou instables est difficile. En effet, l'évolution temporelle d'un tel état est stationnaire. Mais un gain d'énergie du système, même infime, peut, selon le système étudié, conduire cet état à diffuser. De tels cas ne seront pas rencontrés dans les Hamiltoniens étudiés dans cette thèse.

Dans les systèmes atomiques et moléculaires que l'on considère, les résultats permettant de caractériser le spectre essentiel et le spectre discret sont le théorème HVZ, le théorème d'existence d'un état fondamental et les théorèmes sur le nombre de valeurs propres sous le bas du spectre essentiel, incluant le théorème de Zhislin et ses généralisations. Avant de donner les énoncés rigoureux de ces résultats, nous essayons d'abord de les expliquer simplement et d'en donner les implications.

Le théorème HVZ a deux implications. Premièrement l'infimum du spectre essentiel $\inf \sigma_{\text{ess}}(H)$ d'un système avec un noyau de charge eZ, qui interagit avec $N \leq Z$ électrons est égal à l'énérgie la plus basse du système dans lequel $N-1$ électrons sont proches du noyau et un électron bouge librement sans influence des autres particules. En plus le théorème implique que pour un tel système, le spectre essentiel est de la forme $\sigma_{\text{ess}}(H) = [\inf \sigma_{\text{ess}}(H), +\infty)$. En particulier, le spectre essentiel ne contient pas de lacunes.

La première partie de l'énoncé justifie le terme *seuil d'ionisation* pour l'infimum du spectre essentiel. Un système physique tend vers un état où l'énergie totale

est minimale. Une énergie au-dessus de ce seuil d'ionisation provoquera pour le système l'envoi d'un des électrons vers l'infini pour réduire l'énergie du système et obtenir une énergie totale de $\Sigma = \inf \sigma_{\text{ess}}(H)$.

Le deuxième théorème important est l'existence d'un état fondamental. Cela implique l'existence d'une énergie minimale pour un système et permet la construction d'un état lié. Une autre classe de résultats traite du nombre de fonctions propres isolées au-dessous du spectre essentiel, donc le nombre des états excités qui sont stables.

3.1. **Caractérisation du spectre essentiel.** On a déjà discuté l'importance du théorème HVZ et ses implications pour la structure des atomes et des molécules d'une manière non-rigoureuse. Pour énoncer le résultat précisément on doit introduire le concept de *décomposition en clusters* et une notion connexe, la *localisation*. Par souci d'une notation plus facile on discutera le cas d'un seul noyau, une généralisation pour M noyaux étant assez simple.

L'idée d'une décomposition en clusters est de partitionner l'ensemble $\{1, \cdots, N\}$ d'indices des électrons. On décompose l'ensemble en sous-ensembles $\mathcal{C}_0, \cdots, \mathcal{C}_{k-1}$ disjoints tels que

$$\bigcup_{l=0}^{k-1} \mathcal{C}_l = \{1, \cdots, N\}. \tag{3.3}$$

Un sous-ensemble $\mathcal{C}_l$ s'appelle un *cluster* et la collection $(\mathcal{C}_0, \cdots, \mathcal{C}_{k-1})$ s'appelle une *décomposition en k clusters*. Soit $\mathcal{D}_N^k$ l'ensemble de toutes les décompositions possibles de $\{1, \cdots, N\}$ en k clusters. À chaque cluster $\mathcal{C}_l$, on allue und région telle que les particules dans des clusters différents sont loin l'une de l'autre. Dans le cas d'un noyau de charge eZ positionné à l'origine, le système est décrit par le Hamiltonien

$$H_N^Z = \sum_{j=1}^{N} \left(\frac{p_j^2}{2} - \frac{e^2 Z}{|x_j|} \right) + \sum_{1 \le i < j \le N} \frac{e^2}{|x_i - x_j|}. \tag{3.4}$$

On considère les décompositions en 2 clusters $\mathcal{D}_N^2$, dans lesquelles les électrons dans le cluster $\mathcal{C}_0$ ne sont pas associés à un noyau. Pour une décomposition $\beta \in \mathcal{D}_N^2$ on définit *l'interaction intercluster*

$$I_\beta := \sum_{i \in \mathcal{C}_0} -\frac{e^2 Z}{|x_i|} + \sum_{\substack{i \in \mathcal{C}_0 \\ j \in \mathcal{C}_1}} \frac{e^2}{|x_i - x_j|} \tag{3.5}$$

et le *Hamiltonien du cluster β*

$$H_\beta := H_N^Z - I_\beta. \tag{3.6}$$

Theorem 3.2 (Théorème HVZ). *Soit $\beta \in \mathcal{D}_N^2$ une décomposition en clusters. On définit $\mu_\beta := \inf \sigma(H_\beta)$ et $\Sigma := \min_{\substack{\beta \in \mathcal{D}_N^2 \\ \mathcal{C}_0 \neq \emptyset}} \mu_\beta$. Alors*

$$\sigma_{\text{ess}}(H_N^Z) = [\Sigma, \infty). \tag{3.7}$$

Pour expliquer l'idée de la preuve on va introduire un outil important de la mécanique quantique multi-particules.

Definition 3.3. *Un ensemble de fonctions $\mathcal{J} \subset C^1(\mathbb{R}^{3N}; [0,1])$ tel que*

$$\sum_{J \in \mathcal{J}} J^2(x) = 1 \text{ pour tout } x \in \mathbb{R}^{3N} \tag{3.8}$$

s'appelle une partition de l'unité dans $\mathbb{R}^{3N}$.

On rappelle qu'on partitionne par rapport à la somme des *carrés* des fonctions de troncature, ce qui implique pour $\psi \in L^2(\mathbb{R}^{3N})$ et une partition de l'unité $\mathcal{J}$ que l'on a

$$\|\psi\|^2 = \langle \psi, \psi \rangle = \langle \sum_{J \in \mathcal{J}} J^2 \psi, \psi \rangle = \sum_{J \in \mathcal{J}} \|J\psi\|^2. \tag{3.9}$$

Il existe une multitude de partitions de l'unité. Pour des applications différentes on a besoin de choisir une partition spécifique pour chaque scénario. Pour la preuve du théorème HVZ on choisit une partition de l'unité donnée par la classe des fonctions $\{J_\beta\}_{\beta \in \mathcal{D}_N^2}$ telle que sur le support de la fonction J_β, la distance entre les particules dans des clusters differents est au moins R, où R est un paramètre ajustable. De façon détaillée, pour $x = (x_1, \cdots, x_N) \in \mathrm{supp}(J_\beta)$ on a $|x_i - x_j| > R \quad \forall i \in \mathcal{C}_0,\ j \in \mathcal{C}_1$. En plus on suppose que, indépendemment du paramètre R, on a

$$\sum_{\beta \in \mathcal{D}_N^2} J_\beta^2 \equiv 1. \tag{3.10}$$

Un exemple d'une partition qui satisfait ces conditions est une *partition de l'unité Ruelle-Simon*, voir [13, p.32] pour une définition et la preuve de l'existence d'une telle partition.

On va exposer l'idée de la preuve du théorème HVZ d'une manière non rigoureuse. Voir [13] pour une preuve rigoureuse avec énergie non-relativiste, ou chapitre A pour une preuve rigoureuse pour le cadre pseudo-relativiste.

On peut démontrer l'inclusion $\sigma_{\mathrm{ess}}(H_N^Z) \supseteq [\Sigma, \infty)$ par une construction explicite d'une suite de Weyl pour toute valeur $\lambda \in [\Sigma, \infty)$ arbitraire. Pour cette construction il y a plusieurs approches, on présentera une idée attribuée à Hunziker [20], qui aide à comprendre les propriétés spectrales.

Pour la décomposition minimisant β qui donne l'énergie Σ, il existe une fonction normalisée $\phi \in C_0^\infty(\mathbb{R}^{3N})$ telle que pour un quelquonque $\epsilon > 0$

$$\|(H_\beta - \Sigma)\phi\| \leq \epsilon \tag{3.11}$$

puisque C_0^∞ est une partie dense de L^2. On souhaite augmenter l'énergie cinétique par le facteur (positif) $\lambda - \Sigma$. On remarque que par l'application de la transformation unitaire $e^{-i\mu x_j}$ à un état, on augmente l'énergie cinétique du j-ème électron par la quantité μ^2. Comme le potentiel de Coulomb est invariant par rapport à cette transformation $e^{-i\mu x_j}$, qu'on appelle une *translation dans l'espace des moments*, on a construit une fonction propre approximative de H_β qui correspond à la valeur $\lambda = \mu + \Sigma$.

La deuxième partie de la construction est liée à l'interprétation physique du spectre essentiel. Pour l'évolution des fonctions dans la suite on va "imiter" l'évolution temporelle d'un état dans lequel les électrons du cluster $\mathcal{C}_0$ tendent vers l'infini. La translation simultanée de tous les électrons du cluster $\mathcal{C}_0$ par un vecteur $a \in \mathbb{R}^3$ laisse toutes les parties de l'énergie totale inchangées, sauf l'interaction intercluster I_β. Pour vérifier cela, on remarque que $-\Delta$ est invariant par translation et la répulsion entre deux électrons dépend de la position *relative* des électrons qui n'est pas changée.

Pour récapituler, par l'application d'une translation dans l'espace des moments à la fonction ϕ, on a obtenu une fonction propre approximative ψ_0 correspondant à la valeur λ. Puis, en choisissant une translation dans l'espace de position qui repousse le cluster $\mathcal{C}_0$ de plus en plus loin, on construit une suite $(\psi_l)_{l \in \mathbb{N}}$ de fonctions propres approximatives de sorte que $\|(H_\beta - \lambda)\psi_l\| \to 0$ et $\|I_\beta \psi_l\| \to 0$, puisque la distance entre les clusters augmente. Le choix précis des translations $|a_{l-1}| < |a_l|$ cause une convergence faible vers zéro. Donc toutes les conditions du critère de Weyl sont satisfaites.

Pour une construction rigoureuse en suivant la même idée dans un scénario un peu différent, voir chapitre A. Pour une preuve alternative qui utilise des opérateurs d'ondes, voir [38].

L'inclusion $\sigma_{\text{ess}}(H) \subseteq [\Sigma, \infty)$ utilise le concept d'une partition de l'espace d'états par l'application d'une partition de l'unité définie en (3.8), appliquée à l'espace $L^2(\mathbb{R}^{3N})$ (ou de manière équivalente aux opérateurs sur cet espace). On obtient

$$\begin{aligned} H_N^Z &= \sum_{\beta \in \mathcal{D}_N^2} J_\beta^2 (H_\beta + I_\beta) \\ &= \sum_{\beta \in \mathcal{D}_N^2} J_\beta H_\beta J_\beta + \sum_{\beta \in \mathcal{D}_N^2} J_\beta^2 I_\beta - \sum_{\beta \in \mathcal{D}_N^2} J_\beta (H_\beta J_\beta - J_\beta H_\beta), \end{aligned} \tag{3.12}$$

où le troisième terme dans le membre de droite s'appelle *erreur de localisation.* On remarque que la fonction J_β commute avec le potentiel Coulombien, donc

$$J_\beta (H_\beta J_\beta - J_\beta H_\beta) = J_\beta \sum_{j=1}^N [-\Delta_j, J_\beta] = \sum_{j=1}^N |\nabla_j J_\beta|^2. \tag{3.13}$$

La décomposition en (3.12) ainsi que l'identité (3.13) s'appelle *formule de localisation IMS* d'après Ismagilov, Morgan et Simon. Pour plus de détails concernant l'origine et l'application de la formule voir [13, p.28] et les références incluses.

On cherche à démontrer que les "perturbations" $J_\beta^2 I_\beta$ et $\sum_{j=1}^N |\nabla_j J_\beta|^2$ de l'opérateur H_N^Z sont suffisamment "faibles" pour laisser le spectre essentiel invariable, qui implique que

$$\sigma_{\text{ess}}(H_N^Z) = \sigma_{\text{ess}}(H_\beta) \subseteq [\Sigma, \infty). \tag{3.14}$$

L'opérateur $J_\beta^2 I_\beta$ est "petit", par la propriété de changement d'échelle en $1/R$ du potentiel Coulombien, et car les fonctions de troncature de paramètre R des fonctions J_β vérifient alors $J_\beta^2 I_\beta \sim R^{-1}$.

L'erreur de localisation, déterminée par la norme de ∇J_β dépend du choix de la troncature. Cette dépendance du terme d'erreur est prise en compte pour le choix de la troncature. En général on peut choisir la fonction de troncature telle que $|\nabla J_\beta|^2_{\max} \sim CR^{-2}$ pour une constante $C > 0$.

Ces deux inclusions prouvent l'énoncé (3.7). On remarque qu'il existe plusieurs formulations légèrement différentes du théorème HVZ. Une différence significative est que la valeur Σ qui marque l'infimum de $\sigma_{\text{ess}}(H_N^Z)$ est explicitement donnée par l'énergie de l'état de repos du Hamiltonien H_{N-1}^Z, qui décrit un noyau avec le même numéro atomique interagissant avec $N-1$ électrons. Dans cette excursion dans la preuve du théorème HVZ, la partition de l'unité est formulée comme un objet assez vague. Dans le prochain chapitre, on va essayer d'établir une compréhension pratique de la phrase "partition de l'espace d'état".

3.2. **Partition de l'espace d'état, ou localisation.** La partition de l'espace d'état est un des outils principaux qu'on utilise pour obtenir des bornes pour l'espérance de l'énergie totale. C'est par exemple le cas si on estime l'énergie totale de deux atomes à une distance "assez grande" l'un de l'autre, dans le sens où les atomes ne sont pas considérés comme formant une molécule. Dans une perspective physique on s'attend à ce que les interactions entre les électrons contenus dans des atomes distincts sont plus faibles que les interactions entre des particules au sein d'un atome. En termes mathématiques, cela n'est pas tout à fait clair. Bien que l'on dit qu'un électron "tourne autour du noyau", la mécanique quantique n'est pas déterministe et même pour un état lié, un électron peut être dans n'importe quelle position dans l'espace avec une probabilité strictement positive. Comme le potential Coulombien de deux électrons possède une singularité, même une petite probabilité pour deux particules d'être proches pourra introduire des problèmes

pour une estimation rigoureuse. C'est là qu'une localisation par rapport à une partition de l'unité est très utile.

Soit $\psi \in L^2(\mathbb{R}^6)$ un état arbitraire qui décrit deux électrons qui bougent dans un potentiel d'un noyau de charge $2e$ positionné à l'origine. Une possibilité pour estimer l'énergie de cet état est l'utilisation d'une fonction de localisation $u \in C^1(\mathbb{R}^3; [0,1])$ telle que

$$u(z) := \begin{cases} 1 & \text{if } |z| \leq R \\ 0 & \text{if } |z| > 2R \end{cases} \tag{3.15}$$

et une fonction $v \in C^1(\mathbb{R}^3; [0,1])$ telle que $u^2 + v^2 \equiv 1$. En utilisant ces deux fonctions on peut définir une partition de l'unité dans $\mathbb{R}^6$ composée des quatre fonctions

$$\begin{aligned} J_1(x_1, x_2) &:= u(x_1)u(x_2) \\ J_2(x_1, x_2) &:= u(x_1)v(x_2) \\ J_3(x_1, x_2) &:= v(x_1)u(x_2) \\ J_4(x_1, x_2) &:= v(x_1)v(x_2). \end{aligned} \tag{3.16}$$

Évidemment $\sum J_i^2 = 1$ et on peut utiliser ces fonctions pour décomposer n'importe quelle fonction dans l'espace d'état $L^2(\mathbb{R}^6)$. On note que

$$\|\psi\|^2 = \langle \sum_{i=1}^{4} J_i^2 \psi, \psi \rangle = \sum_{i=1}^{4} \langle J_i\psi, J_i\psi \rangle = \sum_{i=1}^{4} \|J_i\psi\|^2. \tag{3.17}$$

En gros, $J_1\psi$ est la partie de l'état ψ où les deux électrons sont proches du noyau. On s'attend à ce que l'énergie de cette partie soit quelque part dans le spectre discret d'un atome d'hélium. Pour les parties $J_2\psi$ et $J_3\psi$ on s'attend à ce que l'énergie soit dans le spectre d'un ion He^+.

Envisageons donc le Hamiltonien H_2^2 qui décrit l'énergie d'un tel système. L'énergie moyenne d'un état $\psi \in L^2(\mathbb{R}^6)$ est donnée par $\langle \psi, H_2^2\psi \rangle$. En utilisant la partition de l'unité définie au-dessus on obtient

$$\langle \psi, H_2^2\psi \rangle = \sum_{i=1}^{4} \langle J_i\psi, H_2^2 J_i\psi \rangle + \sum_{i=1}^{4} |\nabla J_i|^2. \tag{3.18}$$

Comme nous l'avons discuté au-dessus, la localisation nous donne quelques informations: on connait en gros les valeurs moyennes pour les énergies de $J_i\psi$. Mais la décomposition vient avec l'introduction d'un terme positif, l'erreur de localisation. En règle générale, plus on gagne d'informations, plus grande est l'erreur de localisation.

3.3. **Les valeurs propres discrètes.** Pour la structure des atomes et molécules, le spectre discret coïncide avec les niveaux d'énergie des états qui sont physiquement pertinents. L'étude du spectre discret est la détermination du nombre de valeurs propres discrètes, leur localisation et la dimension des espaces propres correspondantes. La valeur propre minimale du Hamiltonien, qui est l'énergie de l'état de repos, est particulièrement intéressante.

On va examiner le Hamiltonien H_N^Z défini en (3.4) qui décrit un noyau qui interagit avec N électrons.

En général il y a trois cas distincts. Un nombre infini de valeurs propres dans $\sigma_{\text{disc}}(H_N^Z)$, un nombre fini de valeurs dans le spectre discret, et le cas où le spectre discret est vide. Il existe des exemples pour tous les cas.

Pour des atomes, donc $N = Z$, le spectre discret de l'opérateur H_Z^Z est fini. Pour le cas avec un noyau fixé qu'on considère, on peut obtenir ce résultat directement du *théorème de Zhislin* quand on tient compte de l'énergie cinétique des noyaux,

voir [38, Theorem XIII.7]. L'énoncé est adapté à la notation utilisée dans cette introduction.

Theorem 3.4 (Théorème de Zhislin). *Soit $H^Z(M)$ un Hamiltonien "atomique" de la forme*

$$H^Z(M) = \sum_{j=1}^{Z} \left(-\frac{\Delta_j^2}{2m} - \frac{Z}{|x_j|} \right) + \sum_{1 \le i < j \le Z} \frac{\nabla_j \cdot \nabla_i}{M} + \frac{1}{|x_i - x_j|} \tag{3.19}$$

où m est la masse d'un électron et M la masse du noyau.
Alors $\sigma_{\text{disc}}(H^Z(M))$ est infini.

Remark 3.5. *Cela implique l'infinitude de $\sigma_{\text{disc}}(H_Z^Z)$ si on laisse $M \to \infty$, par un changement d'échelles du système qui conduit aux unités atomiques, et qu'on utilise dans la définition* (3.4) *de l'opérateur H_Z^Z.*

Intuitivement, si on met de plus en plus d'électrons proches d'un noyau, à un moment donné la répulsion des électrons l'emporte sur l'attraction du noyau, et cela cause l'évasion d'un électron. Cela suggère l'existence d'une valeur $N_{\max}(Z)$, le nombre maximal d'électrons, qui peut être retenu par un noyau de nombre atomique Z. Le théorème de Ruskai-Sigal nous donne une borne supérieure

$$\limsup_{Z \to \infty} \frac{N_{\max}(Z)}{Z} \le 2. \tag{3.20}$$

Cette borne n'est pas assez significative pour le cas physique. Pour des détails sur des améliorations, voir la discussion après [13, Theorem 3.15] et les références incluses. L'amélioration de Lieb de la borne $N_{\max}(Z) < 2Z + 1$ donne la valeur critique qui est expérimentalement attendue seulement dans le cas de l'hydogène. Dans la nature la négativité des ions, en général, ne dépasse jamais la valeur 2. Les résultats théoriques pour une estimation de $N_{\max}(Z)$ d'une manière rigoureuse sont loin de la valeur attendue. D'après [33] la *conjecture d'ionisation*: $N_{\max}(Z) \le Z+1$ ou $N_{\max}(Z) \le Z + 2$ pour tout Z n'a pas encore été démontrée, sauf dans le cas $Z = 1$ discuté ci-dessus.

La collection des résultats suivants est basée sur un article de revue de Nam, voir [33]. Il est connu que le principe de l'exclusion de Pauli est une condition nécessaire pour la conjecture d'ionisation. Pour des électrons "bosoniques", il existe une borne inférieure de Benguria et Lieb [10] qui contredit la conjecture dans ce cas là. La neutralité asymptotique des atomes

$$\limsup_{Z \to \infty} \frac{N_{max}(Z)}{Z} = 1 \tag{3.21}$$

a été prouvé par Lieb, Sigal, Simon et Thirring [26]. Fefferman, Seco [15] et Seco, Sigal, Solovej [41] ont montré que

$$N_{max}(Z) \le Z + CZ^{\frac{5}{7}}.$$

C'est la meilleure borne pour des atomes grands. Pour les atomes de moyenne taille, c'est à dire si $Z \ge 6$, la borne

$$N_{max}(Z) < 1.22Z + 3Z^{\frac{1}{3}}$$

par Nam [34] est une amélioration du résultat de Lieb.

En général le nombre des valeurs propres dans le spectre discret de H_N^Z pour des ions positifs et négatifs est discuté en [13, §3.9] et dépasse le contexte de cette introduction. L'existence d'un spectre discret dans le cas $N \le Z$ est prouvée à l'aide d'une construction facile donnée dans le chapitre suivant.

3.4. **Énergie au-dessous du spectre essentiel.** Dans ce chapitre on va construire un état avec une énergie au-dessous du spectre essentiel de H_N^Z pour des atomes et des ions positifs ($N \leq Z$). Cela implique que le spectre discret $\sigma_{\text{disc}}(H_N^Z)$ n'est pas vide et en particulier, comme le Hamiltonien H_N^Z est semi-borné inférieurement, l'existence d'une énergie de repos avec un espace propre de dimension finie.

D'après une autre formulation du théorème HVZ qui donne plus d'informations sur le bas du spectre essentiel du Hamilonien quand $N \leq Z$, on a

$$\inf \sigma_{\text{ess}}(H_N^Z) = \inf \sigma(H_{N-1}^Z).$$

On utilise cette équation pour démontrer inductivement l'existence d'un état de repos pour un Hamiltonien qui décrit des ions positifs H_k^Z, pour $k = 1, \cdots, Z$. À partir d'un état de repos de H_k^Z dans $L^2(\mathbb{R}^{3k})$, $k < Z$, on construit un état produit dans $L^2(\mathbb{R}^{3(k+1)})$, tel que l'énergie du nouvel état est strictement inférieure, ce qui implique l'existence d'une valeur propre au-dessous du spectre essentiel.

Theorem 3.6. *Soit H_N^Z le Hamiltonien défini en* (3.4)*, alors pour n'importe quel $N \leq Z$, le spectre discret n'est pas vide.*

Encore une fois, on ne va pas présenter une preuve rigoureuse. On va essayer d'expliquer l'orientation générale d'une manière compréhensible. On remarque que l'énoncé de ce théorème est une conséquence immédiate de la finitude ou infinitude du spectre discret, voir p. ex. [13, §3.9] pour une preuve rigoureuse et complète.

On définit

$$\mu_k^Z := \inf \sigma(H_k^Z) \quad \text{et} \quad \Sigma_k^Z := \inf \sigma_{\text{ess}}(H_k^Z).$$

Pour $k = 1$, le spectre discret est infini, voir [38, Theorem XIII.6]. On suppose que l'assertion est prouvée pour un $k < Z$ quelquonque fixé. On choisit une fonction propre approximative, normalisée, $\phi(x_1, \cdots, x_k) \in C_0^\infty(\mathbb{R}^{3k})$ associée à la valeur propre μ_k^Z, avec support sur la boule de rayon δR, $\delta < 1$, centrée à l'origine telle que

$$\langle \phi, (H_k^Z - \mu_k^Z)\phi \rangle \leq d(\delta R) \tag{3.22}$$

pour une fonction $d(\delta R)$ pour laquelle on va démontrer que, pour δ fixé, la fonction $d(\delta R)$ est suffisamment faible pour R grand. On note que

$$H_{k+1}^Z = H_k^Z - \frac{\Delta_{k+1}}{2} - \frac{e^2 Z}{|x_{k+1}|} + \sum_{j=1}^{k} \frac{e^2}{|x_j - x_{k+1}|}. \tag{3.23}$$

On veut trouver une fonction avec une petite valeur moyenne pour les trois derniers termes de (3.23). À cet effet on prend une fonction normalisée $f \in C_0^\infty(\mathbb{R}^3)$ supportée sur la région où $1 < |x_{k+1}| \leq 2$. Heuristiquement on peut imaginer une "bosse" lisse sur $[1, 2]$. On définit une fonction similaire à f, étalée sur l'intervalle $R < |x_{k+1}| \leq 2R$, sans changement de la norme de f:

$$f_R(x_{k+1}) := R^{-\frac{3}{2}} f\left(\frac{x_{k+1}}{R}\right). \tag{3.24}$$

On souligne la similitude avec l'exemple de la fonction Gaussienne du chapitre 2.7. Le comportement de l'énegrie cinétique de la fonction f_R est

$$\langle f_R, -\Delta_{k+1} f_R \rangle \sim CR^{-2}.$$

Dans la prochaine étape on va utiliser la positivité de l'ion décrit par H_k^Z et ϕ. Puisqu'on a $k < Z$, la partie du potentiel concernant le $(k+1)$-ème électron est négative, de l'ordre $\sim -cR^{-1}$ étant donné $\delta < \frac{1}{Z}$. On note que c'est juste, seulement si les deux fonctions ϕ et f_R sont supportées sur des régions où la distance minimale est proportionnelle à R. On omet les détails pour une telle borne, les détails sont données dans le chapitre B.

Comme discuté au-dessus pour l'état produit

$$\psi_R(x_1, \cdots, x_{k+1}) := \phi(x_1, \cdots, x_k) f_R(x_{k+1}) \tag{3.25}$$

on a une valeur moyenne de l'énergie

$$\langle \psi_R, (H^Z_{k+1} - \mu^Z_k)\psi_R \rangle \lesssim d(\delta R) + \frac{C}{R^2} - \frac{c}{R}. \tag{3.26}$$

Il reste à démontrer que $d(\delta R)$ est dominé par le terme négatif $-\frac{c}{R}$ pour R grand. Cette propriété dépend du comportement asymptotique des états de repos. Dans le prochain chapitre nous allons analyser ce comportement asymptotique et nous allons voir que $d(\delta R)$ est en fait exponentiellement petit en R.

3.5. **Comportement asymptotique pour les états liés.** Dans la discussion sur le spectre nous avons établi pour nos modèles que les fonctions propres associées aux énergies dans le spectre discret correspondent à des états dans lesquels les électrons restent essentiellement localisés dans une région confinée autour du noyau; et il est improbable de trouver un électron loin du noyau car nous avons vu dans la construction d'une fonction test dans le théorème HVZ que les états pour lesquels au moins un électron s'échappe à l'infini sont associés à des énergies dans le spectre essentiel. Ainsi, les états propres dans le spectre discret correspondent bien à la physique dans laquelle on n'observe pas d'électron qui quitte l'entourage du noyau pour y revenir sans action extérieure au système. Cette propriété pour les états liés donne à son tour des informations pour les états diffusifs. De plus, cela permet d'estimer plus précisément les interactions entre atomes étant donnés des atomes dans un état d'énergie basse.

Le résultat le plus connu concernant le comportement asymptotique des états bornés pour les opérateurs de Schrödinger à plusieurs particules est le fameux résultat dû à Agmon. Pour le cas discuté en introduction, voir par exemple [2, Theorem 6.4]. Nous donnons ici le résultat tel qu'il est présenté dans [16]

Theorem 3.7 (Agmon). *Soit H^Z_N le Hamiltonien à N-corps donné par* (3.4)*, soit ϕ une fonction propre correspondant à une valeur propre $\lambda < \Sigma^Z_N = \inf \sigma_{\text{ess}}(H^Z_N)$, et choisissons $\beta > 0$ tel que $\beta^2 < \Sigma^Z_N - \lambda$. Alors*

$$\|e^{-\beta^2|x|}\phi\|^2 \leq C \left(\frac{1}{\Sigma^Z_N - \lambda - \beta^2} \right) \|\phi\|^2 \tag{3.27}$$

où C est une constante universelle.

On notera que (3.27) suggère que les états liés de faible énergie décroissent plus rapidement que les états d'énergie proche du spectre essentiel.

Nous avons donné un aperçu général des résultats les plus importants pour les opérateurs de Schrödinger à plusieurs particules H^Z_N. Dans la suite, nous discuterons du cas où l'opérateur énergie cinétique $-\Delta$ est remplacé par un opérateur qui prend en compte le changement de masse pour des particules qui approchent la vitesse de la lumière.

3.6. **L'opérateur d'énergie relativiste.** Jusqu'ici, l'énergie cinétique des électrons a été modélisée par l'opérateur $-\Delta$ déduit de la mécanique newtonnienne. La fameuse équation d'Einstein pour l'énergie totale d'une particule

$$E_{\text{tot}} = \sqrt{p^2c^2 + m^2c^4} \tag{3.28}$$

postule l'augmentation de la masse m pour un objet se déplaçant à une vitesse proche de la vitesse de la lumière c. En soustrayant l'énergie $E_0 = mc^2$ d'une particule au repos, on obtient la partie cinétique, qui donne, dans le contexte de la mécanique quantique, l'opérateur *énergie cinétique relativiste*

$$T_{\text{rel}}(p) := \sqrt{p^2c^2 + m^2c^4} - mc^2. \tag{3.29}$$

Cette correction relativiste devient pertinente dans le cas d'atomes lourds dans lesquels les atomes des couches internes proches du noyau possèdent une énergie cinétique élevée, voir [25, Part VII]. Pour de faibles impulsions p, l'opérateur T_{rel} se comporte de façon similaire à l'opérateur non relativiste $\frac{p^2}{2m}$, alors que pour les impulsions p élevées, les énergies ont un comportement de l'ordre de $|p|$. D'après les équations du mouvement de Newton, la vitesse v d'une particule d'énergie cinétique $T(p)$ est donnée par

$$v = \frac{d}{dp}T(p) \tag{3.30}$$

et en injectant l'opérateur énergie cinétique $T_{\mathrm{rel}}(p)$ défini par (3.29) on obtient

$$v_{\mathrm{rel}} = \frac{pc^2}{\sqrt{p^2c^2 + m^2c^4}}. \tag{3.31}$$

On peut alors voir que $v_{\mathrm{rel}} \xrightarrow{p\to\infty} c$, et que la vitesse de la lumière est une borne supérieure pour la vitesse, en accord avec la relativité spéciale. On notera bien cependant que ce *n'est pas* une théorie relativiste de la mécanique quantique, l'implémentation de la relativité spéciale dans la mécanique quantique restant un des grands problèmes non résolu en physique théorique.

Nous nous concentrerons sur le problème purement mathématique que l'on obtient lorsqu'on remplace l'opérateur $\frac{p^2}{2m}$ par l'opérateur $\sqrt{p^2c^2 + m^2c^4} - mc^2$, et on fera référence aux modèles ayant cette dernière énergie cinétique comme des modèles ou systèmes *pseudo-relativistes*. De façon similaire au cas non-relativiste, on peut effectuer un changement d'échelle tel que dans les unités atomiques on ait $T_{\mathrm{rel}} = \sqrt{p^2+1} - 1$. Nous noterons dans ce chapitre par $\mathfrak{H}_N^Z$ l'opérateur pseudo-relativiste correspondant à l'opérateur H_N^Z défini dans (3.4). Explicitement, cela donne

$$\mathfrak{H}_N^Z := \sum_{j=1}^{N} \left(\sqrt{p_j^2+1} - 1 - \frac{e^2 Z}{|x_j|} \right) + \sum_{1\le i<j\le N} \frac{e^2}{|x_i - x_j|}. \tag{3.32}$$

Les résultats qui ont été introduits pour le Laplacien dans les sous-sections précédentes ne peuvent pas être utilisés directement pour l'opérateur $\mathfrak{H}_N^Z$. Cela inclut en particulier les preuves du caractère (essentiellement) auto-adjoint et les résultats sur la structure du spectre. La preuve du caractère auto-adjoint était basée sur la borne relative du potentiel de Coulomb par rapport au Laplacien. À un niveau fondamental de l'analyse des opérateurs correspondants, on distingue deux différences cruciales entre le cas d'opérateurs non-relativiste et le cas d'opérateurs pseudo-relativistes.

La première différence est que T_{rel} n'est pas un opérateur local. Cela signifie que pour obtenir la valeur de $T_{\mathrm{rel}}\psi$ en un point $x \in \mathbb{R}^3$, cela ne suffit pas de connaître les valeurs de ψ dans un voisinage de x, contrairement au cas du Laplacien. A priori, cela signifie que la manipulation des états, comme la localisation par exemple, peut avoir un effet global non trivial sur l'énergie. Ceci est important en particulier car dans la preuve du théorème HVZ et dans celle de l'existence d'un état fondamental, dans le cas non-relativiste, on peut modifier l'état "loin" du noyau sans changer ses propriétés cinétiques près du noyau.

La seconde différence est que si l'on a une fonction à support compact sur laquelle on effectue un changement d'échelle dans la variable d'espace, on passe pour l'énergie cinétique d'un facteur qui est l'inverse du carré du paramètre de changement d'échelle dans le cas non-relativiste à simplement l'inverse du paramètre dans le cas relativiste. On peut facilement voir ceci par exemple dans le cas des fonctions gaussiennes que l'on a contractées pour illustrer la stabilité de premier type dans le chapitre 2.7. Donc, dans le cas relativiste, ce changement d'échelle donne un

facteur d'échelle du même ordre que celui pour le potentiel Coulombien, et donc on ne peut pas utiliser le même argument de stabilité que dans le cas non-relativiste pour espérer que l'énergie cinétique contrôle la singularité du potentiel de Coulomb, sans contrainte sur les constantes de couplages. Dans la litérature, le Hamiltonien pseudo-relativiste $\mathfrak{H}_N^Z$ est souvent dénommé Hamiltonien de *Herbst*. Herbst a été le premier à étudier cet opérateur. Dans [19], il démontre que le Hamiltonien $\mathfrak{H}_1^Z$ qui décrit un atome hydrogénoïde est auto-adjoint sur l'espace de Sobolev fractionnaire $H^{1/2}(\mathbb{R}^3)$ si le numéro atomique pour le noyau satisfait l'inégalité $Z \leq \frac{2}{\alpha\pi} \approx 87$, où $\alpha \approx \frac{1}{137}$ est la constante de structure fine de Sommerfeld. On notera que la même valeur critique apparait lorsque la stabilité d'un Hamiltonien à plusieurs particules $\mathfrak{H}_N^{\mathcal{Z}}$ est discutée par Lieb et Yau dans [28]. Ces derniers montrent que tout Hamiltonien de la forme $\mathfrak{H}_N^{\mathcal{Z}}$ est *stable* si et seulement si $Z_i \leq \frac{2}{\alpha\pi}$ pour tout $i = 1, \cdots, M$. Ici, la stabilité dont il est question est la stabilité de second type que nous n'évoquons pas ici, mais qui implique en particulier la stabilité de premier type du Chapitre 2.7 et le caractère auto-adjoint.

Une partie de l'article central de cette thèse est dédié à l'estimation de l'erreur de localisation. Voir (3.12) pour le cas non-relativiste pour lequel on utilise la formule de localisation IMS. Dans le cas pseudo-relativiste on peut utiliser une approche similaire, mais l'estimation obtenue est alors très différente. Généralement l'erreur de localisation pour un Hamiltonien $H = T+V$ d'énergie cinétique T et de potentiel multiplicatif V pour une partition de l'unité $\{J_\beta\}$ est donné par

$$\sum_\beta J_\beta[T, J_\beta]. \tag{3.33}$$

On notera que l'erreur de localisation est un opérateur. Mais la terminologie est aussi utilisée pour désigner les quantités

$$\sum_\beta \langle \psi, J_\beta[T, J_\beta]\psi\rangle$$

que l'on devrait plus précisément appeler erreur de localisation dans l'état ψ.

Prendre $T = -\Delta$ dans (3.33) donne lieu à l'opérateur $|\nabla J_\beta|^2$ qui est supporté dans une région où $0 < J_\beta < 1$. Par la propriété de localité de l'opérateur laplacien, l'augmentation de l'énergie moyenne due à la coupure dépend de la densité de probabilité de l'état dans la région de la coupure. Ce *n'est pas* le cas pour l'énergie cinétique relativiste qui donne lieu à un opérateur de localisation global. L'estimation de l'erreur de localisation (pour l'énergie) devient alors une tâche plus compliquée que dans le cas non-relativiste.

Beaucoup des résultats concernant la structure spectrale de Hamiltonien de Schrödinger avec une énergie cinétique non-relativiste que l'on a mentionnés jusqu'ici seront démontrés dans le cas non-relativiste qui nous intéresse, dans les résultats centraux de cette thèse, pour des classes spécifiques d'états physiquement pertinents pour notre problème.

Pour l'opérateur $\mathfrak{H}_N^Z$, dans le cas plus simple où on néglige les symétries du système et le fait que les électrons sont des particules identiques indiscernables, certains des résultats sont connus et sont discutés dans cette thèse. Le cas des états physiquement pertinents est discuté dans le paragraphe suivant. La prise en compte mathématiquement rigoureuse des ces états est traitée dans les résultats centraux de cette thèse à partir du Chapitre 4.

3.7. **Considérations de symétries.** Un aspect important des résultats de cette thèse est la prise en compte des restrictions de l'espace des états à des sous-espaces qui tiennent compte du principe d'exclusion de Pauli qui postule que deux électrons ne peuvent pas se trouver dans le même état quantique. La formulation en

mécanique quantique de ce principe d'exclusion donne une condition mathématique qui nous permet d'exclure des états physiquement impossibles. Le Chapitre G.3 contient des détails sur les symétries de permutation et une introduction à la théorie des représentations, ainsi que ses utilisations en physique mathématique et en analyse spectrale.

Dans ce paragraphe, il nous suffira d'énoncer que pour obtenir un modèle en accord avec la réalité physique, on restreint l'espace des états quantiques à certains espaces mutuellement orthogonaux que l'on peut classifier en fonction de leurs propriétés de symétrie de permutation. On considèrera alors séparément les Hamiltoniens restreints à chacun de ces sous-espaces physiquement pertinents. Selon la structure du spectre, cette approche va nous donner beaucoup d'informations. Souvenons-nous par exemple que la décroissance des états liés ne s'applique pas aux états plongés dans le spectre essentiel. De plus, on ne sait pas montrer le caractère fini ou pas pour la dimension des sous-espaces propres correspondant à des valeurs propres plongées dans le spectre essentiel. Si on étudie nos Hamiltoniens restreints aux sous-espaces physiquement pertinents, cela pourra conduire à translater à droite l'infimum du spectre essentiel et alors permettre l'étude des valeurs propres qui nous intéressent et qui étaient, sans considération de symétrie, au dessus de l'infimum du spectre essentiel.

Dans [50] il est énoncé que si l'on néglige les interactions de Coulomb électron-électron dans H_N^Z, alors, pour $N \geq 4$, "all the eigenvalues corresponding to a realizable symmetry lie in the continuous spectrum. One would expect this is also true for the complete operator H". Cette phrase est à comprendre dans le sens où toutes les valeurs propres correspondant à une symétrie physique, qui correspondent donc aux états liés réels, sont celles qui se trouvent dans le spectre continu de l'opérateur sans restriction de symétrie, et que de plus, cette propriété devrait rester vraie si on inclut les interactions électron-électron. L'étude des valeurs propres plongées n'étant pas possible avec les outils standards de l'analyse spectrale, cela souligne la pertinence de la décomposition de l'espace des états en tenant compte des symétries physiques.

Dans la partie suivante, nous aimerions introduire le problème qui motive le coeur de cette thèse. Notre étude des forces de van der Waals-London dans le contexte pseudo-relativiste en tenant compte des symétries physiques dues au principe de Pauli a mis en évidence que certains des résultats principaux de l'analyse spectrale des opérateurs à plusieurs particules en mécanique quantique n'avaient pas été démontrés dans le cas pseudo-relativiste, ou bien avaient été énoncés sans preuve. Pour démontrer nos résultats, nous utiliserons généralement des techniques bien établies que nous adapterons aux besoins spécifiques des opérateurs pseudo-relativistes et des symétries de permutation considérées. Certaines de ces méthodes ont déjà été expliquées de façon non rigoureuse dans le Chapitre 3, servant ainsi de guide dans la compréhension des preuves détaillées et rigoureuses présentées dans la suite, dans lesquelles des notations compliquées, intrinsèques au problème, sont nécessaires pour rédiger les démonstrations mais peuvent avoir tendance à entraver la compréhension des idées directrices.

4. La force de dispersion de Van der Waals-London

La force de van der Waals joue un rôle vital dans de nombreux phénomènes naturels. C'est l'un des aspects clefs qui permet de comprendre les liaisons moléculaires. Pour donner quelques exemples, la force de van der Waals est nécessaire pour expliquer la condensation de la vapeur, la stabilité structurelle de l'ADN ou les liaisons entre différentes couches de graphène dans le graphite.

La force de van der Waals a une importance qui ne se restreint pas aux aspects microscopiques. Elle est utilisée pour expliquer certains processus biologiques. Il y a aussi de nombreux efforts faits en nanotechnologies pour prendre avantage de ces forces attractives. Pour des exemples, voir l'introduction de l'article [3].

On notera qu'à ce point, il y a une certaine ambiguïté dans le terme "force de van der Waals". Plusieurs interactions inter-atomiques ou inter-moléculaires causées par des effets différents sont appelés "forces de van der Waals". Cela inclut les interactions entre atomes et molécules possédant des monopôles ou multipôles permanents et la répulsion causée par les interactions électron-électron inter-atomiques. Dans ce cas précis, on parlera d'*effet de dispersion* ou *force de dispersion de van der Waals-London (vdW-L)*. C'est donc une interaction inter-atomique ou inter-moléculaire même en l'absence de polarité permanente. La description physique de ce phénomène dans le cas de deux atomes est que les mouvements des électrons dans les différents atomes se synchronisent, créant ainsi des "quasi-dipôles", parfois appelés dipôles *temporaires*, *non-permanents* ou *induits*, qui interagissent entre eux. Afin de coller à la terminologie usuelle, on appellera cette énergie d'interaction la *force* de vdW-L.

De manière surprenante, il existe très peu de résultats mathématiquement rigoureux concernant la force de van der Waals. Dans [29], J. D. Morgan et B. Simon ont obtenu un développement asymptotique de l'énergie d'interaction et ont noté que pour des atomes, en l'absence de moment dipolaire ou quadrupolaire, le terme dominant était en $\mathcal{O}(|D|^{-6})$, où $|D|$ est la distance entre les deux noyaux.

Un autre résultat concernant les interactions de van der Waals a été obtenu dans [27] par E. H. Lieb et W. E. Thirring qui ont construit une fonction test qui montre que l'énergie attractive entre deux atomes sans polarité permanente est au moins de $-C|D|^{-6}$ où C est une certaine constante positive.

Ce résultat a été amélioré par I. Anapolitanos et I. M. Sigal dans [3], qui ont obtenu sous certaines restrictions que nous discuterons plus tard, que le terme dominant est de l'ordre de $|D|^{-6}$ pour l'énergie entre clusters, dans le cas non relativiste, avec une erreur en $\mathcal{O}(|D|^{-7})$.

On notera que tous les résultats précédents ne sont établis que dans le cas d'une énergie cinétique non-relativiste. Pour les atomes lourds il est nécessaire d'inclure les effets relativistes pour les électrons. C'est le but essentiel du travail présenté ici. La méthode de preuve dans [3] est basée pour sa majeure partie sur une méthode de Feshbach-Schur. Notre approche est purement variationnelle et similaire à celle utilisée dans [6, 7, 8, 9, 43, 44, 45] pour obtenir l'asymptotique des valeurs propres pour des opérateurs de Schrödinger à plusieurs particules près du bas du spectre essentiel.

Expliquons maintenant, dans un langage heuristique, l'idée de notre approche dans le cas d'un système diatomique . Nous allons essayer de gagner de l'information sur le spectre discret d'un système diatomique en étudiant le spectre discret de deux atomes identiques, mais isolés l'un de l'autre. Nous estimerons alors l'impact des interactions inter-atomiques en terme d'une énergie d'ensemble minimale. On exprimera la plus petite énergie pour le système diatomique en interaction en terme du développement asymptotique en puissances négatives de $|D|$ "autour" de l'état fondamental du système constitué des deux atomes isolés l'un de l'autre.

Par comparaison avec [3], nous obtenons des termes de plus haut degré dans le développement, et de meilleures estimations pour les restes, que ce soit dans le cas relativiste ou pseudo-relativiste. De plus, notre méthode conduit à une expression explicite des constantes pour les termes jusqu'à l'ordre $|D|^{-9}$ pour l'asymptotique de l'énergie intercluster. Nous avons en outre obtenu une différence notable dans le développement asymptotique pour des systèmes avec plus de deux noyaux. On

montre par exemple que dans le cas deux noyaux, le terme d'ordre $|D|^{-9}$ est nul, alors que pour trois noyaux ou plus, ce terme n'est a priori pas nul. Voir la discussion dans les remarques après le Théorème 9.4.

5. Modèle, notations et résultat principal

On considère une molécule de N électrons de charge $-e$ et de spin $\frac{1}{2}$, et M noyaux de charges eZ_l modélisés par des points fixés (approximation de Born-Oppenheimer) de positions $X_l \in \mathbb{R}^3$. On suppose que le système est neutre. Cela veut dire que $\sum_{l=1}^{M} Z_l = N$. Le Hamiltonien est

$$H := \sum_{i=1}^{N} \left(T_i - \sum_{l=1}^{M} \frac{e^2 Z_l}{|x_i - X_l|} \right) + \sum_{1 \le i < j \le N} \frac{e^2}{|x_i - x_j|} + \sum_{1 \le k < l \le M} \frac{e^2 Z_k Z_l}{|X_k - X_l|} \tag{5.1}$$

et l'opérateur d'énergie cinétique du k-ème électron est

$$T_k := \begin{cases} \sqrt{p_k^2 + 1} - 1 & \text{dans le cas pseudo-relativiste} \\ \frac{p_k^2}{2} & \text{dans le cas non-relativiste} \end{cases} \tag{5.2}$$

Le domaine de forme de H est $H^{1/2}(\mathbb{R}^{3N})$ dans le cas pseudo-relativiste et $H^1(\mathbb{R}^{3N})$ dans le cas non-relativiste. Comme d'habitude $p_k = -i\nabla_{x_k}$ désigne le moment du k-ème électron. Si T_k est pseudo-relativiste, on suppose que $Z_l e^2 \le \frac{2}{\pi}$. Cela implique que le Hamiltonien est semi-borné inférieurement, voir [14, 28].

Dans la partie principale de cet article on va se concentrer sur le cas pseudo-relativiste $T_k = \sqrt{p_k^2 + 1} - 1$ (voir [19] et références incluse) bien que tous les résultats sont vrais de même pour $T_k = \frac{p_k^2}{2}$. Le Hamiltonien est donné en utilisant des unités atomiques où $c = \hbar = m = 1$.

L'espace des phases d'un système de N électrons, prenant en compte le principe de Pauli, est donné par le produit tensoriel antisymétrique de N copies de $L^2(\mathbb{R}^3; \mathbb{C}^2)$, à savoir l'espace $\bigwedge^N L^2(\mathbb{R}^3; \mathbb{C}^2)$ de fonctions antisymétriques par rapport aux transpositions de paires des variables de position et de spin (x_i, s_i) et (x_j, s_j) pour $i \neq j$.

L'opérateur H qu'on considère dépend des variables de position x_i et ne dépend pas des variables de spin s_i. Donc on étudie H agissant sur la projection de $\bigwedge^N L^2(\mathbb{R}^3; \mathbb{C}^2)$ sur l'espace des fonctions dependant seulement des coordonnées, c'est-à-dire l'espace $\mathcal{H}_{\text{Fermi}}$ défini par

$$\mathcal{H}_{\text{Fermi}} := \left\{ \langle \mathfrak{s}, \Psi \rangle_{\text{spin}} | \Psi \in \bigwedge^N L^2(\mathbb{R}^3; \mathbb{C}^2), \mathfrak{s} : \left\{ -\frac{1}{2}, \frac{1}{2} \right\}^N \to \mathbb{C} \right\} \tag{5.3}$$

où

$$\langle \mathfrak{s}, \Psi \rangle_{\text{spin}} := \sum_s \bar{\mathfrak{s}}(s_1, \cdots, s_N) \Psi(x_1, s_1, \cdots, x_N, s_N).$$

On note que $\mathcal{H}_{\text{Fermi}}$ est un sous-espace de $L^2(\mathbb{R}^{3N})$.

La condition d'antisymétrie par rapport aux transpositions des variables des particules implique certaines propriétés symétriques concernant les permutations des variables coordonnées, après découplage des variables de spin. En particulier, les permutations des électrons transforment les fonctions d'après les diagrammes de Young avec deux colonnes au maximum, voir la description en [18, § 7.3.].

De manière plus précise, soit S_N le groupe des permutations de N électrons. Pour tout $\pi \in S_N$ on définit $\mathcal{T}_\pi : \mathcal{H}_{\text{Fermi}} \to \mathcal{H}_{\text{Fermi}}$ où

$$\mathcal{T}_\pi \psi(x_1, \cdots, x_N) := \psi(x_{\pi^{-1}(1)}, \cdots, x_{\pi^{-1}(N)}) \tag{5.4}$$

est l'opérateur qui réalise une permutation des variables de coordonnée.

Soit α n'importe quelle représentation irréductible du groupe S_N et P^α la projection sur le sous-espace des fonctions qui sont transformées par les opérateurs

$\mathcal{T}_\pi$ selon la representation α. Ces projections décomposent l'espace $\mathcal{H}_{\text{Fermi}}$ en une infinitude de sous-espaces orthogonaux $\mathcal{H}^\alpha := P^\alpha \mathcal{H}_{\text{Fermi}}$ tels que

$$\mathcal{H}_{\text{Fermi}} = \bigoplus_{\alpha \in \mathcal{A}} \mathcal{H}^\alpha, \tag{5.5}$$

où $\mathcal{A}$ est l'ensemble de toutes les representations irréductibles du groupe S_N décrit par les diagrammes de Young de deux colonnes au maximum. On note que pour un tel α on a $P^\alpha \mathcal{H}_{\text{Fermi}} = P^\alpha L^2(\mathbb{R}^{3N})$. En fait, l'étude de l'opérateur H sur les sous-espaces $P^\alpha L^2(\mathbb{R}^{3N})$ nous donne toutes les informations sur le spectre de l'opérateur sur $\mathcal{H}_{\text{Fermi}}$. Pour cela on définit

$$H^\alpha := HP^\alpha \tag{5.6}$$

l'opérateur restreint sur l'espace $\mathcal{H}^\alpha$ et

$$E^\alpha_{(X_1,\cdots,X_M)} := \inf \sigma(H^\alpha). \tag{5.7}$$

Dans cet article on va calculer l'énergie d'interaction pour les positions fixées des noyaux, qui est la différence entre $E^\alpha_{(X_1,\cdots,X_M)}$ et la somme des énergies de repos des atomes. Considérons le cas le plus simple d'une molécule diatomique, i.e. $M = 2$. Pour le cas général voir 9.2 dans la partie anglaise.

5.1. **Molécules diatomiques.** Soit $\mathcal{C} \subsetneq \{1, \cdots, N\}$, $\mathcal{C} \neq \emptyset$ un sous-système d'un système de N électrons. On appelle $\mathbb{R}(\mathcal{C})$ l'espace vectoriel des vecteurs de positions $(x_i)_{i\in\mathcal{C}}$ des particules dans $\mathcal{C}$. On note que cet espace est isomorphe à $\mathbb{R}^{3\sharp\mathcal{C}}$, où $\sharp\mathcal{C}$ est le nombre d'éléments de $\mathcal{C}$. On note $L^2(\mathbb{R}(\mathcal{C}))$ l'espace des fonctions L^2 avec arguments dans $\mathbb{R}(\mathcal{C})$. Soit $L^2(\mathbb{R}(\mathcal{C}))^\perp$ le complément orthogonal de $L^2(\mathbb{R}(\mathcal{C}))$ dans $L^2(\mathbb{R}^{3N})$.

L'interaction des particules dans $\mathcal{C}$ avec un noyau fixé à l'origine de charge eZ via un potentiel Coulombien est décrit par le Hamiltonien

$$\tilde{H}^Z_{\mathcal{C}} := \sum_{i\in\mathcal{C}} T_i - \sum_{i\in\mathcal{C}} \frac{e^2 Z}{|x_i|} + \sum_{\substack{i,j\in\mathcal{C}\\ i<j}} \frac{e^2}{|x_i - x_j|} \tag{5.8}$$

sur l'espace $L^2(\mathbb{R}(\mathcal{C}))$. On étend l'opérateur sur $L^2(\mathbb{R}(\mathcal{C}))^\perp$ par l'identité pour que l'on puisse appliquer l'opérateur à une fonction dans $L^2(\mathbb{R}^{3N})$. On va représenter ces deux opérateurs par la notation $\tilde{H}^Z_{\mathcal{C}}$.

Soit $S(\mathcal{C})$ le groupe de permutations de $\mathcal{C}$. Évidemment $S(\mathcal{C})$ est un sous-groupe de S_N. On considère une représentation irréductible $\alpha_{\mathcal{C}}$ de $S(\mathcal{C})$.

Definition 5.1. *Pour un type de représentation irréductible α de S_N on dit que $\alpha'_{\mathcal{C}}$ est induit par α, noté $\alpha'_{\mathcal{C}} \prec \alpha$, si $\alpha'_{\mathcal{C}}$ est contenu dans la restriction de α sur $S(\mathcal{C})$, voir* [18, p. 94-98].

Dans le même esprit que pour la décomposition de $\mathcal{H}_{\text{Fermi}}$ en sous-espaces $\mathcal{H}^\alpha$ on peut décomposer le sous-espace de Fermi correspondant de $L^2(\mathbb{R}(\mathcal{C}))$ en sous-espaces $P^{\alpha_{\mathcal{C}}} L^2(\mathbb{R}(\mathcal{C}))$ pour toutes représentations irréductibles $\alpha_{\mathcal{C}}$ de $S(\mathcal{C})$ décrites par un diagramme de Young avec deux colonnes au maximum.

On considère une décomposition en deux clusters $\beta = (\mathcal{C}_1, \mathcal{C}_2)$ du système d'électrons $\{1, \cdots, N\}$ telle que $\mathcal{C}_1 \cup \mathcal{C}_2 = \{1, \cdots, N\}$ et $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$. On définit $\mathcal{D}^2_N$ comme l'ensemble de toutes telles décompositions. On dit qu'une décomposition est *atomique* si pour le nombre des électrons en $\mathcal{C}_1$, $\sharp\mathcal{C}_1 = Z_1$ et pour le nombre des électrons en $\mathcal{C}_2$ on a $\sharp\mathcal{C}_2 = Z_2$. L'ensemble de ces décompositions atomiques est noté $\mathcal{D}^{at} \subset \mathcal{D}^2_N$.

Pour la décomposition $\beta = (\mathcal{C}_1, \mathcal{C}_2)$ on définit l'interaction inter-cluster

$$I_\beta := \sum_{i\in\mathcal{C}_1} \frac{-e^2 Z_2}{|x_i - X_2|} + \sum_{j\in\mathcal{C}_2} \frac{-e^2 Z_1}{|x_j - X_1|} + \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} \frac{e^2}{|x_i - x_j|} + \frac{e^2 Z_1 Z_2}{|X_2 - X_1|} \tag{5.9}$$

et le Hamiltonien de cluster β par

$$H_\beta := H - I_\beta. \tag{5.10}$$

En d'autres termes, H_β est l'opérateur où les particules des sous-systèmes differents n'interagisent pas. On note que pour tout $\beta \in \mathcal{D}_N^2$ on a $L^2(\mathbb{R}(\mathcal{C}_1))^\perp = L^2(\mathbb{R}(\mathcal{C}_2))$. Le groupe de symétrie de ce Hamiltonien qu'on considère est le groupe $S_\beta := S(\mathcal{C}_1)\times S(\mathcal{C}_2) \subset S_N$ de permutations qui laisse la décomposition en clusters intacte. On utilise la même notion d'induction de représentations qu'au-dessus. Comme S_β est un produit direct de deux groupes, les représentations irréductibles α_β de S_β sont des produits également. En particulier, pour toute représentation irréductible $\alpha'_\beta \prec \alpha$ de S_β il existe une paire unique $\alpha'_{\mathcal{C}_1} \prec \alpha$ et $\alpha'_{\mathcal{C}_2} \prec \alpha$ telle que

$$\alpha'_{\mathcal{C}_1} \otimes \alpha'_{\mathcal{C}_2} \cong \alpha'_\beta, \tag{5.11}$$

voir [18, p. 110-114]. Soit $P^{\alpha'_\beta}$ la projection de $\mathcal{H}_{\text{Fermi}}$ sur les fonctions de type de symétrie α'_β. On pose

$$H_\beta^{\alpha'_\beta} := H_\beta P^{\alpha'_\beta} \quad \text{et} \quad H_\beta^\alpha := \sum_{\alpha'_\beta \prec \alpha} H_\beta^{\alpha'_\beta}, \tag{5.12}$$

on définit

$$\mu_\beta^\alpha := \min_{\alpha'_\beta \prec \alpha} \inf \sigma(H_\beta^{\alpha'_\beta}) \tag{5.13}$$

et

$$\mu^\alpha := \min_{\beta\in\mathcal{D}_N^2} \mu_\beta^\alpha. \tag{5.14}$$

Par invariance par rotation et translation du Hamiltonien pour $M = 2$, $E^\alpha_{(X_1,X_2)}$ dépend seulement de $|D|$, où $D := X_2 - X_1$. On écrira $E^\alpha_{|D|}$ au lieu de $E^\alpha_{(X_1,X_2)}$. Dans les deux cas, non-relativiste ou pseudo-relativiste, il n'est pas difficle de voir que $\mu^\alpha = \lim_{|D|\to\infty} E^\alpha_{|D|}$.

Pour un point fixé $X \in \mathbb{R}^3$, qui représentera la position de l'un des noyaux, et pour la variable $x \in \mathbb{R}^{3N}$, on définit la translation unitaire par X dans la variable pour la i-ème particule comme

$$\mathcal{U}_X^{(i)} : \begin{cases} L^2(\mathbb{R}^{3N}) \to L^2(\mathbb{R}^{3N}) \\ \mathcal{U}_X^{(i)}\varphi(x) \mapsto \varphi(x_1, \cdots, x_{i-1}, x_i + X, x_{i+1}, \cdots, x_N). \end{cases} \tag{5.15}$$

Pour $\beta = (\mathcal{C}_1, \mathcal{C}_2) \in \mathcal{D}_N^2$ et X_1, X_2 étant les positions respectives des noyaux, on définit l'opérateur de translation

$$\mathcal{U}_\beta := \prod_{i\in\mathcal{C}_1} \mathcal{U}_{X_1}^{(i)} \prod_{j\in\mathcal{C}_2} \mathcal{U}_{X_2}^{(j)}. \tag{5.16}$$

On pose

$$\tilde{H}_\beta := \mathcal{U}_\beta H_\beta \mathcal{U}_\beta^*. \tag{5.17}$$

On notera que $\tilde{H}_\beta$ est unitairement équivalent à H_β et

$$\tilde{H}_\beta = \tilde{H}_{\mathcal{C}_1}^{Z_1} + \tilde{H}_{\mathcal{C}_2}^{Z_2}.$$

On définit, pour $\beta \in \mathcal{D}_N^2$ les fonctions $f_2, f_3 \in C^\infty(\mathbb{R}^{3N})$ par

$$f_2(x) := \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} -e^2\big(3(x_i\cdot e_D)(x_j\cdot e_D) - x_i\cdot x_j\big), \tag{5.18}$$

$$f_3(x) := \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} \frac{e^2}{2}\Big(3(x_i - x_j)\cdot e_D\big[2(x_i\cdot x_j) - 5(x_i\cdot e_D)(x_j\cdot e_D)\big] \\ + 3|x_i|^2(x_j\cdot e_D) - 3|x_j|^2(x_i\cdot e_D)\Big), \tag{5.19}$$

où $e_D := \frac{D}{|D|}$ est le vecteur unitaire dans la direction de X_1 vers X_2. On notera que les fonctions f_2, f_3 dependent de la décomposition en clusters β.

A partir de maintenant, fixons $\beta \in \mathcal{D}^{at}$. Nous montrerons dans l'Appendice B que μ^α est une valeur propre discrète de H^α_β. Par équivalence unitaire, μ^α est aussi une valeur propre discrète de

$$\tilde{H}^\alpha_\beta := \sum_{\alpha'_\beta \prec \alpha} \tilde{H}^{\alpha'_\beta}_\beta := \sum_{\alpha'_\beta \prec \alpha} \tilde{H}_\beta P^{\alpha'_\beta} \tag{5.20}$$

où la somme est sur toutes les représentations irréductibles induites $\alpha'_\beta \prec \alpha$. On désigne par $\tilde{\mathcal{W}}^\alpha_\beta \subset \mathcal{H}^\alpha$ le sous-espace de $\tilde{H}^\alpha_\beta$ qui correspond à μ^α et on définit

$$a_1(\beta) := \max_{\substack{\phi\in\tilde{\mathcal{W}}^\alpha_\beta\\ \|\phi\|=1}} \|(\tilde{H}_\beta - \mu^\alpha)^{-\frac{1}{2}} f_2\phi\|^2. \tag{5.21}$$

Bien que μ^α soit une valeur propre de $\tilde{H}_\beta$, la valeur $a_1(\beta)$ est bien définie car $f_2\phi$ est orthogonal au sous-espace correspondant (voir Lemma E.2). On définit $\tilde{\mathcal{V}}^\alpha_\beta \subset \tilde{\mathcal{W}}^\alpha_\beta$ comme le sous-espace de tous les ϕ tels que $\|(\tilde{H}_\beta - \mu^\alpha)^{-\frac{1}{2}} f_2\phi\|^2 = a_1(\beta)$ et

$$a_2(\beta) := \max_{\substack{\phi\in\tilde{\mathcal{V}}^\alpha_\beta\\ \|\phi\|=1}} \|(\tilde{H}_\beta - \mu^\alpha)^{-\frac{1}{2}} f_3\phi\|^2. \tag{5.22}$$

De façon similaire, le Lemme E.2 assure aussi que $a_2(\beta)$ est bien défini. Par la symétrie de permutation, pour tout $\beta_1, \beta_2 \in \mathcal{D}^{at}$ on a $a_1(\beta_1) = a_1(\beta_2)$ et $a_2(\beta_1) = a_2(\beta_2)$. Ainsi, on omettra l'argument β dans la définition et on écrira a_1 et a_2 dans la suite. Pour les molécules diatomiques, notre résultat principal est

Theorem 5.2 (Interaction de Van der Waals pour les molécules diatomiques)**.** *Soit α une représentation irréductible de S_N correspondant à un diagramme de Young avec au plus deux colonnes et supposons que les conditions suivantes sont vraies:*
1) Pour tout $\beta \in \mathcal{D}^2_N \setminus \mathcal{D}^{at}$

$$\mu^\alpha_\beta > \mu^\alpha.$$

2) Pour tout $\beta \in \mathcal{D}^{at}$ et toute représentation irréductible α^*_β du groupe S_β avec $\alpha^*_\beta \prec \alpha$ tel que $P^{\alpha^*_\beta}\tilde{\mathcal{W}}^\alpha_\beta \neq \emptyset$,

$$\dim(P^{\alpha^*_\beta}\tilde{\mathcal{W}}^\alpha_\beta) = \dim \alpha^*_\beta.$$

Alors

$$E^\alpha_{|D|} - \mu^\alpha = -\frac{a_1}{|D|^6} - \frac{a_2}{|D|^8} + \mathcal{O}(|D|^{-10}) \tag{5.23}$$

où a_1 and a_2 sont respectivement définis par (5.21) *et* (5.22)*.*

6. Quantum mechanics

In this small introduction to quantum mechanics, we follow the book by Gustafson and Sigal [17] for the fundamentals of quantum mechanics. We will start with the comparatively simple case of one particle. After establishing a foundation, we give an overview of the most important results for the multiparticle case to construct a framework for the results that are proven in our paper, which constitutes the core part of this thesis. Given the format, we can only scratch the surface of the topic. For a more indepth introduction to quantum mechanics with proofs and many useful explanations we refer to the above mentioned book.

As the famous double-slit experiments show, electrons and other elementary particles exhibit wave-like behaviour and do not have a deterministic position, whereas in classical mechanics an object has an exact position and momentum (at least on a relevant scale). In the classical framework we can express measurable quantities such as kinetic energy, potential energies and the likes as functions that depend on deterministic properties of the object such as position, mass or momentum.

To accurately replicate reality, the position of a quantum mechanical particle must be given in terms of a probability distribution.

The *state* of a particle is described by a complex-valued function of position and time $\psi(x,t)$ with $x \in \mathbb{R}^3,\ t \in \mathbb{R}$. We normalize this function such that

$$\int_{\mathbb{R}^3} |\psi(x,t)|^2 dx = 1 \tag{6.1}$$

and interpret $|\psi(\cdot,t)|^2$ as the probability distribution of the particle position at time t. Note that for many appications it is sufficient to require $\psi(\cdot,t) \in L^2(\mathbb{R}^3)$, which implies that the function is *normalizable* instead of it being *normalized* as it is stipulated in (6.1). Furthermore, to justify the name *wave function*, it should be a solution of some wave equation. In our case the Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{6.2}$$

with initial condition

$$\psi\big|_{t=0} = \psi_0 \tag{6.3}$$

where

$$H = T + V \tag{6.4}$$

is the *Hamiltonian*, *Schrödinger* or *energy operator* acting on a Hilbert space $\mathcal{H}$. The first summand in (6.4) describes the kinetic energy of a particle and V is a potential. For our purpose we consider the Hilbert space $L^2(\mathbb{R}^3)$. The operators T and V will be specified later, there are several possibilities and the specific choice depends on which properties are the subject of interest.

6.1. Observables.

The experimentally measurable properties must mathematically be evaluated in a non-deterministic way. That is, each value is observed with a certain probability and never with absolute certainty. So we have to express the classical observables in a probabilistic context, where the values are assessed with respect to the probability distribution $|\psi(\cdot,t)|^2$ for a state ψ. In the Schrödinger setting the development of the system over time is modelled purely in the change of the wave function. Thus we realize these observables as (time independent) operators acting on functions in $L^2(\mathbb{R}^3)$. Furthermore we require *self-adjointness*; by requiring self-adjointness, among many other properties, we ensure that the operator has purely real-valued spectrum (so that they correspond to some measurement in the real world). Such an operator is called a (quantum mechanical) *observable*. The restriction to self-adjoint operators will become clearer as we progress further into the topic.

6.2. **Time evolution.** In this section we talk more about self-adjointness, in particular for the energy of a system which corresponds to the *observable* given by the Schrödinger operator H. If we can not ensure self-adjointness of H, the standard "machinery" of quantum mechanics would fail and the physical interpretations would be questionable at best. We will address the question of self-adjointness for operators of the form (6.4) in Chapter 6.6.

Without going too much into the detail we will motivate some of these fundamental results of quantum mechanics in a non-rigorous way. The canonical solution of a differential equation of the form

$$\frac{d}{dt}f(t) = ikf(t) \tag{6.5}$$

for some $k \in \mathbb{R}$ is $f(t) = e^{ikt}f(0)$. We would like to use a similar approach when we replace k by an operator H in the differential equation (6.2). For a bounded operator B on the Hilbert space $\mathcal{H}$ we can define the exponential e^{itB} via the convergent power series

$$e^{itB} := \sum_{n=0}^{\infty} \frac{(it)^n B^n}{n!}. \tag{6.6}$$

For an *unbounded* operator A (as is the case for our operator H), the convergence of such a power series is not clear. Nonetheless we can formally define the operator e^{itA} via functional calculus with the help of the spectral theorem, see [40, Chapter VII], and Stone's theorem [40, Theorem VIII.8]. These results rely on self-adjointness of H and give rise to the so-called *one parameter unitary group* $U(t)$, in addition they prove a one-to-one correspondence between a self-adjoint operator H and $U(t)$. Furthermore, for all times $t \in \mathbb{R}$ the operator $U(t)$ is unitary, strongly continuous and for a state ψ_0 in the domain of H we can even differentiate $U(t)\psi_0$ with respect to t with $\partial_t U(t)\psi = iH\psi$. This justifies the physical interpretation as a time evolution and has all the properties that we would expect from an operator of the form e^{itH}. In abuse of notation we define $e^{itH} := U(t)$ and we call $\psi(x,t) = U(t)\psi_0(x)$ the *time evolution* of the state ψ_0.

6.3. **Spectrum of an operator.** The spectrum of an operator A is defined as

$$\sigma(A) = \{\lambda \in \mathbb{C} \big| A - \lambda \text{ is not invertible (has no bounded inverse)}\}. \tag{6.7}$$

Note that for a self-adjoint operator A the spectrum lies on the real axis, $\sigma(A) \subset \mathbb{R}$. We say that $\lambda \in \sigma(A)$ is an *isolated eigenvalue* if λ is an eigenvalue such that for some $\epsilon > 0$ we have $(\lambda - \epsilon, \lambda + \epsilon) \cap \sigma(A) = \{\lambda\}$. Typically in spectral analysis of atomic and molecular systems, we distinguish two types of values in the spectrum. We define the *discrete spectrum* as

$$\sigma_{\text{disc}}(A) := \big\{\lambda \in \sigma(A) \big| \lambda \text{ is an isolated eigenvalue with finite multiplicity}\}. \tag{6.8}$$

The *essential spectrum* is defined as

$$\sigma_{\text{ess}}(A) := \sigma(A) \setminus \sigma_{\text{disc}}(A). \tag{6.9}$$

There are several ways to categorize the spectrum. This breakdown into essential and discrete spectrum is the preferred choice for spectral analysis of atoms and molecules, since the essential spectrum is stable under some class of perturbations, unlike other classifications.

We will focus our attention to a Schrödinger operator H with (non-relativistic) kinetic energy operator $-\Delta$ and Coulomb potential V. In this case the potential V is viewed as a perturbation of the kinetic energy operator that consequently does not influence the essential spectrum of the kinetic energy operator. Note that the

kinetic energy operator has purely essential spectrum coinciding with the positive real axis: $\sigma_{\text{ess}}(-\Delta) = [0, \infty)$.

The eigenvalues $E \in \sigma_{\text{disc}}(H)$ correspond to energy levels that are realized by some normalizable function $\psi \in L^2(\mathbb{R}^3)$. These "energy levels" are visualized by the spectral absorption lines of an atom, as they are depicted in highschool level physics books. These absorption spectra correspond to the energy difference between the ground state and the first excited states of an atom.

6.4. **Bound States and Decaying States.** In this section we will discuss the classification into bound states and decaying states. The following definitions are from [17, §6.2] with some additional remarks.

Let us look at solutions of the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi \tag{6.10}$$

with initial condition

$$\psi\big|_{t=0} = \psi_0. \tag{6.11}$$

Assume that $\psi_0 \in \{\text{span of eigenfunctions of } H\}$. Then a solution of (6.10) is given by $\psi(x,t) = U(t)\psi_0(x)$, for the one parameter unitary group $U(t)$ described in Chapter 6.2, formally identified with the exponential $e^{-\frac{i}{\hbar}tH}$. Then by [39, Theorem XI.115] for any $\epsilon > 0$ there is an R such that,

$$\inf_t \int_{|x|\leq R} |\psi(x,t)|^2 dx \geq 1 - \epsilon. \tag{6.12}$$

A state ψ that fulfills (6.12) is called a *bound state*, as it remains essentially localized in space for all times.

Note that for an eigenfunction ϕ_0 of H corresponding to the energy E we have

$$\phi(x,t) = e^{-\frac{i}{\hbar}tH}\phi_0(x) = e^{-\frac{i}{\hbar}tE}\phi_0(x) \tag{6.13}$$

and consequently the probability distribution $|\phi(\cdot,t)|^2$ is constant in time and thus

$$\int_{|x|\geq R} |\phi(x,t)|^2 dx = \int_{|x|\geq R} |\phi_0(x)|^2 dx \to 0 \tag{6.14}$$

as $R \to \infty$.

On the other hand, if $\psi_0 \in \{\text{span of eigenfunctions of } H\}^\perp$ the time evolution $\psi = U(t)\psi_0$ will leave any ball of radius R. To be precise

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \int_{|x|\leq R} |\psi(x,t)|^2 dx\, dt = 0. \tag{6.15}$$

Such states are called *decaying states*, since they describe a particle "moving away to infinity".

The above discussion, based on the Wiener and the RAGE theorem (see e.g.[39, p.340ff]) is associated to a decomposition of the spectrum into point spectrum and continuous spectrum. For models of atoms and molecules, when we are interested in ground states and excited states, as it is the case in the thesis at hand, the natural decomposition is the one into discrete and essential spectrum as discussed in Paragraph 6.3. In this case, any eigenstate associated to a value in the discrete spectrum is in the point spectrum, thus fulfils (6.12) and is a bound state. Indeed, any state with energy in the discrete spectrum enjoys the much stronger property (6.13).

Remark 6.1. *A priori there might be eigenvalues embedded in the essential spectrum. With the methods discussed in this paper it is however not possible to study the eigenvalues in the essential spectrum or decide on their relevance for the physical system.*

6.5. **Fixed nucleus approximation.** The mass of a nucleus in comparison to the mass of an electron is huge. In many cases we can simplify the investigation of a system, by assuming that a nucleus is fixed in place, i.e. the nucleus position is treated as a parameter, since the kinetic energy contribution of the nuclei to the total energy of the system are negligible. The following discussion is a non-rigorous motivation for the fixed nucleus approximation.

In a system of two bodies rotating around each-other, as would be the case for a hydrogen atom, the momenta of the bodies must have the same value. Classically, the kinetic energy of a particle is given by $E_{\text{kin}} = \frac{p^2}{m}$, where m is the mass of the particle and p is it's momentum. For a hydrogen atom modelled classically as a two-body problem this would imply equal momenta for the electron and the nucleus (consisting of one proton). Thus the relation between the kinetic energies is

$$E_{\text{kin}}^{(pr)} m_{pr} = E_{\text{kin}}^{(e)} m_e \tag{6.16}$$

where indices pr and e stand for proton and electron respectively. Since $m_{pr} \approx 2 \cdot 10^3 m_e$, we have the inverse relation for the kinetic energies. For bigger atoms this difference in kinetic energy is even greater. This has two reasons: Firstly, the nucleus of a bigger atom contains neutrons which makes the difference in mass more severe, in fact the table of elements shows that roughly the first 20 elements contain about one neutron for each proton, while larger atoms contain more neutrons than protons.

Secondly, the momentum of a nucleus of an atom would increase at most linearly with the number of electrons. Linear scaling would require that all electrons are clumped at one place and move synchronously. In general the nucleus momentum increases much slower, since electrons tend to "spread out" and the effect that the momenta of electrons have on the nucleus can cancel out.

Since we have fixed the nucleus position we are in a setting, where the mass of the nucleus is mostly irrelevant. By changing the length scale of the system by the factor $\frac{2m_e}{\hbar}$ we can get rid of the prefactor for the kinetic energy of the electron. Let $-e$ denote the charge of an electron. The Hamiltonian of a system with one electron with position vector $x \in \mathbb{R}^3$ interacting via Coulomb potential with a nucleus of charge eZ that has it's position fixed at the origin is given as

$$H_1^Z := -\Delta - \frac{e^2 Z}{|x|}. \tag{6.17}$$

For $Z = 1$ the operator in (6.17) describes the hydrogen atom. For other integer values of Z such a system is called *hydrogen-like*. The subscript in (6.17) indicates the number of electrons in the system, for now we stick to the case of one electron.

The fixed nucleus approximation becomes more important for bigger systems, especially in the case of multiple nuclei that will be discussed in Section 7 where this approximation reduces the complexity of the system significantly.

6.6. **Self-adjointness of the Hamiltonian.** In order to have a well-defined quantum mechanical system we need to ensure that H is self-adjoint. We will not give any detail concerning the proof of self-adjointness for Hamiltonian operators, since the proofs are rather technical and very well known. We shall nonetheless state the famous Kato-Rellich theorem. We use the formulation as in [37, Theorem X.12] to which we refer for the proof. The Kato-Rellich theorem proves self-adjointness for a wide class of relevant operators.

Definition 6.2. *Let A and B be densely defined linear operators on a Hilbert space $\mathcal{H}$ with domains given respectively by $\mathcal{D}(A)$ and $\mathcal{D}(B)$. Suppose that*

- $\mathcal{D}(B) \supset \mathcal{D}(A)$
- *For some a and b in $\mathbb{R}$ and all $\varphi \in \mathcal{D}(A)$,*

$$\|B\varphi\| \leq a\|A\varphi\| + b\|\varphi\|. \tag{6.18}$$

Then B is said to be A-bounded. The infimum of such a is called relative bound of B with respect to A.

Theorem 6.3 (Kato-Rellich theorem)**.** *Suppose that A is self-adjoint, B is symmetric, and B is A-bounded with relative bound $a < 1$. Then $A + B$ is self-adjoint on $\mathcal{D}(A)$ and essentially self-adjoint on any core of A.*

This theorem says that if we perturb the self-adjoint operator A in a sufficiently "gentle" way by the operator B, the resulting operator $A + B$ is essentially self-adjoint. For the one particle case, note that the operator $-\Delta$ is self-adjoint on the Sobolev space $H^2(\mathbb{R}^3)$. The self-adjointness of H_1^Z is obtained from a corollary of the Kato-Rellich theorem:

Corollary 6.4. *Let $V \in L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ be real-valued. Then $-\Delta + V$ is essentially self-adjoint on $L^2(\mathbb{R}^3)$.*

This ensures self-adjointness for a variety of physically interesting potentials. We will not go into detail regarding the physical interpretation of self-adjointness, but we will discuss a physically more concrete problem which is somewhat related: the question of stability. In particular the question of: "What keeps the electrons from falling into the nucleus?" This relation that we refer to is based on the observation that the kinetic energy "controls" the potential energy, which is the fundament of the stability discussed in the next section.

6.7. **Stability of matter.** We have so far discussed two possible scenarios for the time evolution of states describing a one-electron system: bound states, where the electron remains essentially localized for all times and decaying states, where an electron moves away to infinity. But we did not discuss the possibility that an electron comes very close to the nucleus, potentially falling into the nucleus. Such a scenario is called *instability of the first kind.*

Definition 6.5. *A system described by the quantum Hamiltonian H_1^Z acting on $L^2(\mathbb{R}^3)$ exhibits stability of the first kind iff*

$$\inf \sigma(H_1^Z) > -\infty. \tag{6.19}$$

While the kinetic energy operator is non-negative, the Coulomb potential can yield (a priori huge) negative energies for a state. A violation of (6.19) is thus theoretically possible for a state where an electron can be located arbitrarily close to the nucleus with non-vanishing probability. This would cause the energy to be very negative, due to the singularity of the Coulomb potential $V(x) := -\frac{e^2 Z}{|x|}$.

However, the kinetic energy operator controls the (negative) singularity of the Coulomb potential. In [25, Part III] the proof of stability of the first kind for an operator H_1^Z as defined in (6.17) uses a type of Sobolev inequality to prove that

$$\|\nabla\psi\|^2 \geq 3\left(\frac{\pi}{2}\right)^{\frac{3}{4}} \|\varrho(\psi)\|_3 \tag{6.20}$$

where $\varrho(\psi) = |\psi|^2$ is the probability density of the electron position, furthermore

$$\langle \psi, V\psi \rangle = -e^2 Z \int_{\mathbb{R}^3} \frac{\varrho(\psi)(x)}{|x|} dx. \tag{6.21}$$

In particular

$$\langle \psi, H\psi \rangle \geq \inf_{\substack{\rho \geq 0 \\ \|\rho\|_1 = 1}} \left\{ 3 \left(\frac{\pi}{2} \right)^{\frac{3}{4}} \|\rho\|_3 - e^2 Z \int_{\mathbb{R}^3} \frac{\rho(x)}{|x|} dx \right\} \tag{6.22}$$

where we allow any non-negative function with $\|\rho\|_1 = 1$, i.e. not necessarily a probability density obtained from a wave function. Minimization of the right hand side of (6.22) yields

$$\langle \psi, H\psi \rangle \geq -\frac{e^4 Z^2}{3}. \tag{6.23}$$

Now unfortunately this proof is not particularly enlightening with regards to fundamentally understanding why the electron would not fall into the nucleus. Inspired by the heuristic explanation in [25] let us investigate the relation between kinetic energy and Coulomb energy for a family of normalized Gaussian functions

$$\varphi_k(x) := C_k e^{-\frac{|x|^2}{k^2}}. \tag{6.24}$$

The function φ_k roughly has the width k. For the kinetic energy we have

$$\langle \varphi_k, -\Delta \varphi_k \rangle \sim \frac{1}{k^2}$$

and for the Coulomb energy

$$\langle \varphi_k, V\varphi_k \rangle \sim -\frac{e^2 Z}{k}.$$

So in order for the electron to be closer to the origin to lower the Coulomb energy by the factor $\frac{e^2 Z}{k}$ the "price" in kinetic energy gain is around $\frac{1}{k^2}$. The latter will obviously dominate the Coulomb energy for small values of k.

7. Multiparticle systems

So far we have only looked at a wave function describing the behaviour of a single particle. We will extend the model so that we can describe systems with multiple electrons and nuclei. Throughout the thesis we will model nuclei as particles with fixed position, neglecting their kinetic energy in the computation of the total energy of the system. We do not however neglect their electrostatic interaction with other particles. We will introduce a system we can use to describe atoms and molecules. Let us study the case of M nuclei at fixed positions $X_1, \cdots, X_M \in \mathbb{R}^3$ and charge $eZ_1, \cdots eZ_M$ interacting via Coulomb potential with N electrons at positions $x_1, \cdots, x_N \in \mathbb{R}^3$ with identical mass and charge $-e$. We denote by the caligraphic $\mathcal{Z} = (Z_1, \cdots, Z_M)$ the vector containing the atomic numbers of the M nuclei. The Schrödinger operator acting on $L^2(\mathbb{R}^{3N})$ is given by

$$H_N^{\mathcal{Z}} = \sum_{j=1}^{N} \frac{p_j^2}{2} - \sum_{j=1}^{N} \sum_{m=1}^{M} \frac{e^2 Z_m}{|x_j - X_m|} + \sum_{1 \leq i < j \leq N} \frac{e^2}{|x_i - x_j|} + \sum_{1 \leq m < k \leq M} \frac{e^2 Z_m Z_k}{|X_m - X_k|} \tag{7.1}$$

where $p_j = -i\hbar \nabla_{x_j}$ is the momentum of the jth electron. We distinguish here the notation H_N^Z and $H_N^{\mathcal{Z}}$. In the first case we refer to a system of N electrons interacting with a single nucleus with atomic number Z, where the latter Hamiltonian refers to a system with M nuclei with sizes given by $\mathcal{Z} = (Z_1, \cdots, Z_M)$. For a function $\psi \in L^2(\mathbb{R}^{3N})$ the probability density of the electrons being in positions $x_1, \cdots, x_N$ is given by $|\psi|^2$.

One of the first questions that arises is whether such an operator is self-adjoint. In the one particle case the Kato-Rellich theorem and applications thereof yield that for a class of potentials $V = L^2(\mathbb{R}^3) + L^\infty(\mathbb{R}^3)$ (which includes the Coulomb potential), the operator $-\Delta + V$ is (essentially) self-adjoint. The canonical generalization

of the proof to N particle potentials $V = L^2(\mathbb{R}^{3N}) + L^\infty(\mathbb{R}^{3N})$ only includes the case where electron-electron interaction is neglected. For an operator of the form (7.1), we have to show in addition, that the electron-electron interactions are relative bounded with respect to the Laplace operator. This relative bound is equal to zero and the Kato-Rellich theorem yields essential self-adjointness of H_N^Z, defined in (7.1).

Similarly, the relative boundedness of the multiparticle Coulomb potential can be used to extend the stability of the first kind for the one particle case to the N electron case.

The construction of a solution of a multiparticle Schrödinger equation has so far not been achieved. But the analysis of the spectrum of an operator can give us a great deal of information about the system. Some of them have been discussed in Chapters 6.3 and 6.4. The essential spectrum, in particular the bottom of the essential spectrum can for example give us information about the scattering and decaying properties of a system. On the discrete part of the spectrum, comparing the ground state energies for varying fixed nucleus positions allows to determine molecule structures, as the system will generally tend towards an allocation that minimizes overall energy. As we have seen in (6.15), a state which is orthogonal to the eigenspaces corresponds to a decaying state, meaning that at least one particle of the system will move to infinity. Note that there might exist eigenfunctions with eigenvalues embedded in the essential spectrum, the question whether such states are to be considered as stable or unstable is difficult to decide. Theoretically the time evolution of such a state is locally contained. On the other hand a tiny amount of perturbation in the energy could cause such a state to scatter. The answer depends on the spectral structure of a specific system and can not be decided in a general way and is beyond the scope of this thesis.

In the atomic and molecular systems that we consider, the results that allow such conclusions about atomic and molecular structure are the HVZ theorem named after Hunziker, van Winter and Zhislin, the existence of a ground state and the theorems on the number of discrete eigenvalues below the essential spectrum, including Zhislin's theorem and generalizations thereof. Before giving the statements of the respective results we try to describe the statements in words and discuss the implications.

The HVZ theorem has two implications. The first is that the infimum of the essential spectrum $\inf \sigma_{\text{ess}}(H)$ of a system with one nucleus with charge eZ interacting with $N \leq Z$ electrons is given by the lowest energy of the system where $N-1$ electrons are located around the nucleus and one electron moves freely without influence of the other particles. Furthermore it states that for such a system the essential spectrum has the form $\sigma_{\text{ess}}(H) = [\inf \sigma_{\text{ess}}(H), \infty)$. That is, the essential spectrum has no gaps.

The first part of the statement justifies the term *ionization threshold* for the bottom of the essential spectrum. Since a system tends to a state where the overall energy is minimized, an energy above this threshold would cause the system to shed one of the electrons to overall lower the energy and obtain the energy $\Sigma = \inf \sigma_{\text{ess}}(H)$.

The second important theorem is the existence of a ground state. This implies the existence of a minimal energy for a system, which in turn allows the construction of a bound state. Another important class of results is concerned with the number of isolated eigenvalues below the essential spectrum (thus the number of stable excited states).

7.1. **Characterization of the essential spectrum.** We have already discussed the importance of the HVZ theorem and it's implications for the structure of atoms and molecules in a non-rigorous way. In order to give the accurate statement we need to introduce the concept of a *cluster decomposition* and the related concept of *localization*. For the sake of easier notation we will stick to the one nucleus case. Extention for M nuclei is rather straightforward.

The idea of a cluster decomposition is that we devide the set of electron labels $\{1, \cdots, N\}$ into pairwise disjoint subsets $\mathcal{C}_0, \cdots, \mathcal{C}_{k-1}$ such that

$$\bigcup_{l=0}^{k-1} \mathcal{C}_l = \{1, \cdots, N\}. \tag{7.2}$$

Such a subset $\mathcal{C}_l$ is called a *cluster* and the collection $(\mathcal{C}_0, \cdots, \mathcal{C}_{k-1})$ is called a *k cluster decomposition*. Let $\mathcal{D}_N^k$ denote the set of all possible decompositions of $\{1, \cdots, N\}$ into k clusters. The idea of this decomposition is that we allocate a region to each cluster $\mathcal{C}_l$ such that particles in different clusters are far from each-other. In the case of one nucleus with charge eZ located at the origin, the system is described by the Hamiltonian

$$H_N^Z = \sum_{j=1}^{N} \left(\frac{p_j^2}{2} - \frac{e^2 Z}{|x_j|} \right) + \sum_{1 \le i < j \le N} \frac{e^2}{|x_i - x_j|}. \tag{7.3}$$

We consider 2 cluster decompositions $\mathcal{D}_N^2$ and electrons in cluster $\mathcal{C}_1$ are gathered with the nucleus, whereas electrons in cluster $\mathcal{C}_0$ are not associated with a nucleus. For any cluster decomposition $\beta \in \mathcal{D}_N^2$ we define the *intercluster interaction*

$$I_\beta := \sum_{i \in \mathcal{C}_0} -\frac{e^2 Z}{|x_i|} + \sum_{\substack{i \in \mathcal{C}_0 \\ j \in \mathcal{C}_1}} \frac{e^2}{|x_i - x_j|} \tag{7.4}$$

and the *cluster Hamiltonian*

$$H_\beta := H_N^Z - I_\beta. \tag{7.5}$$

Theorem 7.1 (HVZ theorem)**.** *For a cluster decomposition $\beta \in \mathcal{D}_N^2$, define $\mu_\beta := \inf \sigma(H_\beta)$ and $\Sigma := \min_{\substack{\beta \in \mathcal{D}_N^2 \\ \mathcal{C}_0 \neq \emptyset}} \mu_\beta$, then*

$$\sigma_{\text{ess}}(H_N^Z) = [\Sigma, \infty). \tag{7.6}$$

To explain the idea of the proof we will introduce an important tool of multi-particle quantum mechanics.

Definition 7.2. *A set of functions $\mathcal{J} \subset C^1(\mathbb{R}^{3N}; [0,1])$ with*

$$\sum_{J \in \mathcal{J}} J^2(x) = 1 \text{ for all } x \in \mathbb{R}^{3N} \tag{7.7}$$

is called a partition of unity or a partition of one on $\mathbb{R}^{3N}$.

Note that we partition one with respect to the *square* sum of the cutoff functions, such that for $\psi \in L^2(\mathbb{R}^{3N})$ and a partition of unity $\mathcal{J}$ we have

$$\|\psi\|^2 = \langle \psi, \psi \rangle = \langle \sum_{J \in \mathcal{J}} J^2 \psi, \psi \rangle = \sum_{J \in \mathcal{J}} \|J\psi\|^2. \tag{7.8}$$

There exist countless different partitions of unity, as different applications require a specific choice of such a partition. In order to prove the HVZ theorem we choose a partition of unity given by the family of functions $\{J_\beta\}_{\beta \in \mathcal{D}_N^2}$ such that on the support of the function J_β, the distance between particles in different clusters is at

least R for some adjustable parameter R. More explicitly for $x = (x_1, \cdots, x_N) \in \operatorname{supp}(J_\beta)$ we have $|x_i - x_j| > R \quad \forall i \in \mathcal{C}_0,\ j \in \mathcal{C}_1$. Furthermore we require

$$\sum_{\beta \in \mathcal{D}_N^2} J_\beta^2 \equiv 1 \tag{7.9}$$

independently of the choice of R. An example of a partition that satisfies the requirements is a so-called *Ruelle-Simon partition of unity*, see [13, p.32] for the definition and the proof of the existence of such a partition.

We will sketch the idea of the proof for the HVZ theorem in a non-rigorous way. See [13] for the rigorous proof with nonrelativistic kinetic energy, or Chapter A for a rigorous proof for the pseudo-relativistic setting.

The inclusion $\sigma_{\text{ess}}(H_N^Z) \supseteq [\Sigma, \infty)$ can be shown by explicit construction of a Weyl sequence for any $\lambda \in [\Sigma, \infty)$. For this construction there are different approaches, we will present an idea attributed to Hunziker [20], which illustrates some important ideas of spectral analysis in a comprehensible way.

For the minimizing decomposition β yielding the energy Σ, there exists a normalized function $\phi \in C_0^\infty(\mathbb{R}^{3N})$ such that for any $\epsilon > 0$

$$\|(H_\beta - \Sigma)\phi\| \leq \epsilon \tag{7.10}$$

since C_0^∞ lies dense in L^2. We now want to raise the kinetic energy of the state by the (positive) factor $\lambda - \Sigma$. Note that by applying the (unitary) transform $e^{-i\mu x_j}$ to a state, we increase the kinetic energy of the jth electron by the amount μ^2. Since the Coulomb potential is invariant under this unitary transform $e^{-i\mu x_j}$ which is sometimes referred to as a *shift in momentum space*, we have constructed an approximate eigenfunction of H_β corresponding to the value λ.

The second part of the construction is linked to the physical interpretation of the essential spectrum. For an increasing index in the terms of the sequence we will "mimic" the evolution in time of a state where the electrons in cluster $\mathcal{C}_0$ move away to infinity. Simultaneous translation of all electrons in cluster $\mathcal{C}_0$ by some vector $a \in \mathbb{R}^3$ leaves all parts of the total energy unchanged, except for the intercluster interaction I_β. To verify this, note that $-\Delta$ is translation invariant and the electron-electron repulsion depends on the *relative* position of the electrons which is left unchanged.

To recapitulate, by applying a shift in momentum space to ϕ, we obtain a function ψ_0 approximating the energy λ. By then choosing a shift (in position space) that drives the cluster $\mathcal{C}_0$ further and further away, we construct a sequence $(\psi_l)_{l \in \mathbb{N}}$ of approximate eigenfunctions with $\|(H_\beta - \lambda)\psi_l\| \to 0$, while for the intercluster interaction $\|I_\beta \psi_l\| \to 0$ since the distance between the clusters increases. The accurate choice of bigger and bigger shifts ($|a_{l-1}| < |a_l|$) causes the sequence to weakly converge to zero, thus satisfying all condition of Weyl's criterion.

For a rigorous construction following the same idea in a sightly different scenario, see Chapter A. For an alternative proof using wave operators, see [38].

The inclusion $\sigma_{\text{ess}}(H) \subseteq [\Sigma, \infty)$ uses the concept of partitioning the state space by applying a partition of unity, defined in (7.7), to the state space $L^2(\mathbb{R}^{3N})$ (or equivalently to the operators acting on this space) to get

$$\begin{aligned} H_N^Z &= \sum_{\beta \in \mathcal{D}_N^2} J_\beta^2 H_N^Z \\ &= \sum_{\beta \in \mathcal{D}_N^2} J_\beta H_N^Z J_\beta - \sum_{\beta \in \mathcal{D}_N^2} J_\beta (H_N^Z J_\beta - J_\beta H_N^Z). \end{aligned} \tag{7.11}$$

Similarly, replacing $H_N^Z = \sum_{\beta \in \mathcal{D}_N^2} H_N^Z J_\beta^2$ and summing the two resulting equalities we obtain

$$H_N^Z = \sum_{\beta \in \mathcal{D}_N^2} J_\beta H_N^Z J_\beta + \frac{1}{2} \sum_{\beta \in \mathcal{D}_N^2} [J_\beta, [J_\beta, H_N^Z]] \tag{7.12}$$

where the second term on right hand side is called the *localization error*. Note that the function J_β commutes with the Coulomb potential and thus

$$[J_\beta, [J_\beta, H_N^Z]] = [J_\beta, [J_\beta, -\Delta]] = -2|\nabla J_\beta|^2. \tag{7.13}$$

The decomposition in (7.12) together with the identity (7.13) is called *IMS localization formula* named after Ismagilov, Morgan and Simon. For details on the origin and use of the formula see [13, p.28] and references therein.

We aim to show that the "perturbations" $J_\beta^2 I_\beta$ and $\sum_{j=1}^N (\nabla_j J_\beta)^2$ of the operator H_N^Z are sufficiently "gentle" so that they leave the essential spectrum invariant, which then implies that

$$\sigma_{\text{ess}}(H_N^Z) = \sigma_{\text{ess}}(H_\beta) \subseteq [\Sigma, \infty). \tag{7.14}$$

The operator $J_\beta^2 I_\beta$ is "small", since by the scaling of the Coulomb potential and the scaling parameter R in the cutoff functions, we have $J_\beta^2 I_\beta \sim R^{-1}$.

The localization error, determined by the absolute value of ∇J_β depends on the choice of the cutoff. This dependence of the error term is taken into account in the choice of the cutoff. Generally the cutoff function can be chosen such that $|\nabla J_\beta|^2_{\max} \sim CR^{-2}$ for some constant $C > 0$. Consequently, by letting R tend to infinity we prove the second inclusion as it is stated in (7.14).

These two inclusions yield equality in (7.6) and prove the theorem. Note that there exist several slightly different formulations of the HVZ theorem. One notable formulation is the one, where the value Σ marking the infimum of $\sigma_{\text{ess}}(H_N^Z)$ is explicitly proven to be the ground state energy of the Hamiltonian H_{N-1}^Z, describing a nucleus with the same atomic number interacting with $N-1$ electrons. In this excursus on the proof of the HVZ theorem, the partition of unity has been formulated as a rather vague object. In the next chapter we will try to establish a more demonstrative view on a partition of the state space.

7.2. **Partition of the state space, or localization.** The partitioning of the state space is one of the main tools that we use to obtain bounds on the expected value of the total energy. For example, if we were to estimate the total energy of two atoms at a "reasonably large" distance one to the other (in the sense that they would not be considered to form a molecule). From a physical perspective we expect that the interactions between electrons in different atoms should be much weaker than the interactions of particles within the same atom. Mathematically, this is a priori not clear: Even though we say that the electron "moves around the nucleus", quantum mechanics is not deterministic and even for a bound state an electron can be at any location in space with nonzero probability. Since the Coulomb potential for two electrons has a singularity, even a very small probability of two particles being very close to each-other might cause problems. This is where localization by a partition of unity comes in handy.

Let us consider an arbitrary state $\psi \in L^2(\mathbb{R}^6)$ describing two electrons moving in the potential of a nucleus of charge $2e$ located at the origin. A possibility to estimate the energy of this state is to use a localization function $u \in C^1(\mathbb{R}^3; [0,1])$ with

$$u(z) := \begin{cases} 1 & \text{if } |z| \leq R \\ 0 & \text{if } |z| > 2R \end{cases} \tag{7.15}$$

together with a function $v \in C^1(\mathbb{R}^3; [0,1])$, such that $u^2 + v^2 \equiv 1$. With these two functions we can define a partition of unity on $\mathbb{R}^6$ consisting of the four functions

$$\begin{aligned} J_1(x_1, x_2) &:= u(x_1)u(x_2) \\ J_2(x_1, x_2) &:= u(x_1)v(x_2) \\ J_3(x_1, x_2) &:= v(x_1)u(x_2) \\ J_4(x_1, x_2) &:= v(x_1)v(x_2). \end{aligned} \tag{7.16}$$

Obviously $\sum J_i^2 = 1$ and we can use these functions to decompose any function in the state space $L^2(\mathbb{R}^6)$ by noting that

$$\|\psi\|^2 = \langle \sum_{i=1}^{4} J_i^2 \psi, \psi \rangle = \sum_{i=1}^{4} \langle J_i \psi, J_i \psi \rangle = \sum_{i=1}^{4} \|J_i \psi\|^2. \tag{7.17}$$

Very crudely speaking, $J_1\psi$ is the part of the state ψ where both electrons are close to the nucleus. We would thus expect the energy of this part somewhere in the discrete spectrum of a helium atom. The parts $J_2\psi$ and $J_3\psi$ would be expected to have the energy within the spectrum of a He^+ ion.

Let us look at the Hamiltonian H_2^2 that describes the energy of such a system. The expected energy of a state $\psi \in L^2(\mathbb{R}^6)$ is given by $\langle \psi, H_2^2 \psi \rangle$. Using the partition of unity defined above, we get

$$\langle \psi, H_2^2 \psi \rangle = \sum_{i=1}^{4} \langle J_i \psi, H_2^2 J_i \psi \rangle + \sum_{i=1}^{4} |\nabla J_i|^2. \tag{7.18}$$

As discussed above the localization yields some information, we know roughly what the expected energy values of $J_i\psi$ should be. But the decomposition comes with the introduction of a positive term, the localization error. As a rule of thumb, the more information the localization yields, the bigger is the localization error.

7.3. **The discrete eigenvalues.** For the structure of atoms and molecules, the discrete spectrum coincides with the energy levels of the physically relevant states. The investigation of the discrete spectrum is concerned with determining the number of discrete eigenvalues, their location and the dimension of their corresponding eigenspaces. Of particular interest is the lowest eigenvalue of the Hamiltonian, the ground state energy.

Let us look at the Hamiltonian H_N^Z defined as in (7.3) describing a nucleus with atomic number Z located at the origin interacting with N electrons.

Generally there are three main cases that are distinguished. An infinite number of eigenvalues in $\sigma_{\text{disc}}(H_N^Z)$, a finite number of eigenvalues in the discrete spectrum, and the case where the discrete spectrum is empty. There are examples for all cases.

For atoms, that is for $N = Z$, the discrete spectrum of the operator H_Z^Z is infinite. For the fixed nucleus case that we consider this result can be obtained from *Zhislin's theorem*, where the kinetic energy of the nucleus is considered, see [38, Theorem XIII.7], the statement is adapted to the notation of this introduction.

Theorem 7.3 (Zhislin's theorem)**.** *Let $H^Z(M)$ be an "atomic" Hamiltonian of the form*

$$H^Z(M) = \sum_{j=1}^{Z} \left(-\frac{\Delta_j}{2m} - \frac{Z}{|x_j|} \right) + \sum_{1 \le i < j \le Z} \left(\frac{\nabla_i \cdot \nabla_j}{M} + \frac{1}{|x_i - x_j|} \right) \tag{7.19}$$

where m is the mass of an electron and M the mass of the nucleus in a system, where the center of mass motion is removed.
Then $\sigma_{\text{disc}}(H^Z(M))$ is infinite.

Remark 7.4. *This implies the infiniteness of $\sigma_{\text{disc}}(H_Z^Z)$, by letting $M \to \infty$ and rescaling the system to the atomic units that we use in the definition of the operator H_Z^Z in (7.3).*

Intuitively, if we put more and more electrons close to a nucleus, at some point the repulsion of the electrons should outscale the attraction of the nucleus, causing one electron to scatter to infinity. This suggests the existence of $N_{\max}(Z)$, the maximal number of electrons that a nucleus of atomic number Z can bind before one electron is hurled to infinity. The Ruskai-Sigal theorem, see [13, Theorem 3.15] proves that this critical value $N_{\max}(Z)$ exists in the sense that for $N > N_{\max}(Z)$ the spectrum of H_N^Z contains no discrete spectrum. The Ruskai-Sigal theorem also gives an upper bound

$$\limsup_{Z\to\infty} \frac{N_{\max}(Z)}{Z} \leq 2. \tag{7.20}$$

This bound is not particularly significant for the physical cases. For details on the refinements of this bound see the discussion after [13, Theorem 3.15] and references therein. The refinement due to Lieb that yields $N_{\max}(Z) < 2Z + 1$ yields the experimentally expected critival value only for hydrogen. In nature the negativity of ions does generally not exceed 2. The theoretical results that estimate $N_{\max}(Z)$ in a rigorous way are all far from the anticipated value. According to [33] the *ionization conjecture*, which can be stated as $N_{\max}(Z) \leq Z+1$ or $N_{\max}(Z) \leq Z+2$, for all Z has not been proven, except for the case of $Z = 1$, as we have discusse above.

The following collection of results is based on the review by Nam, see [33]. It is known that the Pauli exclusion principle is a necessary condition for the ionization conjecture, since for "bosonic" electrons there exists a lower bound by Benguria and Lieb [10] which shows, that the conjecture could not be true for bosonic matter. The asymptotic neutrality of atoms

$$\limsup_{Z\to\infty} \frac{N_{max}(Z)}{Z} = 1 \tag{7.21}$$

has been proven by Lieb, Sigal, Simon and Thirring [26]. Feffernam, Seco [15] and Seco, Sigal, Solovej [41] show that

$$N_{max}(Z) \leq Z + CZ^{\frac{5}{7}}$$

which is the best-known bound for large atoms. For "medium" atoms with $Z \geq 6$ the bound

$$N_{max}(Z) < 1.22Z + 3Z^{\frac{1}{3}}$$

by Nam [34] is an improvement to the bound by Lieb.

The general number of eigenvalues in the discrete spectrum of H_N^Z for positive and negative ions of overall charge $e(Z - N)$ is discussed in [13, §3.9] and exceeds the context of this introduction. The existence of discrete spectrum in the cases $N \leq Z$ can be answered by a rather straightforward construction shown in the next chapter.

7.4. **Energy below the essential spectrum.** In this section we will construct a state with energy below the essential spectrum of H_N^Z for atoms and positive ions ($N \leq Z$). This implies that the discrete spectrum $\sigma_{\text{disc}}(H_N^Z)$ is not empty and in particular, since the Hamiltonian H_N^Z is semi-bounded below the existence of

a ground state energy and finite dimensionality of the corresponding ground state space.

By an alternative formulation of the HVZ theorem, which reveals slightly more information about the bottom of the essential spectrum of a Hamiltonian with $N \leq Z$ we get that

$$\inf \sigma_{\text{ess}}(H_N^Z) = \inf \sigma(H_{N-1}^Z).$$

We use this equation to prove inductively the existence of a ground state for Hamiltonians describing non-negative ions H_k^Z, for $k = 1, \cdots, Z$. Based on a ground state of H_k^Z, $k < Z$ in $L^2(\mathbb{R}^{3k})$ we will construct a product state in $L^2(\mathbb{R}^{3(k+1)})$ such that the energy of the new state is strictly smaller, thus yielding the existence of an eigenvalue below the essential spectrum.

Theorem 7.5. *Let H_N^Z be defined as in* (7.3)*, then for any $N \leq Z$ the discrete spectrum is not empty.*

Again, we will not reproduce a rigorous proof, but try to explain the general approach in a comprehensible way. Note that this theorem is an immediate consequence of the finiteness or infiniteness of the discrete spectrum, for a complete and rigorous proof thereof the reader can see e.g. [13, §3.9].

Let us define

$$\mu_k^Z := \inf \sigma(H_k^Z) \quad \text{and} \quad \Sigma_k^Z := \inf \sigma_{\text{ess}}(H_k^Z).$$

For $k = 1$ the discrete spectrum is infinite, see for example [38, Theorem XIII.6]. Assume that the assertion has been proven for arbitrary but fixed $k < Z$. We pick an approximate normalized eigenfunction $\phi(x_1, \cdots, x_k) \in C_0^\infty(\mathbb{R}^{3k})$, supported on a ball of radius δR, $\delta < 1$ around the origin, corresponding to the eigenvalue μ_k^Z with

$$\langle \phi, (H_k^Z - \mu_k^Z)\phi \rangle \leq d(\delta R) \tag{7.22}$$

for some function $d(\delta R)$. We will later see that for some fixed δ, the function $d(\delta R)$ is sufficiently small for big R. Notice that

$$H_{k+1}^Z = H_k^Z - \frac{\Delta_{k+1}}{2} - \frac{e^2 Z}{|x_{k+1}|} + \sum_{j=1}^{k} \frac{e^2}{|x_j - x_{k+1}|}. \tag{7.23}$$

We aim to find a function that is small regarding the expected value of the last three terms in (7.23). For this purpose we take a normalized function $f \in C_0^\infty(\mathbb{R}^3)$ that is supported in the region where $1 < |x_{k+1}| \leq 2$ (imagine a smooth "bump" on $[1, 2]$ for example). We define a function similar to f, that is spread to the interval $R < |x_{k+1}| \leq 2R$ and has the same norm as f:

$$f_R(x_{k+1}) := R^{-\frac{3}{2}} f\left(\frac{x_{k+1}}{R}\right). \tag{7.24}$$

Note that similar to the example of the Gaussian functions in Chapter 6.7, the behaviour of the kinetic energy of f_R is

$$\langle f_R, -\Delta_{k+1} f_R \rangle \sim CR^{-2}.$$

In the next step we will use the positivity of the ion described by H_k^Z and ϕ. Since $k < Z$ the part of the potential concerning the $(k+1)$st electron is overall negative of order $\sim -cR^{-1}$ under the condition that $\delta < \frac{1}{Z}$. Note that this is only true, since both ϕ and f_R are compactly supported on regions with minimal distance that scales linearly in R. The details to obtain such a bound rigorously are omitted here, since they are given in Chapter B.

By the above discussion, the product state

$$\psi_R(x_1, \cdots, x_{k+1}) := \phi(x_1, \cdots, x_k) f_R(x_{k+1}) \tag{7.25}$$

yields an energy

$$\langle \psi_R, (H_{k+1}^Z - \mu_k^Z)\psi_R \rangle \lesssim d(\delta R) + \frac{C}{R^2} - \frac{c}{R}. \tag{7.26}$$

We are left to show that $d(\delta R)$ is dominated by the negative term $-\frac{c}{R}$ for large R which depends on the asymptotic behaviour of bound states that we will discuss in the next chapter. We will see that $d(\delta R)$ is in fact exponentially small in R.

7.5. **Asymptotic behaviour of bound states.** In the discussion of the spectrum we have established, that for eigenfunctions corresponding to the energies in the discrete spectrum, the electrons are very likely found in some region around the nucleus where the potential energies are relatively small. It is highly unlikely that they are found very far from the nucleus, since states where one electron is sent far away will yield an energy in the essential spectrum as we have seen in the construction of a trial state for the HVZ theorem. A result which shows that the probability density of a bound state is small far from the nucleus would be in accordance with the physical world, where we do not observe electrons leaving and re-entering the surrounding of the nucleus on a regular basis without external influence. Such a result for bound states yields in turn useful information about scattering states as a counterpart. Furthermore it allows to estimate interactions between atoms more accurately, given that the atoms are in a low energy state.

On the field of asymptotic behaviour of bound states for multiparticle Schrödinger operators, the most famous results are due to Agmon. Concerning the case discussed in this introduction, see [2, Theorem 6.4]. We give the statement similar to the way it is presented in [16].

Theorem 7.6 (Agmon)**.** *Let H_N^Z be the N-body Hamiltonian in* (7.3) *and let ϕ be an eigenfunction corresonding to the eigenvalue $\lambda < \Sigma_N^Z = \inf \sigma_{\text{ess}}(H_N^Z)$ and choose $\beta > 0$ with $\beta^2 < \Sigma_N^Z - \lambda$. Then*

$$\|e^{-\beta^2|x|}\phi\|^2 \leq C \left(\frac{1}{\Sigma_N^Z - \lambda - \beta^2} \right) \|\phi\|^2 \tag{7.27}$$

for some constant C.

Note that the (7.27) suggests that bound states with low energy decay faster than states with energy close to the essential spectrum.

We have now given an overview of the most important results for the multiparticle Schrödinger operator H_N^Z. In the next part of the introduction we will discuss the case where the kinetic energy operator $-\Delta$ is replaced by an operator that accounts for the change of mass for velocities approaching the speed of light.

7.6. **Relativistic kinetic energy operator.** So far we have modelled the kinetic energy for electrons by the operator $-\Delta$ deduced from Newtonian mechanics. The famous equation by Einstein for the total energy of a particle

$$E_{\text{tot}} = \sqrt{p^2c^2 + m^2c^4} \tag{7.28}$$

postulates the increase of mass m for an object moving at velocities close to the speed of light c. By subtracting the energy of a particle at rest $E_0 = mc^2$ we obtain the kinetic portion of the total energy, or in the context of distributions and expected values, the *relativistic kinetic energy operator*

$$T_{\text{rel}}(p) := \sqrt{p^2c^2 + m^2c^4} - mc^2. \tag{7.29}$$

This relativistic correction becomes relevant in the context of larger atoms, where the innermost electrons have high kinetic energy, see [25, Part VII]. For small momenta p, the operator T_{rel} behaves similar to the nonrelativistic case $\frac{p^2}{2m}$, whereas

for large momenta the energy scales like $|p|$. By Newton's equation of motion the velocity v of a particle with kinetic energy $T(p)$ is given by

$$v = \frac{d}{dp}T(p) \tag{7.30}$$

and by plugging in the kinetic energy operator $T_{\text{rel}}(p)$ defined in (7.29) we get

$$v_{\text{rel}} = \frac{pc^2}{\sqrt{p^2c^2 + m^2c^4}}. \tag{7.31}$$

We can see that $v_{\text{rel}} \xrightarrow{p\to\infty} c$, and the speed of light is an upper bound for the velocity in accordance with special relativity. Note that this is *not* a relativistic theory of quantum mechanics, the implementation of special relativity into the theory of quantum mechanics is still one of the big unsolved problems of theoretical physics.

We will focus on the purely mathematical problems that arise if we replace the operator $\frac{p^2}{2}$ by the operator $\sqrt{p^2c^2 + m^2c^4} - mc^2$, we will refer to models with the latter kinetic energy operator as a *pseudo-relativistic* model or system. Similar to the nonrelativistic case we can rescale the system, such that in atomic units we have $T_{\text{rel}} = \sqrt{p^2+1} - 1$. We will denote in this chapter by $\mathfrak{H}_N^Z$ the pseudo-relativistic equivalent to the Hamiltonian H_N^Z defined in (7.3). More explicitly

$$\mathfrak{H}_N^Z := \sum_{j=1}^{N}\left(\sqrt{p_j^2+1} - 1 - \frac{e^2Z}{|x_j|}\right) + \sum_{1\le i<j\le N}\frac{e^2}{|x_i - x_j|}. \tag{7.32}$$

The results that were introduced for the Laplacian in the previous subsections of the introduction can not be used directly for the operator $\mathfrak{H}_N^Z$. That includes the proof of (essential) self-adjointness and the results on the spectral structure. The proof of self-adjointness was based on the relative boundedness of the Coulomb potential with respect to the Laplacian. On the fundamental level of the analysis of the operator there are two crucial differences between the nonrelativistic and the relativistic operator.

The first being that T_{rel} is not a local operator. This means that in order to obtain the value of $T_{\text{rel}}\psi$ at some point $x \in \mathbb{R}^3$ it is not sufficient to know the values in a small neighborhood around x, as is the case for the Laplacian. A priori this means, that manipulations of a state, such as localization for example, might have a global effect on the expected energy of the state. This is in particular relevant as in the proofs for the HVZ theorem and the existence of a ground state we could modify the state "far away" from the nucleus without changing the energy of the state close to the nucleus.

Secondly, if take a function on some compact support and spread the function similar to the construction in (7.24), instead of a square inverse scaling in the nonrelativistic case we get a linear inverse scaling in the relativistic case. Compare also the example of the contracted Gaussian function that we used to illustrate the stability of the first kind of the Hamiltonian in Chapter 6.7. Consequently the scaling of the relativistic kinetic energy operator is of the same order as the Coulomb potential. In that sense we can not expect the kinetic energy to control the singularity of the Coulomb potential without constraints.

In the literature the pseudo-relativistic Hamiltonian $\mathfrak{H}_N^Z$ is often called the *Herbst* Hamiltonian. Herbst was one of the first to study this operator, in [19] he shows that the Hamiltonian $\mathfrak{H}_1^Z$ describing a hydrogen-like atom, can be self-adjointly realized on the fractional Sobolev space $H^{1/2}(\mathbb{R}^3)$ if the atomic number of the nucleus satisfies $Z \le \frac{2}{\alpha\pi} \approx 87$, where α is the *Sommerfeld fine structure constant*. Note that the same critical value appears, when the stability of a multiparticle Hamiltonian

$\mathfrak{H}_N^{\mathcal{Z}}$ is discussed by Lieb and Yau [28]. They prove that the Hamiltonian of the form $\mathfrak{H}_N^{\mathcal{Z}}$ is *stable* if and only if $Z_i \leq \frac{2}{\alpha\pi}$ for all $i = 1, \cdots, M$. Here, stability implies in particular stability of the first kind discussed in Chapter 6.7.

A substantial part of the paper is the estimate of the localization error (compare (7.12) for the relativistic case where we use the IMS localization formula). We can use a very similar approach, but the resulting localization error will be different. Generally, the localization error for a Hamiltonian with kinetic energy operator T and multiplicative potential V for the partition of unity $\{J_\beta\}$ is given by

$$\sum_\beta J_\beta[T, J_\beta]. \tag{7.33}$$

Note that the localization error is an operator, however the terminology is used interchangably for both the operator and

$$\sum_\beta \langle \psi, J_\beta[T, J_\beta]\psi\rangle$$

which should accurately be called the expected localization error energy of the state ψ. Setting $T = -\Delta$ in (7.33) gives rise to the operator $|\nabla J_\beta|^2$ which is supported in the region where $0 < J_\beta < 1$. By locality of the Laplacian the expected energy increase due to the cutoff depends on the probability density of the state in the region of the cutoff. This is *not* the case for the relativistic kinetic energy T_{rel} that gives rise to a *globally supported* localization error operator. Estimating the localization error (energy) thus becomes a more difficult task.

Many of the results regarding the spectral structure of the Schrödinger Hamiltonian (with nonrelativistic kinetic energy) that we have mentioned and discussed in the introduction will be proven over the course of this thesis for a specific class of physically relevant states. For the operator $\mathfrak{H}_N^{\mathcal{Z}}$ where we neglect the fact that electrons are identical particles and subject to the Pauli exclusion principle these results are known and will be discussed in a later section. The restriction to physically relevant states is discussed shortly in the next part and the extention of the known results to this physically relevant setting is tackled in the core part of the thesis.

7.7. **Symmetry considerations.** A defining part of this thesis is the consideration of a restriction of the state space to subspaces that are in accord with the Pauli exclusion principle, which postulates that two electrons can never be simultaneous in the same quantum mechanical state. The quantum mechanical formulation of the exclusion principle gives a mathematical condition that allows us to exclude the physically impossible states from our considerations. Chapter G.3 contains more detail on permutation symmetry and an introduction to representation theory for it's application in mathematical physics and spectral analysis.

At this point it shall suffice to state that in order to have a model that is in better accordance with reality, we restrict the state space to certain mutually orthogonal subspaces that can be classified by their permutational symmetric properties. We will consider the reduced Hamiltonians for each of these physically relevant subspaces separately. This will give us complete information of the spectrum for the collection of physically relevant states. Depending on the structure of the spectrum, this approach can reveal more information about the spectral properties of the system. Recall that exponential decay of bound states does not apply to eigenvalues embedded in the essential spectrum, furthermore we do not get finite dimensionality of the corresponding eigenspace for embedded eigenvalues. In general, the restriction of the state space will yield bigger values for the spectrum. In particular, the bottom of the essential spectrum will be further to the right on the real

line. This process could thus "unembed" (or isolate) certain eigenvalues of finite multiplicity and show that the corresponding eigenfunctions enjoy the properties shown for states with energy in the discrete spectrum, like exponential decay.

In [50] it is stated that neglecting the electron-electron Coulomb interaction in H_N^Z and for $N \geq 4$ "all the eigenvalues corresponding to a realizable symmetry lie in the continuous spectrum. One would expect this is also true for the complete operator H". In this case, the "realizable symmetry" refers to a states that satisfies the Pauli exclusion principle and "the continous spectrum" in which these values are embedded refers to the spectrum of the operator on the *full* state space. Studying the embedded eigenvalues is not possible with the usual tools of spectral analysis, which emphasizes the relevance of this decomposition of the state space.

In the next part we would like to introduce the problem that motivates the core of this thesis. The attempt to study the van der Waals-London dispersion in the context of the pseudo-relativistic operator and state space restrictions according to the Pauli principle, have revealed that some of the main results in spectral analysis of multiparticle quantum mechanics have not been proven for this specific case, or have been announced without proof. We generally use established methods adapted to the specific requirements of the pseudo-relativistic operator and the permutational symmetry considerations. Some of these methods have already been explained in Chapter 7 in a nonrigorous way. These explanations are supposed to serve as a guiding principle for the case that a rather complitated notation which is intrinsic to this type multiparticle quantum mechanics obstructs the view on the essentials.

8. The Van der Waals - London dispersion

The van der Waals force plays a vital role in many natural phenomena. It is a key aspect of understanding molecular bindings. To give a few examples, the van der Waals force is needed to explain the condensation of water from vapor, the structural stability of DNA, and the binding between several layers of graphene to form graphite.

The importance of the van der Waals force is not restricted to the microscopic scale. The van der Waals forces are used to explain some biological processes and there are efforts in nanotechnology to take advantage of this attractive force. For further examples, see the introductory discussion in [3].

Let us note at this point, that there is a certain degree of ambiguity in the term "van der Waals force". Several interatomic and -molecular interactions caused by different effects are collectively called "van der Waals force". This includes interactions of atoms and molecules with permanent mono- and multipoles and the repulsion caused by interatomic electron-electron interactions. In this context we will talk about the so-called *van der Waals-London (vdW-L) dispersion*. An interatomic and -molecular interaction even in the absence of permanent polarity. The physical discription of the phenomenon in the example of two atoms is that the movements of the electrons in different atoms will synchronize, thus creating "quasi-dipoles", sometimes referred to as *temporary*, *non-permament* or *induced dipoles* that interact with each-other. In order to stick to the usual terminology, we will abusively call this interaction energy the vdW-L *force*. Note that the vdW-L force is also present for molecules with permanent multipoles.

Surprisingly enough there are only few mathematically rigorous results concerning the van der Waals force. In [29], J. D. Morgan and B. Simon obtain an asymptotic expansion of the interaction energy and notice that for atoms with the absence of dipole and quadrupole momentum, the leading behaviour is $\mathcal{O}(|D|^{-6})$, where $|D|$ is the distance between two nuclei.

Another result concerning van der Waals interaction is obtained in [27] by E. H. Lieb and W. E. Thirring where they constructed a trial function to show that attractive energy between two atoms without permanent polarity is at least $-C|D|^{-6}$ for some positive constant C.

This result was improved by I. Anapolitanos and I. M. Sigal in [3], who obtain under some restrictions, which we will discuss later, the leading term of order $|D|^{-6}$ for the intercluster energy in the nonrelativistic case with an error $\mathcal{O}(|D|^{-7})$.

Note that all previous results only hold for nonrelativistic kinetic energies. For heavy atoms one should include relativistic effects for the electrons. This is the main goal of the work at hand. The method of the proof developed in [3] is based in its essential part on the Feshbach-Schur method. Our approach is purely variational and similar to the one used in [6, 7, 8, 9, 43, 44, 45] to obtain asymptotics of eigenvalues of multiparticle Schrödinger operators near the bottom of the essential spectrum.

Let us explain the idea of the approach for a diatomic system in a heuristic language in a diatomic setting. We will try to gain information about the discrete spectrum of the diatomic (interacting) system, by looking at the discrete spectrum of the same, but isolated (non-interacting) atoms. We then estimate the interatomic interaction by comparing the lowest energy of the isolated system to the minimal overall energy for the interacting system. We will find that the lowest energy for the interacting diatomic system can be expressed in terms of an asymptotic expansion in negative powers of $|D|$ "around" the ground state energy of the isolated system.

In comparison to [3], we get higher order terms in the asymptotic expansion and better estimates for the remainder terms in both, the nonrelativistic and the pseudo-relativistic case. Furthermore, our method yields an explicit expression for the coefficients for terms up to the order $|D|^{-9}$ in the intercluster energy expansion. In addition, we have obtained a remarkable difference in the asymptotic expansion for a system with more than two nuclei. While in the case of two nuclei the coefficient of order $|D|^{-9}$ is zero, for three or more nuclei it is a priori not. See the discussion in the remarks after Theorem 9.4.

9. The model, notations and main results

We consider a molecule with N electrons of charge $-e$ and spin $\frac{1}{2}$, and M pointwise nuclei with charges eZ_l located at positions X_l in $\mathbb{R}^3$, which we suppose to be fixed (Born-Oppenheimer approximation). We assume that the system is neutral, which means that $\sum_{l=1}^{M} Z_l = N$. The corresponding Hamiltonian is

$$H := \sum_{i=1}^{N}\left(T_i - \sum_{l=1}^{M}\frac{e^2 Z_l}{|x_i - X_l|}\right) + \sum_{1\le i<j\le N}\frac{e^2}{|x_i - x_j|} + \sum_{1\le k<l\le M}\frac{e^2 Z_k Z_l}{|X_k - X_l|} \tag{9.1}$$

with k-th electron kinetic energy operator

$$T_k := \begin{cases} \sqrt{p_k^2+1} - 1 & \text{in the pseudo-relativistic case} \\ \frac{p_k^2}{2} & \text{in the nonrelativistic case} \end{cases} \tag{9.2}$$

and form domain $H^{1/2}(\mathbb{R}^{3N})$ in the pseudo-relativistic case and $H^1(\mathbb{R}^{3N})$ in the nonrelativistic case. As usual $p_k = -i\nabla_{x_k}$ denotes the momentum of the k-th electron. If T_k is pseudo-relativistic, we assumed $Z_l e^2 \le \frac{2}{\pi}$, which ensures that the Hamiltonian is semi-bounded from below, see [14, 28].

In the main part of the paper we will focus on the pseudo-relativistic kinetic energy case $T_k = \sqrt{p_k^2+1} - 1$ (see [19] and references therein) although all the results hold for $T_k = \frac{p_k^2}{2}$ likewise. Here the Hamiltonian is written in atomic units, i.e. $c = \hbar = m = 1$.

The phase space for a system of N electrons, taking into account the Pauli-principle, is the antisymmetric tensor product of N copies of $L^2(\mathbb{R}^3;\mathbb{C}^2)$, namely the space $\bigwedge^N L^2(\mathbb{R}^3;\mathbb{C}^2)$ of functions in $\bigotimes^N L^2(\mathbb{R}^3;\mathbb{C}^2)$ that are antisymmetric with respect to transpositions of pairs of position and spin particle variables (x_i, s_i) and (x_j, s_j), for $i \neq j$.

The operator H we consider does only depend on the coordinate variables x_i, but not on spin variables s_i. Hence we consider H to act on the projection of $\bigwedge^N L^2(\mathbb{R}^3;\mathbb{C}^2)$ onto the space of functions depending on coordinates alone, that is, on the space $\mathcal{H}_{\text{Fermi}}$ defined by

$$\mathcal{H}_{\text{Fermi}} := \Big\{ \langle \mathfrak{s}, \Psi \rangle_{\text{spin}} | \Psi \in \bigwedge^N L^2(\mathbb{R}^3;\mathbb{C}^2), \mathfrak{s} : \Big\{ -\frac{1}{2}, \frac{1}{2} \Big\}^N \to \mathbb{C} \Big\} \tag{9.3}$$

where

$$\langle \mathfrak{s}, \Psi \rangle_{\text{spin}} := \sum_s \bar{\mathfrak{s}}(s_1, \cdots, s_N) \Psi(x_1, s_1, \cdots, x_N, s_N).$$

Note that $\mathcal{H}_{\text{Fermi}}$ is a subspace of $L^2(\mathbb{R}^{3N})$.

The condition of antisymmetry with respect to transposition of the particle variables implies certain symmetry properties for permutations of coordinate variables after decoupling of the spin variables. Namely, permutations of electrons transform the functions according to a Young pattern with at most two columns as described in [18, § 7.3.]. Note that for more than two particles a function which is completely symmetric under transposition of coordinate variables can never be antisymmetric under transposition of the full particle variables, since the spin can only attain two values.

More precisely, let S_N be the group of permutations of N electrons. For any $\pi \in S_N$ let $\mathcal{T}_\pi : \mathcal{H}_{\text{Fermi}} \to \mathcal{H}_{\text{Fermi}}$ with

$$\mathcal{T}_\pi \psi(x_1, \cdots, x_N) := \psi(x_{\pi^{-1}(1)}, \cdots, x_{\pi^{-1}(N)}) \tag{9.4}$$

be the operator that realizes a permutation on the particle variables.

Let α be an irreducible representation of the group S_N and P^α the projection on the subspace of functions transformed under the action of operators $\mathcal{T}_\pi$ according to the representation α. These projections decompose the space $\mathcal{H}_{\text{Fermi}}$ into a finite number of orthogonal subspaces $\mathcal{H}^\alpha := P^\alpha \mathcal{H}_{\text{Fermi}}$ such that

$$\mathcal{H}_{\text{Fermi}} = \bigoplus_{\alpha \in \mathcal{A}} \mathcal{H}^\alpha, \tag{9.5}$$

where $\mathcal{A}$ is the set of all irreducible representations of the group S_N corresponding to a Young pattern with at most two columns. Note that for such α, we have $P^\alpha \mathcal{H}_{\text{Fermi}} = P^\alpha L^2(\mathbb{R}^{3N})$. In fact, studying the operator H on the subspaces $P^\alpha L^2(\mathbb{R}^{3N})$ gives us complete information on the spectrum of the operator on $\mathcal{H}_{\text{Fermi}}$. To that end let

$$H^\alpha := H P^\alpha \tag{9.6}$$

be the operator H restricted to the space $\mathcal{H}^\alpha$ and

$$E^\alpha_{(X_1, \cdots, X_M)} := \inf \sigma(H^\alpha). \tag{9.7}$$

In the work at hand, we will compute the interaction energy for fixed positions of the nuclei, which is the difference between $E^\alpha_{(X_1, \cdots, X_M)}$ and the sum of ground state energies of atoms. Let us start with the simplest case of a diatomic molecule, i.e. $M = 2$.

9.1. **Diatomic molecules.** Let $\mathcal{C} \subsetneq \{1, \cdots, N\}$, $\mathcal{C} \neq \emptyset$ be an arbitrary subsystem of a system of N electrons. We define $\mathbb{R}(\mathcal{C})$ as the vector space of position vectors $(x_i)_{i\in\mathcal{C}}$ of particles in $\mathcal{C}$. Note that this space is isomorphic to $\mathbb{R}^{3\sharp\mathcal{C}}$, where $\sharp\mathcal{C}$ is the number of elements in $\mathcal{C}$. We let $L^2(\mathbb{R}(\mathcal{C}))$ be the space of L^2-functions with arguments in $\mathbb{R}(\mathcal{C})$. Denote by $L^2(\mathbb{R}(\mathcal{C}))^\perp$ the orthogonal complement in $L^2(\mathbb{R}^{3N})$ of $L^2(\mathbb{R}(\mathcal{C}))$.

For particles in $\mathcal{C}$ interacting via Coulomb potential with a nucleus at the origin of charge eZ we define the Hamiltonian

$$\tilde{H}^Z_{\mathcal{C}} := \sum_{i\in\mathcal{C}} T_i - \sum_{i\in\mathcal{C}} \frac{e^2 Z}{|x_i|} + \sum_{\substack{i,j\in\mathcal{C}\\ i<j}} \frac{e^2}{|x_i - x_j|} \tag{9.8}$$

acting on $L^2(\mathbb{R}(\mathcal{C}))$. We extend the operator by the identity in $L^2(\mathbb{R}(\mathcal{C}))^\perp$ to an operator acting on functions in $L^2(\mathbb{R}^{3N})$. In abuse of notation we will write $\tilde{H}^Z_{\mathcal{C}}$ for both, the one acting on $L^2(\mathbb{R}(\mathcal{C}))$ and the operator acting on $L^2(\mathbb{R}(\mathcal{C}))\oplus L^2(\mathbb{R}(\mathcal{C}))^\perp$.

Let $S(\mathcal{C})$ be the group of permutations within $\mathcal{C}$. Obviously $S(\mathcal{C})$ is a subgroup of S_N. Consider $\alpha_{\mathcal{C}}$ to be an irreducible representation of $S(\mathcal{C})$.

Definition 9.1. *For α a type of irreducible representation of S_N, we say that $\alpha'_{\mathcal{C}}$ is* induced *by α and write $\alpha'_{\mathcal{C}} \prec \alpha$, if $\alpha'_{\mathcal{C}}$ is contained in the restriction of α to $S(\mathcal{C})$, see* [18, p. 94-98].

In the same way as the space $\mathcal{H}_{\text{Fermi}}$ can be decomposed into the spaces $\mathcal{H}^\alpha$, the corresponding Fermi subspace of $L^2(\mathbb{R}(\mathcal{C}))$ can be decomposed into subspaces

$$P^{\alpha_{\mathcal{C}}} L^2(\mathbb{R}(\mathcal{C}))$$

where $\alpha_{\mathcal{C}}$ runs over all irreducible representations of $S(\mathcal{C})$ corresponding to a Young pattern of at most two columns.

We will consider a cluster decomposition $\beta = (\mathcal{C}_1, \mathcal{C}_2)$ of the original system $\{1, \cdots, N\}$ into clusters $\mathcal{C}_1$ and $\mathcal{C}_2$ such that $\mathcal{C}_1 \cup \mathcal{C}_2 = \{1, \cdots, N\}$ and $\mathcal{C}_1 \cap \mathcal{C}_2 = \emptyset$. Define $\mathcal{D}^2_N$ as the set of all such decompositions. Decompositions where the number of electrons in $\mathcal{C}_1$, $\sharp\mathcal{C}_1 = Z_1$ and the number of electrons in $\mathcal{C}_2$, $\sharp\mathcal{C}_2 = Z_2$ will be called *atomic* decomposition $\mathcal{D}^{at} \subset \mathcal{D}^2_N$.

For the decomposition $\beta = (\mathcal{C}_1, \mathcal{C}_2)$ we define the intercluster interaction

$$I_\beta := \sum_{i\in\mathcal{C}_1} \frac{-e^2 Z_2}{|x_i - X_2|} + \sum_{j\in\mathcal{C}_2} \frac{-e^2 Z_1}{|x_j - X_1|} + \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} \frac{e^2}{|x_i - x_j|} + \frac{e^2 Z_1 Z_2}{|X_2 - X_1|} \tag{9.9}$$

and set the cluster Hamiltonian H_β to be

$$H_\beta := H - I_\beta. \tag{9.10}$$

In other words, H_β is the operator where particles from different subsystems do not interact. Note that for each $\beta \in \mathcal{D}^2_N$ we have $L^2(\mathbb{R}(\mathcal{C}_1))^\perp = L^2(\mathbb{R}(\mathcal{C}_2))$. The symmetry group of this Hamiltonian we consider is $S_\beta := S(\mathcal{C}_1) \times S(\mathcal{C}_2) \subset S_N$, the group of permutations which leave the cluster decomposition β intact. We use the same notion of inducing of representations as above. Since S_β is a direct product of two groups, the irreducible representations α_β of S_β are direct products too. In particular, for any irreducible representation $\alpha'_\beta \prec \alpha$ of S_β there exists a unique pair $\alpha'_{\mathcal{C}_1} \prec \alpha$ and $\alpha'_{\mathcal{C}_2} \prec \alpha$ such that

$$\alpha'_{\mathcal{C}_1} \otimes \alpha'_{\mathcal{C}_2} \cong \alpha'_\beta, \tag{9.11}$$

see [18, p. 110-114]. We take $P^{\alpha'_\beta}$ to be the projection in $\mathcal{H}_{\text{Fermi}}$ onto functions of symmetry type α'_β. Letting

$$H_\beta^{\alpha'_\beta} := H_\beta P^{\alpha'_\beta} \quad \text{and} \quad H_\beta^\alpha := \sum_{\alpha'_\beta \prec \alpha} H_\beta^{\alpha'_\beta}, \tag{9.12}$$

we define

$$\mu_\beta^\alpha := \min_{\alpha'_\beta \prec \alpha} \inf \sigma(H_\beta^{\alpha'_\beta}) \tag{9.13}$$

and

$$\mu^\alpha := \min_{\beta \in \mathcal{D}_N^2} \mu_\beta^\alpha. \tag{9.14}$$

By translation and rotation invariance of the Hamiltonian for $M = 2$, $E^\alpha_{(X_1,X_2)}$ only depends on $|D|$, where $D := X_2 - X_1$. We will write $E^\alpha_{|D|}$ instead of $E^\alpha_{(X_1,X_2)}$. In both, the pseudo-relativistic and the nonrelativistic, cases it is not difficult to see that $\mu^\alpha = \lim_{|D|\to\infty} E^\alpha_{|D|}$.

For some fixed point $X \in \mathbb{R}^3$, which will be the position of one of the nuclei, and the variable $x \in \mathbb{R}^{3N}$, we define the unitary shift by X in the i-th particle variable as

$$\mathcal{U}_X^{(i)} : \begin{cases} L^2(\mathbb{R}^{3N}) \to L^2(\mathbb{R}^{3N}) \\ \mathcal{U}_X^{(i)}\varphi(x) \mapsto \varphi(x_1, \cdots, x_{i-1}, x_i + X, x_{i+1}, \cdots, x_N). \end{cases} \tag{9.15}$$

For $\beta = (\mathcal{C}_1, \mathcal{C}_2) \in \mathcal{D}_N^2$ and X_1, X_2 being the positions of the nuclei we define the shift operators

$$\mathcal{U}_\beta := \prod_{i\in\mathcal{C}_1} \mathcal{U}_{X_1}^{(i)} \prod_{j\in\mathcal{C}_2} \mathcal{U}_{X_2}^{(j)}. \tag{9.16}$$

We set

$$\tilde{H}_\beta := \mathcal{U}_\beta H_\beta \mathcal{U}_\beta^*. \tag{9.17}$$

Note that $\tilde{H}_\beta$ is unitary equivalent to H_β and

$$\tilde{H}_\beta = \tilde{H}_{\mathcal{C}_1}^{Z_1} + \tilde{H}_{\mathcal{C}_2}^{Z_2}.$$

We define for $\beta \in \mathcal{D}_N^2$ the functions $f_2, f_3 \in C^\infty(\mathbb{R}^{3N})$ as

$$f_2(x) := \sum_{\substack{i\in\mathcal{C}_1 \\ j\in\mathcal{C}_2}} -e^2\big(3(x_i \cdot e_D)(x_j \cdot e_D) - x_i \cdot x_j\big), \tag{9.18}$$

$$\begin{aligned} f_3(x) := \sum_{\substack{i\in\mathcal{C}_1 \\ j\in\mathcal{C}_2}} \frac{e^2}{2}\Big(&3(x_i - x_j)\cdot e_D\big[2(x_i \cdot x_j) - 5(x_i \cdot e_D)(x_j \cdot e_D)\big] \\ &+ 3|x_i|^2(x_j \cdot e_D) - 3|x_j|^2(x_i \cdot e_D)\Big), \end{aligned} \tag{9.19}$$

where $e_D := \frac{D}{|D|}$, a unit vector in the direction from X_1 to X_2. Note that the functions f_2, f_3 depend on the cluster decomposition β.

For now, let us fix any $\beta \in \mathcal{D}^{at}$. We will show in Appendix B that μ^α is a discrete eigenvalue of H_β^α. By unitary equivalence μ^α is also a discrete eigenvalue of

$$\tilde{H}_\beta^\alpha := \sum_{\alpha'_\beta \prec \alpha} \tilde{H}_\beta^{\alpha'_\beta} := \sum_{\alpha'_\beta \prec \alpha} \tilde{H}_\beta P^{\alpha'_\beta} \tag{9.20}$$

where the sum is over all induced irreducible representations $\alpha'_\beta \prec \alpha$. Denote by $\tilde{\mathcal{W}}_\beta^\alpha \subset \mathcal{H}^\alpha$ the eigenspace of $\tilde{H}_\beta^\alpha$ corresponding to μ^α and let

$$a_1(\beta) := \max_{\substack{\phi\in\tilde{\mathcal{W}}_\beta^\alpha \\ \|\phi\|=1}} \|(\tilde{H}_\beta - \mu^\alpha)^{-\frac{1}{2}} f_2 \phi\|^2. \tag{9.21}$$

Although μ^α is an eigenvalue of $\tilde{H}_\beta$ the value $a_1(\beta)$ is well-defined since $f_2\phi$ is orthogonal to the corresponding eigenspace, see Lemma E.2. We define $\tilde{\mathcal{V}}_\beta^\alpha \subset \tilde{\mathcal{W}}_\beta^\alpha$ as the subspace of all ϕ such that $\|(\tilde{H}_\beta - \mu^\alpha)^{-\frac{1}{2}} f_2\phi\|^2 = a_1(\beta)$ and

$$a_2(\beta) := \max_{\substack{\phi\in\tilde{\mathcal{V}}_\beta^\alpha \\ \|\phi\|=1}} \|(\tilde{H}_\beta - \mu^\alpha)^{-\frac{1}{2}} f_3\phi\|^2. \tag{9.22}$$

Similarly, Lemma E.2 ensures that also $a_2(\beta)$ is well-defined. Due to permutational symmetry, for any $\beta_1, \beta_2 \in \mathcal{D}^{at}$ we have $a_1(\beta_1) = a_1(\beta_2)$ and $a_2(\beta_1) = a_2(\beta_2)$. Hence we omit the argument β in the definition and write a_1 and a_2 throughout the paper. For diatomic molecules our main result is

Theorem 9.2 (Van der Waals interaction for diatomic molecules)**.** *Let α be an irreducible representation of S_N corresponding to a Young pattern with at most two columns and let the following conditions hold:*
1) For all $\beta \in \mathcal{D}_N^2 \setminus \mathcal{D}^{at}$

$$\mu_\beta^\alpha > \mu^\alpha.$$

2) For each $\beta \in \mathcal{D}^{at}$ and each irreducible representation α_β^* of the group S_β with $\alpha_\beta^* \prec \alpha$ such that $P^{\alpha_\beta^*}\tilde{\mathcal{W}}_\beta^\alpha \neq \emptyset$,

$$\dim(P^{\alpha_\beta^*}\tilde{\mathcal{W}}_\beta^\alpha) = \dim \alpha_\beta^*.$$

Then

$$E_{|D|}^\alpha - \mu^\alpha = -\frac{a_1}{|D|^6} - \frac{a_2}{|D|^8} + \mathcal{O}(|D|^{-10}) \tag{9.23}$$

where a_1 and a_2 are defined in (9.21) *and* (9.22) *respectively.*

Remarks 9.3.

- *Conditions 1) and 2) of Theorem 9.2 are the same as in the previous work* [3] *by I. Anapolitanos and I. M. Sigal, where they obtained an asymptotic expansion of $E_{|D|}^\alpha - \mu^\alpha$ in the nonrelativistic case with an error of order $\mathcal{O}(|D|^{-7})$.*
- *The physical meaning of Condition 1) is that the lowest energy of the non-interacting system occurs when the electrons are allocated neutrally. It is important to mention that if Condition 1) does not hold, then $E_{|D|}^\alpha - \mu^\alpha$ is dominated by Coulomb interaction which decays like $|D|^{-1}$ and is thus much stronger than the van der Waals-London interaction. Experimental data suggests that for any tuple of stable atoms this condition is true, see the discussion in the introduction of* [3]*. Further detail concerning this condition can be found in Chapter H of this thesis.*
- *Condition 2) is only needed in the proof of the lower bound for the vdW-L interaction. It imposes restrictions on the rotational symmetry of the atoms in the diatomic molecule. In particular the ground state space of $\tilde{H}_\beta^\alpha$ only contains functions which transform according to the irreducible representation of the group $SO(3)$ of degree $\ell = 0$. To see this, notice that the Hamiltonian $\tilde{H}_\beta^\alpha$ is invariant under rotations $R \in SO(3)$. Thus for any eigenfunction $\phi \in P^{\alpha'_\beta}\tilde{\mathcal{W}}_\beta^\alpha$ the rotated function $T_R\phi$ is an eigenfunction corresponding to the same value. Rotation and permutation operators commute, thus $T_R\phi \in P^{\alpha'_\beta}\tilde{\mathcal{W}}_\beta^\alpha$. So by* [18, §3.19] *the dimension of $P^{\alpha'_\beta}\tilde{\mathcal{W}}_\beta^\alpha$ is an integer multiple of the dimension of α'_β and the dimension of a representation of the $SO(3)$ group. By Condition 2) $\dim(P^{\alpha'_\beta}\tilde{\mathcal{W}}_\beta^\alpha) = \dim \alpha'_\beta$ so the dimension of the representation of $SO(3)$ describing the symmetry of ϕ is one. So it must be the irreducible representation of degree $\ell = 0$.*

- *Our method allows to obtain the expansion of $E^{\alpha}_{|D|} - \mu^{\alpha}$ up to arbitrary negative power of $|D|$. In particular, for diatomic molecules this expansion does not include odd powers $|D|^{-7}$ and $|D|^{-9}$ in both the pseudo-relativistic and nonrelativistic case. This seems to be new and unknown even in the physics literature on the van der Waals force.*
- *In the definition of functions f_2, f_3 and therefore in the definition of a_1 and a_2, we use the vector e_D. By the $SO(3)$ symmetry of $\tilde{H}_{\beta}$ and Condition 2), the values of a_1 and a_2 will not change if we replace e_D in* (9.18) *and* (9.19) *with an arbitrary normalized vector in $\mathbb{R}^3$.*
- *The functions f_2, f_3 are, respectively, the second- and third-order coefficients in the Taylor expansion of the intercluster interaction (see Appendix D). They are invariant under permutations in S_{β} and hence for any irreducible representation $\alpha'_{\beta} \prec \alpha$ of S_{β}, we have $f_l P^{\alpha'_{\beta}} = P^{\alpha'_{\beta}} f_l$, for $l = 2, 3$.*
- *Numerical calculations in* [47] *for the coefficients of the interaction energy of H, He and Li are consistent with the statement and predict the behaviour given in* (9.23).

Strategy of the proof of Theorem 9.2 To prove the main result, we derive estimates of the difference $E^{\alpha}_{|D|} - \mu^{\alpha}$ from above and from below. These bounds coincide up to an order $\mathcal{O}(|D|^{-10})$. To get an estimate from below for the interaction energy, we apply a partition of unity to the configuration space, and minimize the functionals in the corresponding regions. To obtain an upper bound, we construct a suitable trial function.

More precisely, let $\beta^0 = (\{1, \cdots, Z_1\}, \{Z_1 + 1, \cdots, N\})$. By permutation symmetry of the operator $\tilde{H}^{\alpha}_{\beta^0}$, the ground state space $\tilde{\mathcal{W}}^{\alpha}_{\beta^0}$ of $\tilde{H}^{\alpha}_{\beta^0}$ can be written as a direct sum of subspaces transforming according to the induced irreducible representations $\alpha'_{\beta^0} \prec \alpha$, more explicitly

$$\tilde{\mathcal{W}}^{\alpha}_{\beta^0} = \bigoplus_{\alpha'_{\beta^0} \prec \alpha} P^{\alpha'_{\beta^0}} \tilde{\mathcal{W}}^{\alpha}_{\beta^0}.$$

Thus there is at least one $\alpha^*_{\beta^0} \prec \alpha$ such that there exists $\phi \in P^{\alpha^*_{\beta^0}} \tilde{\mathcal{W}}^{\alpha}_{\beta^0}$ that realises the maxima a_1 and a_2 with $\|\phi\| = 1$. For such a $\phi \in P^{\alpha^*_{\beta^0}} \tilde{\mathcal{W}}^{\alpha}_{\beta^0}$ we define

$$\Upsilon := \chi_o(x) \Big\{ \phi(x) - (\tilde{H}_{\beta^0} - \mu^{\alpha})^{-1} \Big(\frac{f_2(x)}{|D|^3} + \frac{f_3(x)}{|D|^4} \Big) \phi(x) \Big\} \tag{9.24}$$

where $\chi_o(x)$ is a smooth function which localizes each particle in a ball of radius $|D|^{\frac{3}{4}}$, centered at the origin. As a trial function, which yields the required estimate of $E^{\alpha}_{|D|} - \mu^{\alpha}$ from above, we define $\Upsilon_{\text{trial}} := P^{\alpha} \mathcal{U}^*_{\beta^0} \Upsilon$.

To prove the estimate from above, we need to show that applying the cutoff function $\chi_o(x)$ increases the energy only by an exponentially small amount. To this end we need to prove exponential decay of ϕ, $(\tilde{H}_{\beta^0} - \mu^{\alpha})^{-1} f_2 \phi$, and $(\tilde{H}_{\beta^0} - \mu^{\alpha})^{-1} f_3 \phi$, which is done in Section 10. In addition, we need a suitable estimate for the so-called localization error for the pseudo-relativistic kinetic energy. Such an estimate is obtained in Section 11. In both cases, the proof of exponential decay and the estimate of the localization error, the main difficulty arises from the non-locality of the pseudo-relativistic kinetic energy operator.

For the estimate from below we consider all possible cluster decompositions into three clusters $\beta = (\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2)$. Some of the clusters may be empty. Particles in $\mathcal{C}_0$ are far from the nucleus. Electrons in $\mathcal{C}_1$ and $\mathcal{C}_2$ are close to X_1 and X_2 respectively. We apply a partition of unity of the configurations space with smooth functions J_{β}

cutting the configuration space according to the clusters in β. If $\mathcal{C}_0 \neq \emptyset$ or if $\mathcal{C}_1$ and $\mathcal{C}_2$ are not neutral atoms, the infimum of the spectrum of the cluster Hamiltonian corresponding to this β on the subspace $\mathcal{H}^{\alpha'_\beta}$ is, by assumption, strictly greater than μ^α for all $\alpha'_\beta \prec \alpha$. For sufficiently large $|D|$, this implies

$$\langle J_\beta \psi, (H^\alpha - \mu^\alpha) J_\beta \psi \rangle \geq 0.$$

Now consider β for which $\mathcal{C}_0 = \emptyset$, and $(\mathcal{C}_1, \mathcal{C}_2) \in \mathcal{D}^{at}$. Similar to [6, 7, 8, 9, 43, 44, 45] we define a bilinear form

$$\langle \varphi, \psi \rangle_1 := \langle \varphi, (\tilde{H}_\beta - \mu^\alpha)\psi \rangle$$

and the corresponding semi-norm

$$\|\varphi\|_1^2 := \langle \varphi, \varphi \rangle_1.$$

Then we project the state $\mathcal{U}_\beta J_\beta \psi$ onto the ground state subspace $\tilde{\mathcal{W}}_\beta^\alpha$ of the operator $\tilde{H}_\beta^\alpha$ to get

$$\mathcal{U}_\beta J_\beta \psi = \gamma_1 \phi + \mathcal{R}$$

for a normalized state $\phi \in \tilde{\mathcal{W}}_\beta^\alpha$. We proceed by projecting the rest term $\mathcal{R}$ onto the functions

$$\begin{aligned} \phi_2 &= (\tilde{H}_\beta - \mu^\alpha)^{-1} f_2 \phi, \\ \phi_3 &= (\tilde{H}_\beta - \mu^\alpha)^{-1} f_3 \phi \end{aligned} \tag{9.25}$$

consecutively, with respect to $\langle \cdot, \cdot \rangle_1$. Note that by Corollary E.3 these states are well-defined. For the state $J_\beta \psi$ we arrive at the following representation

$$J_\beta \psi = \mathcal{U}_\beta^* \big(\gamma_1 \phi + |D|^{-3} \gamma_2 \phi_2 + |D|^{-4} \gamma_3 \phi_3 + g \big) \tag{9.26}$$

for a suitable function g. We substitute (9.26) into the quadratic form of

$$(H - \mu^\alpha) P^\alpha = (H_\beta - \mu^\alpha + I_\beta) P^\alpha.$$

Then we expand I_β as a Taylor series and do a simple minimization in parameters $\gamma_1, \gamma_2, \gamma_3$, using orthogonality relations proven in Appendix D. It turns out that $\|g\|$ will be very small and $\gamma_1, \gamma_2, \gamma_3$ close to the coefficients of the trial function, which we used to get the upper bound, when ψ is close to a minimizer of the energy.

Finally, in analogy to the estimate from above, the localization error is small on ϕ, ϕ_2, and ϕ_3 due to their exponential decay.

9.2. **Extension to M-atomic molecules.** We can extend the result of Theorem 9.2, stated for a diatomic molecule, to larger systems.

We will assume that the distances between atoms are simultaneously scaled by a parameter $d > 0$. For all $1 \leq k < l \leq M$, we write $X_k - X_l =: dD_{k,l}$, where vectors $D_{k,l}$ are assumed to be fixed. The scaling parameter d will tend to infinity. The operator H can be written as

$$H = \sum_{i=1}^N \left(T_i - \sum_{k=1}^M \frac{e^2 Z_k}{|x_i - X_k|} \right) + \sum_{1 \leq i < j \leq N} \frac{e^2}{|x_i - x_j|} + \sum_{1 \leq k < l \leq M} \frac{e^2 Z_k Z_l}{d|D_{k,l}|}. \tag{9.27}$$

We let

$$E_d^\alpha := \inf \sigma(H^\alpha) \tag{9.28}$$

denote the infimum of the spectrum of H restricted to the space $\mathcal{H}^\alpha = P^\alpha \mathcal{H}_{\text{Fermi}}$.

Consider the cluster decomposition $\beta_M := (\mathcal{C}_1, \cdots, \mathcal{C}_M)$ of the original system into M clusters such that $\bigcup_{k=1}^M \mathcal{C}_k = \{1, \cdots, N\}$ and $\mathcal{C}_k \cap \mathcal{C}_l = \emptyset$ for all $k \neq l$. We define the set $\mathcal{D}_N^M$ as the collection of all such decompositions. Let

$$\tilde{H}_{\beta_M} := \sum_{k=1}^M \tilde{H}_{\mathcal{C}_k}^{Z_k} \tag{9.29}$$

where $\tilde{H}^{Z_k}_{\mathcal{C}_k}$ defined according to (9.8), acting on the space $L^2(\mathbb{R}^{3N})$. The symmetry group of this Hamiltonian is $S_{\beta_M} := S(\mathcal{C}_1) \times \cdots \times S(\mathcal{C}_M) \subset S_N$, the group of permutations which leave the cluster decomposition β_M intact. Once again, the irreducible representations of S_{β_M} can be expressed as direct products of irreducible representations $\alpha'_{\mathcal{C}}$ of $S(\mathcal{C})$. In particular, for any irreducible representation $\alpha'_{\beta_M} \prec \alpha$ of S_{β_M} there exists a unique M-tuple of irreducible representations $\alpha'_{\mathcal{C}_k} \prec \alpha$ such that

$$\bigotimes_{k=1}^{M} \alpha'_{\mathcal{C}_k} \cong \alpha'_{\beta_M}. \tag{9.30}$$

We take $P^{\alpha'_{\beta_M}}$ to be the projection in $\mathcal{H}_{\text{Fermi}}$ onto functions belonging to the irreducible representation α'_{β_M}. Letting $\tilde{H}^{\alpha'_{\beta_M}}_{\beta_M} := \tilde{H}_{\beta_M} P^{\alpha'_{\beta_M}}$ we define

$$\mu^{\alpha}_{\beta_M} := \min_{\alpha'_{\beta_M} \prec \alpha} \inf \sigma\big(\tilde{H}^{\alpha'_{\beta_M}}_{\beta_M}\big) \tag{9.31}$$

and

$$\mu^{\alpha}_{M} := \min_{\beta_M \in \mathcal{D}^M_N} \mu^{\alpha}_{\beta_M}. \tag{9.32}$$

Similar to the diatomic case $\mu^{\alpha}_M = \lim_{d\to\infty} E^{\alpha}_d$. We define the functions $f_2^{(k,l)}$, $f_3^{(k,l)} \in C^\infty(\mathbb{R}^{3N})$ as

$$f_2^{(k,l)}(x) := \sum_{\substack{i\in\mathcal{C}_k \\ j\in\mathcal{C}_l}} -e^2\big(3(x_i\cdot e_{D_{k,l}})(x_j\cdot e_{D_{k,l}}) - x_i\cdot x_j\big), \tag{9.33}$$

$$\begin{aligned} f_3^{(k,l)}(x) := \sum_{\substack{i\in\mathcal{C}_k \\ j\in\mathcal{C}_l}} \frac{e^2}{2}\Big(&3(x_i - x_j)\cdot e_{D_{k,l}}\big[2(x_i\cdot x_j) - 5(x_i\cdot e_{D_{k,l}})(x_j\cdot e_{D_{k,l}})\big] \\ &+ 3|x_i|^2(x_j\cdot e_{D_{k,l}}) - 3|x_j|^2(x_i\cdot e_{D_{k,l}})\Big), \end{aligned} \tag{9.34}$$

where $e_{D_{k,l}} := \frac{D_{k,l}}{|D_{k,l}|}$ is the unit vector in the direction from nucleus k to nucleus l. The functions $f_2^{(k,l)}$ and $f_3^{(k,l)}$ are related to the second- and third-order coefficients in the Taylor expansion of the intercluster interaction of cluster k with cluster l, see Appendix D for details.

The value μ^{α}_M defined in (9.32) is a discrete eigenvalue of the operator $\tilde{H}^{\alpha}_{\beta_M}$, see Theorem B.1. Denote by $\tilde{\mathcal{W}}^{\alpha}_{\beta_M} \subset \mathcal{H}^{\alpha}$ the eigenspace of $\tilde{H}^{\alpha}_{\beta_M}$ corresponding to μ^{α}_M and let

$$a_1^M := \max_{\substack{\phi\in\tilde{\mathcal{W}}^{\alpha}_{\beta_M} \\ \|\phi\|=1}} \|(\tilde{H}_{\beta_M} - \mu^{\alpha}_M)^{-\frac{1}{2}} \sum_{1\le k<l\le M} |D_{k,l}|^{-3} f_2^{(k,l)}\phi\|^2. \tag{9.35}$$

We define $\tilde{\mathcal{V}}^{\alpha}_{\beta_M} \subset \tilde{\mathcal{W}}^{\alpha}_{\beta_M}$ the subspace of all ϕ such that

$$\|(\tilde{H}_{\beta_M} - \mu^{\alpha}_M)^{-\frac{1}{2}} \sum_{1\le k<l\le M} |D_{k,l}|^{-3} f_2^{(k,l)}\phi\|^2 = a_1^M$$

and

$$a_2^M := \max_{\substack{\phi\in\tilde{\mathcal{V}}^{\alpha}_{\beta_M} \\ \|\phi\|=1}} \|(\tilde{H}_{\beta_M} - \mu^{\alpha}_M)^{-\frac{1}{2}} \sum_{1\le k<l\le M} |D_{k,l}|^{-4} f_3^{(k,l)}\phi\|^2. \tag{9.36}$$

Slightly abusing notation, for $\beta \in \mathcal{D}^M_N$ we write $\beta \in \mathcal{D}^{at}$ iff for all $k \in \{1, \cdots, M\}$ one hase $\sharp\mathcal{C}_k = Z_k$.

Theorem 9.4. *Let α be an irreducible representation of S_N corresponding to a Young pattern with at most two columns and let the following conditions hold:*
1') For all $\beta \in \mathcal{D}_N^M \setminus \mathcal{D}^{at}$

$$\mu^\alpha_{\beta_M} > \mu^\alpha_M. \tag{9.37}$$

2') For each induced irredducible representation $\alpha^*_{\beta_M} \prec \alpha$ of the group S_N such that $P^{\alpha^*_{\beta_M}} \tilde{\mathcal{W}}^\alpha_{\beta_M} \neq \emptyset$,

$$\dim(P^{\alpha^*_{\beta_M}} \tilde{\mathcal{W}}^\alpha_{\beta_M}) = \dim \alpha^*_{\beta_M}. \tag{9.38}$$

Then

$$E^\alpha_d - \mu^\alpha_M = -\frac{a_1^M}{d^6} - \frac{a_2^M}{d^8} + \frac{a_3^M}{d^9} + \mathcal{O}(d^{-10}).$$

where

$$a_3^M = \sum_{\substack{k\neq l \\ l\neq n, n\neq k}} \frac{\langle (\tilde{H}_{\beta_M} - \mu^\alpha_M)^{-1} f_2^{(k,l)}\phi, f_2^{(l,n)} (\tilde{H}_{\beta_M} - \mu^\alpha_M)^{-1} f_2^{(n,k)}\phi\rangle}{8|D_{k,l}|^3 |D_{l,n}|^3 |D_{n,k}|^3}. \tag{9.39}$$

Remarks 9.5.

- *The term of order d^{-6} is a sum of the corresponding terms in Theorem 9.2. Again no term of order d^{-7} appears. The main difference to the diatomic case is the appearance of the term of order d^{-9}. This term stems from an interaction of three atoms. Each of the atoms induces dipole momenta in the other two atoms of this triplet. Their interaction is proportional to d^{-9}.*
- *In the diatomic case the result will not change if we replace the vector e_D in the definition f_2, f_3, (9.18) and (9.19) by an arbitrary normalized vector. In contrast to that, in the multi-atomic case the term of order d^{-9} depends on the angles between vectors $D_{k,l}, D_{l,n}$ and $D_{n,k}$.*

The paper is organized as follows.

In Section 10 we prove exponential decay of functions $\phi, (\tilde{H}_\beta - \mu^\alpha)^{-1} f_2\phi$, and $(\tilde{H}_\beta - \mu^\alpha)^{-1} f_3\phi$, which play a crucial role in the proof of Theorems 9.2 and 9.4.

In Section 11 we prove a localization error estimate for the pseudo-relativistic kinetic energy, which shows that outside the region, where the derivative of the cutoff function is non-zero, the localization error is exponentially small.

In Sections 12 and 13 we prove Theorems 9.2 and 9.4 respectively.

In Appendix A and B we prove the HVZ theorem for atoms and atomic ions and the existence of a ground state for pseudo-relativistic atoms and positive ions on spaces with fixed permutation symmetry. This result was announced by G. Zhislin in [49]. For convenience of the reader we give a complete proof of these statements.

In Appendix C and D we prove several technical estimates, which we use in Sections 10 and 12, respectively.

Finally, in Appendix E we prove orthogonality relations, which are due to the symmetry of functions ϕ and I_β.

10. Exponential decay of eigenfunctions

In the nonrelativistic case, exponential decay of eigenfunctions with given permutation symmetry is well-known (see e.g. [2]). The exponential decay of eigenfunctions of a Hamiltonian with pseudo-relativistic kinetic energy proved by Carmona, Masters and Simon in [12] does not apply for Coulomb potentials, however. For pseudo-relativistic kinetic energy and Coulomb potentials, exponential decay of eigenfunctions was shown by Nardini in [35] for the two body case. He extended his results to the N-body case in [36]. However, in the proof he uses a method

which destroys permutational symmetries. To prove Theorem 9.2 we need exponential decay of ground states $\phi \in \tilde{\mathcal{W}}^{\alpha}_{\beta}$ of $\tilde{H}_\beta$ and exponential decay of functions of the form $(\tilde{H}_\beta - \mu^\alpha)^{-1} f_l \phi$, $l = 2, 3$ where $\phi \in \tilde{\mathcal{W}}^{\alpha}_{\beta}$ is a ground state. To this end we will apply a modification of Agmon's method (see [2]), adapted to the nonlocal presudo-relativistic kinetic energy, which preserves symmetry.

Let $\alpha_{\mathcal{C}}$ be an irreducible representation of $S(\mathcal{C})$. We define

$$\Sigma^{\alpha_{\mathcal{C}}} := \lim_{R\to\infty} \inf_{\substack{\psi \in P^{\alpha_{\mathcal{C}}} H^{1/2}(\mathbb{R}(\mathcal{C})) \\ \mathrm{supp}(\psi)\cap B_R(0)=\emptyset}} \|\psi\|^{-2} \langle \psi, \tilde{H}^Z_{\mathcal{C}} \psi\rangle, \tag{10.1}$$

where $B_R(0)$ is the ball in $\mathbb{R}(\mathcal{C})$ of radius R centered at 0 and $\tilde{H}^Z_{\mathcal{C}}$ was defined in (9.8). Everywhere in this section we treat the pseudo-relativistic kinetic energy operator $T_i = \sqrt{p_i^2 + 1} - 1$ only.

Theorem 10.1. *For any fixed $\mu < \Sigma^{\alpha_{\mathcal{C}}}$, assume that $\Upsilon \in H^{1/2}(\mathbb{R}(\mathcal{C}))$ satisfies $P^{\alpha_{\mathcal{C}}}\Upsilon = \Upsilon$ and $(\tilde{H}^Z_{\mathcal{C}} - \mu)\Upsilon = \Gamma$, where Γ is a function with $e^{a|\cdot|}\Gamma \in L^2(\mathbb{R}(\mathcal{C}))$ for some $a > 0$. Then there exists $b > 0$ such that*

$$e^{b|\cdot|}\Upsilon \in L^2(\mathbb{R}(\mathcal{C})). \tag{10.2}$$

Remark 10.2. *Choosing $\Gamma = 0$ in the above theorem implies that any eigenfunction Υ of $\tilde{H}^Z_{\mathcal{C}}$ with associated eigenvalue $\mu < \Sigma^{\alpha_{\mathcal{C}}}$ is exponentially decaying.*

In addition to Theorem 10.1 we will need a similar statement for cluster Hamiltonians $\tilde{H}_\beta$ corresponding to a cluster decomposition β into two clusters.

Proposition 10.3. *Let α_β be an irreducible representation of S_β and let*

$$\Sigma^{\alpha_\beta} := \lim_{R\to\infty} \inf_{\substack{\varphi \in P^{\alpha_\beta} H^{1/2}(\mathbb{R}^{3N}) \\ \mathrm{supp}(\varphi)\cap B_R(0)=\emptyset}} \|\varphi\|^{-2} \langle \varphi, \tilde{H}_\beta \varphi\rangle, \tag{10.3}$$

where $B_R(0)$ is the ball in $\mathbb{R}^{3N}$ with radius R centered at 0. For any fixed $\tilde{\mu} < \Sigma^{\alpha_\beta}$, assume that $\tilde{\Upsilon} \in L^2(\mathbb{R}^{3N})$ satisfies $P^{\alpha_\beta}\tilde{\Upsilon} = \tilde{\Upsilon}$ and $(\tilde{H}_\beta - \tilde{\mu})\tilde{\Upsilon} = \tilde{\Gamma}$, where $\tilde{\Gamma}$ is a function with $e^{a|\cdot|}\tilde{\Gamma} \in L^2(\mathbb{R}^{3N})$ for some $a > 0$. Then there exists $b > 0$ such that

$$e^{b|\cdot|}\tilde{\Upsilon} \in L^2(\mathbb{R}^{3N}). \tag{10.4}$$

The proof of Proposition 10.3 follows immediately from Theorem 10.1.

Corollary 10.4. *Let $\beta \in \mathcal{D}^{at}$ and $\tilde{\mathcal{W}}^{\alpha}_{\beta}$ be the ground state subspace of the Hamiltonian $\tilde{H}^{\alpha}_{\beta} = \tilde{H}_\beta P^\alpha$ corresponding to the energy μ^α. Then for any normalized function $\phi \in \tilde{\mathcal{W}}^{\alpha}_{\beta}$ the functions ϕ, $\phi_2 = (\tilde{H}_\beta - \mu^\alpha)^{-1} f_2 \phi$, and $\phi_3 = (\tilde{H}_\beta - \mu^\alpha)^{-1} f_3 \phi$ with f_2, f_3 defined in (9.18),(9.19) and some $b_1, b_2, b_3 > 0$ holds*

$$e^{b_1|\cdot|}\phi,\ e^{b_2|\cdot|}\phi_2,\ e^{b_3|\cdot|}\phi_3 \in L^2(\mathbb{R}^{3N}). \tag{10.5}$$

Since ϕ is a ground state, the existence of a $b_1 > 0$ follows immediately from Proposition 10.3. Notice that for $l = 2, 3$ we have

$$(\tilde{H}_\beta - \mu^\alpha)\phi_l = (\tilde{H}_\beta - \mu^\alpha)(\tilde{H}_\beta - \mu^\alpha)^{-1} f_l \phi = f_l \phi. \tag{10.6}$$

The functions f_l grow at most polynomially in $|x_i|$ which is controlled by the exponential decay of ϕ so we can apply Proposition 10.3 with $\tilde{\Gamma} = f_l \phi$ to obtain the result.

Proof of Theorem 10.1. Let $k := \sharp\mathcal{C}$ be the number of electrons in the cluster $\mathcal{C}$. To simplify the notation assume $\mathcal{C} = \{1, \cdots, k\}$. Let $\Upsilon \in H^{1/2}(\mathbb{R}(\mathcal{C}))$ be a solution of the equation $(\tilde{H}^Z_{\mathcal{C}} - \mu)\Upsilon = \Gamma$ and $\xi_\epsilon \in C^\infty(\mathbb{R}(\mathcal{C}); \mathbb{R})$ be a family of functions with the following properties:

- ξ_ϵ is bounded for any $\epsilon > 0$

- ξ_ϵ is invariant under all permutations of the variables $x_i \in \mathbb{R}^3$
- $\mathrm{supp}(\xi_\epsilon) \cap B_R(0) = \emptyset$ for some $R > 0$, which will be chosen later
- $|\xi_\epsilon| \le Ce^{\frac{a}{2}|\cdot|}$ for some constant $C > 0$, independent of ϵ

By the definition of Σ^{α_C} in (10.1) and since $\mathrm{supp}(\xi_\epsilon) \cap B_R(0) = \emptyset$, there exists a function $\vartheta(R)$ such that $\lim_{R\to\infty} \vartheta(R) = 0$ and

$$\big(\Sigma^{\alpha_C} - \mu - \vartheta(R)\big)\|\xi_\epsilon \Upsilon\|^2 \le \langle \xi_\epsilon \Upsilon, (\tilde{H}^Z_{\mathcal{C}} - \mu)\xi_\epsilon \Upsilon\rangle. \tag{10.7}$$

Since $(\tilde{H}^Z_{\mathcal{C}} - \mu)\Upsilon = \Gamma$, we can write

$$\begin{aligned} \langle \xi_\epsilon \Upsilon, (\tilde{H}^Z_{\mathcal{C}} - \mu)\xi_\epsilon \Upsilon\rangle &= \mathrm{Re}\langle \xi_\epsilon \Upsilon, (\tilde{H}^Z_{\mathcal{C}} - \mu)\xi_\epsilon \Upsilon\rangle \\ &= \mathrm{Re}\langle \Upsilon, \xi_\epsilon^2 \Gamma\rangle + \mathrm{Re}\langle \xi_\epsilon \Upsilon, [\xi_\epsilon, \tilde{H}^Z_{\mathcal{C}}]\Upsilon\rangle \end{aligned} \tag{10.8}$$

where the absolute value of the first term on the r.h.s. is bounded above by $C_1\|\Upsilon\|$ for some constant $C_1 < \infty$, since $e^{a|\cdot|}\Gamma \in L^2(\mathbb{R}(\mathcal{C}))$. Together with (10.7) we get

$$\big(\Sigma^{\alpha_C} - \mu - \vartheta(R)\big)\|\xi_\epsilon \Upsilon\|^2 \le C_1\|\Upsilon\| + |\,\mathrm{Re}\langle \xi_\epsilon \Upsilon, [\xi_\epsilon, \tilde{H}^Z_{\mathcal{C}}]\Upsilon\rangle|. \tag{10.9}$$

We now specify the choice of ξ_ϵ: For $\nu \in (0, \frac{a}{2})$ and for $\epsilon \in (0,1)$ we set

$$F(r) = F_{\nu,\epsilon}(r) := \frac{\nu r}{1 + \epsilon r}. \tag{10.10}$$

Given $R > 0$ pick $h_R \in C^\infty(\mathbb{R}_+; [0,1])$ such that

$$h_R(r) := \begin{cases} 1 & \text{if} \quad r < R \\ 0 & \text{if} \quad r > 2R \end{cases}$$

and define for $x \in \mathbb{R}^{3k}$ the function $\chi_R \in C^\infty(\mathbb{R}^{3k}; [0,1])$ with

$$\chi_R(x) := 1 - \prod_{i=1}^k h_R(|x_i|). \tag{10.11}$$

Let

$$\xi_\epsilon(x) := \chi_R(x) \sum_{i=1}^k e^{F(|x_i|)}. \tag{10.12}$$

To estimate the second term on the r.h.s. of (10.9) we will use the following lemma, which is proved in Appendix C.

Lemma 10.5. *Let the functions $F_{\nu,\epsilon}$, and ξ_ϵ be as in* (10.10) *and* (10.12)*. For fixed $R > 0$ and arbitrary $\hat{\delta} > 0$ there exists $\nu_0 > 0$ and $C_{\nu_0,R} < \infty$ such that for all $\nu \in (0, \nu_0)$, all $\epsilon \in (0,1)$, and all $\varphi \in H^{1/2}(\mathbb{R}(\mathcal{C}))$*

$$|\,\mathrm{Re}\langle \xi_\epsilon \varphi, [\xi_\epsilon, \tilde{H}^Z_{\mathcal{C}}]\varphi\rangle| \le \hat{\delta}\|\sum_{i=1}^k e^{F_{\nu,\epsilon}(|x_i|)}\varphi\|^2 + C_{\nu_0,R}\|\varphi\|^2. \tag{10.13}$$

Substituting (10.13) in (10.9) yields

$$\begin{aligned} &(\Sigma^{\alpha_C} - \mu - \vartheta(R))\|\xi_\epsilon \Upsilon\|^2 \\ &\le C_1\|\Upsilon\| + \hat{\delta}\|\sum_{i=1}^k e^{F_{\nu,\epsilon}(|x_i|)}\Upsilon\|^2 + C_{\nu_0,R}\|\Upsilon\|^2 \\ &\le C_1\|\Upsilon\| + \hat{\delta}\|\xi_\epsilon \Upsilon\|^2 + \hat{\delta}\|(1-\chi_R)\sum_{i=1}^k e^{F_{\nu,\epsilon}(|x_i|)}\Upsilon\|^2 + C_{\nu_0,R}\|\Upsilon\|^2 \end{aligned}$$

for any $\nu < \nu_0$. Moreover for all $i \in \{1, \cdots, k\}$ we have $(1-\chi_R)e^{F(|x_i|)} \le e^{2\nu R}$. Hence

$$(\Sigma^{\alpha_C} - \mu - \vartheta(R))\|\xi_\epsilon \Upsilon\|^2 \le C_1\|\Upsilon\| + \hat{\delta}\|\xi_\epsilon \Upsilon\|^2 + \hat{\delta}k^2 e^{4\nu R}\|\Upsilon\|^2 + C_{\nu_0,R}\|\Upsilon\|^2$$

which is equivalent to

$$(\Sigma^{\alpha_C} - \mu - \vartheta(R) - \hat{\delta})\|\xi_\epsilon \Upsilon\|^2 \leq C_1\|\Upsilon\| + \hat{\delta}k^2 e^{4\nu R}\|\Upsilon\|^2 + C_{\nu_0,R}\|\Upsilon\|^2. \tag{10.14}$$

Choosing $R > 0$ such that $\vartheta(R) < \frac{\Sigma^{\alpha_C} - \mu}{3}$ and picking $\hat{\delta}$ such that $\hat{\delta} < \frac{\Sigma^{\alpha_C} - \mu}{3}$, we arrive at

$$(\Sigma^{\alpha_C} - \mu - \vartheta(R) - \hat{\delta}) > \frac{\Sigma^{\alpha_C} - \mu}{3} =: \gamma.$$

From (10.14) we get

$$\gamma\|\xi_\epsilon \Upsilon\|^2 < C_1\|\Upsilon\| + \hat{\delta}k^2 e^{4\nu R}\|\Upsilon\|^2 + C_{\nu_0,R}\|\Upsilon\|^2 \tag{10.15}$$

and deviding by γ on both sides yields

$$\|\xi_\epsilon \Upsilon\|^2 < \frac{C_1}{\gamma}\|\Upsilon\| + \frac{k^2\hat{\delta}e^{4\nu R} + C_{\nu_0,R}}{\gamma}\|\Upsilon\|^2. \tag{10.16}$$

Note that the bound (10.16) is uniform in ϵ. Since ξ_ϵ converges monotonically to $e^{\nu|\cdot|}$ as $\epsilon \to 0$, the monotone convergence theorem completes the proof of Theorem 10.1. □

11. Localization error estimate

In the proof of Theorem 9.2 and Theorem 9.4 we will use a partition of unity of the configuration space. In addition to this, we use a cutoff function in our construction of the trial function which we will use to bound the intercluster energy from above (see introduction in Section 9). To obtain the required upper bound we need to show that cutting the ground states of the subsystems leads to an exponentially small increase in the expectation value of the intercluster energy. Therefore we need a suitable estimate of the so-called localization error. Note that in contrast to the nonrelativistic kinetic energy operator, the pesudo-relativistic operator is not local. Consequently the localization error is non-zero everywhere, including the regions where derivatives of the cutoff functions vanish. Of course, there exist several estimates for the localization error of the pseudo-relativistic kinetic energy. However none of them are precise enough for the proof of the van der Waals-London law. The main difference between the estimate for the localization error given below in Theorem 11.3 and the known results (see for example [11, 21, 22, 24, 28]) is, we show that the localization error decays exponentially with the distance to the support of the derivatives of the cutoff functions.

Let $u_{\boldsymbol{\rho}} \in C^1(\mathbb{R}^3; [0,1])$ such that

$$u_{\boldsymbol{\rho}}(z) := \begin{cases} 1 & \text{if } |z| \leq \frac{\boldsymbol{\rho}}{8} \\ 0 & \text{if } |z| > \frac{\boldsymbol{\rho}}{4} \end{cases} \tag{11.1}$$

and pick $v_{\boldsymbol{\rho}} \in C^1(\mathbb{R}^3; [0,1])$ such that $u_{\boldsymbol{\rho}}^2 + v_{\boldsymbol{\rho}}^2 = 1$ and $h \in H^{1/2}(\mathbb{R}^3)$, the one electron localization error $\mathcal{LE}_1[h]$ is given by

$$\mathcal{LE}_1[h] = \langle u_{\boldsymbol{\rho}} h, T_1 u_{\boldsymbol{\rho}} h\rangle + \langle v_{\boldsymbol{\rho}} h, T_1 v_{\boldsymbol{\rho}} h\rangle - \langle h, T_1 h\rangle \tag{11.2}$$

where $T_1 = \sqrt{p^2 - 1} + 1$.

Theorem 11.1 (Localization error estimate for one electron)**.** *For any fixed $\boldsymbol{\rho}_0 > 0$ there exists a constant $C < \infty$ depending on $\boldsymbol{\rho}_0$, such that for all $\boldsymbol{\rho} \geq \boldsymbol{\rho}_0$ and for all $h \in H^{1/2}(\mathbb{R}^3)$*

$$|\mathcal{LE}_1[h]| \leq C\Big(\frac{1}{\boldsymbol{\rho}^2}\|\chi_{\boldsymbol{\rho}} h\|^2 + e^{-\frac{\boldsymbol{\rho}}{64}}\|h\|^2\Big) \tag{11.3}$$

where

$$\chi_{\boldsymbol{\rho}}(z) := \begin{cases} 1 & \textit{if } 3\boldsymbol{\rho}/32 < |z| \leq 9\boldsymbol{\rho}/32 \\ 0 & \textit{elsewhere}\ . \end{cases} \tag{11.4}$$

Remark 11.2. *The first term in the r.h.s. of* (11.4) *corresponds to a region slightly bigger than the region where the derivatives of the cutoff functions* $u_{\boldsymbol{\rho}}$ *and* $v_{\boldsymbol{\rho}}$ *are supported. It is important that the localization error is exponentially small outside of this region.*

Proof. According to [28, Theorem 9] we have

$$\mathcal{LE}_1[h] = \int_{\mathbb{R}^6} \frac{K_2(|z_1-z_2|)}{4\pi^2|z_1-z_2|^2}\Big[\big(u_{\boldsymbol{\rho}}(z_1)-u_{\boldsymbol{\rho}}(z_2)\big)^2+\big(v_{\boldsymbol{\rho}}(z_1)-v_{\boldsymbol{\rho}}(z_2)\big)^2\Big]h(z_1)\overline{h(z_2)}dz_1dz_2. \tag{11.5}$$

This integral is invariant under permutation of z_1 and z_2, so we can consider only the region where $|z_2| \le |z_1|$. Let us define the following sets covering the region $|z_2| \le |z_1|$.

$$\begin{aligned}
A_1 &:= \big\{(z_1,z_2)\in\mathbb{R}^3\times\mathbb{R}^3 \big| |z_2|\le|z_1|, |z_1|\le \frac{\boldsymbol{\rho}}{8}\big\}\\
A_2 &:= \big\{(z_1,z_2)\in\mathbb{R}^3\times\mathbb{R}^3 \big| |z_2|\le|z_1|, \frac{\boldsymbol{\rho}}{8}<|z_1|\le\frac{9\boldsymbol{\rho}}{32}, |z_1-z_2|\le\frac{\boldsymbol{\rho}}{64}\big\}\\
A_3 &:= \big\{(z_1,z_2)\in\mathbb{R}^3\times\mathbb{R}^3 \big| |z_2|\le|z_1|, \frac{9\boldsymbol{\rho}}{32}<|z_1|, |z_1-z_2|\le\frac{\boldsymbol{\rho}}{64}\big\}\\
A_4 &:= \big\{(z_1,z_2)\in\mathbb{R}^3\times\mathbb{R}^3 \big| |z_2|\le|z_1|, \frac{\boldsymbol{\rho}}{8}<|z_1|, \frac{\boldsymbol{\rho}}{64}<|z_1-z_2|\big\}.
\end{aligned} \tag{11.6}$$

We define

$$I_{A_j} := \int_{A_j} \frac{K_2(|z_1-z_2|)}{4\pi^2|z_1-z_2|^2}\Big[\big(u_{\boldsymbol{\rho}}(z_1)-u_{\boldsymbol{\rho}}(z_2)\big)^2+\big(v_{\boldsymbol{\rho}}(z_1)-v_{\boldsymbol{\rho}}(z_2)\big)^2\Big]h(z_1)\overline{h(z_2)}dz_1dz_2. \tag{11.7}$$

For $(z_1,z_2)\in A_1$ we have $u_{\boldsymbol{\rho}}(z_1)=u_{\boldsymbol{\rho}}(z_2)=1$ and we find

$$I_{A_1}=0. \tag{11.8}$$

Next, by the scaling in the cutoff functions we have

$$|u_{\boldsymbol{\rho}}(z_1)-u_{\boldsymbol{\rho}}(z_2)| \le \frac{1}{\boldsymbol{\rho}}|z_1-z_2|\sup_{\xi}|\nabla u_{\boldsymbol{\rho}}(\xi)|,$$

and a similar inequality for $v_{\boldsymbol{\rho}}$. Using these estimates we obtain

$$|I_{A_2}| \le \frac{C}{\boldsymbol{\rho}^2}\int_{A_2} K_2(|z_1-z_2|)|h(z_1)|\ |h(z_2)|dz_1dz_2 \tag{11.9}$$

for some constant C. Note that $K_2(|\cdot|)\in L^1(\mathbb{R}^3)$ (for details see [1]). For $(z_1,z_2)\in A_2$ we have $\chi_{\boldsymbol{\rho}}(z_2)=1$ and $\frac{\boldsymbol{\rho}}{8}\le|z_1|\le\frac{9\boldsymbol{\rho}}{32}$ which, together with Young's convolution inequality, yields

$$\begin{aligned}
|I_{A_2}| &\le \frac{C}{\boldsymbol{\rho}^2}\int_{\boldsymbol{\rho}/8<|z_1|\le 9\boldsymbol{\rho}/32}|h(z_1)|\left(\int_{\mathbb{R}^3}K_2(|z_1-z_2|)\chi_{\boldsymbol{\rho}}(z_2)|h(z_2)|dz_2\right)dz_1\\
&\le \frac{C}{\boldsymbol{\rho}^2}\|K_2(|\cdot|)\|_{L^1}\|\chi_{\boldsymbol{\rho}}h\|^2.
\end{aligned} \tag{11.10}$$

Similar to the integral over the region A_1, we get

$$I_{A_3}=0. \tag{11.11}$$

In the region A_4 we use

$$\big(u_{\boldsymbol{\rho}}(z_1)-u_{\boldsymbol{\rho}}(z_2)\big)^2+\big(v_{\boldsymbol{\rho}}(z_1)-v_{\boldsymbol{\rho}}(z_2)\big)^2\le 2$$

to get

$$
\begin{aligned}
|I_{A_4}| &\leq \frac{1}{2\pi^2}\int_{A_4} \frac{K_2(|z_1-z_2|)}{|z_1-z_2|^2}|h(z_1)|\ |h(z_2)|dz_1dz_2 \\
&\leq \frac{1}{2\pi^2}\int_{\mathbb{R}^3} |h(z_1)|\Big[\Big(\frac{\mathbf{1}_{\{|\cdot|>\frac{\rho}{64}\}}K_2(|\cdot|)}{|\cdot|^2}\Big) * |h|\Big](z_1)dz_1.
\end{aligned}
\tag{11.12}
$$

Using Young's convolution inequality for the r.h.s. of (11.12) we get

$$
|I_{A_4}| \leq \frac{1}{2\pi^2}\|h\|^2 \int_{\mathbb{R}^3} \mathbf{1}_{\{|z_1|>\frac{\rho}{64}\}}\frac{K_2(|z_1|)}{|z_1|^2}dz_1. \tag{11.13}
$$

From [1, (9.7.2)] and the remark after [1, (9.7.4)], for large $r \in \mathbb{R}$

$$
K_2(r) \leq 10e^{-r}. \tag{11.14}
$$

For more detail on the asymptotic expansion of Bessel functions see [4, §4.8] and [4, (4.12.6)]. By (11.14) there exists a constant $C < \infty$ such that

$$
|I_{A_4}| \leq C\|h\|^2 e^{-\frac{\rho}{64}}. \tag{11.15}
$$

Collecting (11.8), (11.10), (11.11) and (11.15) concludes the proof. □

Theorem 11.1 can be used to estimate the localization error introduced by the partition of unity of the configuration space in the proof of Theorem 9.2 and Theorem 9.4. We will only prove this for the case of two nuclei (Theorem 9.2). The generalization to the case of M nuclei is straightforward.

11.1. **Multiparticle localization error estimate.** Let $\mathcal{D}_N^3$ be the collection of decompositions β of $\{1,\cdots,N\}$ into three clusters $(\mathcal{C}_0,\mathcal{C}_1,\mathcal{C}_2)$, with $\mathcal{C}_k \cap \mathcal{C}_l = \emptyset$ for all $k \neq l$ and $\bigcup_{k=0}^2 \mathcal{C}_k = \{1,\cdots,N\}$. We will assume that $\mathcal{C}_0$ contains particles far from both nuclei. Clusters $\mathcal{C}_1$ and $\mathcal{C}_2$ contain electrons localized near X_1 and X_2 respectively. For $z \in \mathbb{R}^3$, using $u_{\boldsymbol{\rho}}$ defined in (11.1) with $\boldsymbol{\rho} = |D|^{\frac{3}{4}}$, we set

$$
\begin{aligned}
w_1(z) &:= u_{|D|^{\frac{3}{4}}}(z-X_1) \\
w_2(z) &:= u_{|D|^{\frac{3}{4}}}(z-X_2)
\end{aligned}
\tag{11.16}
$$

and pick $w_0(z) \in C^1(\mathbb{R}^3;[0,1])$ such that $w_0^2+w_1^2+w_2^2 \equiv 1$. For $x \in \mathbb{R}^{3N}$ we define a family of cutoff functions $J_\beta \in C^1(\mathbb{R}^{3N};[0,1])$ with

$$
J_\beta(x) := \prod_{i\in\mathcal{C}_1} w_1(x_i)\prod_{j\in\mathcal{C}_2} w_2(x_j)\prod_{k\in\mathcal{C}_0} w_0(x_k). \tag{11.17}
$$

It is easy to see that these functions form a partition of unity, such that for all $x \in \mathbb{R}^{3N}$

$$
\sum_{\beta\in\mathcal{D}_N^3} J_\beta^2(x) = 1. \tag{11.18}
$$

The localization error for some state $\psi \in H^{1/2}(\mathbb{R}^{3N})$ and the partition of unity defined by the functions $J_\beta \in C^1(\mathbb{R}^{3N};[0,1])$ is given by

$$
\mathcal{LE}[\psi] := \sum_{\beta\in\mathcal{D}_N^3} \langle J_\beta\psi, HJ_\beta\psi\rangle - \langle\psi,H\psi\rangle. \tag{11.19}
$$

For $z \in \mathbb{R}^3$ we set

$$
\chi^{(D)}(z) := \begin{cases} 1 & \text{if } 3|D|^{\frac{3}{4}}/32 \leq |z| \leq 9|D|^{\frac{3}{4}}/32 \\ 0 & \text{elsewhere} \end{cases} \tag{11.20}
$$

and for $x \in \mathbb{R}^{3N}$ we define

$$
\chi^{(i)}(x) := \Big(\chi^{(D)}(x_i-X_1)+\chi^{(D)}(x_i-X_2)\Big). \tag{11.21}
$$

Theorem 11.3 (N electron localization error estimate)**.** *There exists $C > 0$ such that for any $\psi \in H^{1/2}(\mathbb{R}^{3N})$ we have*

$$|\mathcal{LE}[\psi]| \leq C \left(|D|^{-\frac{3}{2}} \sum_{i=1}^{N} \|\chi^{(i)}\psi\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}} \|\psi\|^2 \right). \tag{11.22}$$

For the proof of this theorem we need the following result.

Proposition 11.4. *Let $w_0, w_1, w_2 \in C^1(\mathbb{R}^3; [0,1])$ be as defined in* (11.16)*. Then there exists $D_0 > 0$ and a constant $C < \infty$, such that for all $|D| \geq D_0$ and for all $h \in H^{1/2}(\mathbb{R}^3)$*

$$\begin{aligned} &|\sum_{l=0}^{2} \langle w_l h, T_1 w_l h\rangle - \langle h, T_1 h\rangle| \\ &\leq C\Big(|D|^{-\frac{3}{2}} \Big\| \big(\chi^{(D)}(\cdot - X_1) + \chi^{(D)}(\cdot - X_2)\big) h \Big\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}} \|h\|^2 \Big) \end{aligned} \tag{11.23}$$

Proof. According to Theorem 11.1 we have

$$\begin{aligned} &\Big| \langle w_1 h, T_1 w_1 h\rangle + \Big\langle \sqrt{1-w_1^2} h, T_1 \sqrt{1-w_1^2} h \Big\rangle - \langle h, T_1 h\rangle \Big| \\ &\quad \leq C\Big(|D|^{-\frac{3}{2}} \big\|\chi^{(D)}(\cdot - X_1) h\big\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}} \|h\|^2 \Big). \end{aligned} \tag{11.24}$$

Applying Theorem 11.1 once more for $\sqrt{1-w_1^2} h \in H^{1/2}(\mathbb{R}^3)$ and the cutoff function $w_2 \in C^1(\mathbb{R}^3; [0,1])$ we get

$$\begin{aligned} &\Big| \Big\langle w_2 \sqrt{1-w_1^2} h, T_1 w_2 \sqrt{1-w_1^2} h \Big\rangle - \Big\langle \sqrt{1-w_1^2} h, T_1 \sqrt{1-w_1^2} h \Big\rangle \Big| \\ &\quad + \Big\langle \sqrt{(1-w_2^2)(1-w_1^2)} h, T_1 \sqrt{(1-w_2^2)(1-w_1^2)} h \Big\rangle \\ &\leq C\Big(|D|^{-\frac{3}{2}} \big\|\chi^{(D)}(\cdot - X_2) \sqrt{1-w_1^2} h\big\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}} \|\sqrt{1-w_1^2} h\|^2 \Big). \end{aligned} \tag{11.25}$$

Since $\mathrm{supp}(w_1) \cap \mathrm{supp}(w_2) = \emptyset$ we find

$$w_2 \sqrt{1-w_1} = w_2$$

and

$$\sqrt{(1-w_2^2)(1-w_1^2)} = w_0.$$

Notice that $\|\sqrt{1-w_1^2} h\|^2 \leq \|h\|^2$, together with (11.24) and (11.25) this yields the result. □

Proof of Theorem 11.3. The Coulomb potential, as a multiplicative operator, commutes with the functions J_β. The operator T_k only acts in the k-th particle variable, meaning that it commutes with functions $w_l(x_j)$ for $l = 0, 1, 2$ and $j \neq k$ and we have

$$\mathcal{LE}[\psi] = \sum_{i=1}^{N} \left(\sum_{l=0}^{2} \langle w_l(x_i)\psi, T_i w_l(x_i)\psi\rangle - \langle \psi, T_i \psi\rangle \right). \tag{11.26}$$

Applying Proposition 11.4 for each term on the r.h.s. yields the result. □

12. Proof of Theorem 9.2

12.1. **Estimate from below.** Let $\psi \in \mathcal{H}^\alpha$ with $\|\psi\| = 1$ and a_1, a_2 defined in (9.21) and (9.22). We have to show that there exists a constant $0 < C < \infty$ such that

$$\langle \psi, (H - \mu^\alpha)\psi\rangle \geq -\frac{a_1}{|D|^6} - \frac{a_2}{|D|^8} - \frac{C}{|D|^{10}}. \tag{12.1}$$

We decompose an arbitrary state $\psi \in \mathcal{H}^\alpha$ with respect to the partition of unity given by J_β defined in (11.17) according to the cluster decompositions in $\mathcal{D}_N^3$ to get

$$\langle \psi, (H-\mu^\alpha)\psi\rangle = \sum_{\beta\in\mathcal{D}_N^3} \langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle - \mathcal{LE}[\psi] \tag{12.2}$$

where $\mathcal{LE}[\psi]$ is the localization error defined in (11.19). By Theorem 11.3 there exists a constant $0 < C < \infty$ such that

$$-\mathcal{LE}[\psi] \geq -C\left(|D|^{-\frac{3}{2}}\sum_{i=1}^N \|\chi^{(i)}\psi\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}}\|\psi\|^2\right) \tag{12.3}$$

where $\chi^{(i)}$ was defined in (11.21). Let

$$L[J_\beta\psi] := \langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle - C\left(|D|^{-\frac{3}{2}}\sum_{i=1}^N \|\chi^{(i)}J_\beta\psi\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}}\|J_\beta\psi\|^2\right) \tag{12.4}$$

with C given by (12.3). According to (11.18) we have

$$\|\psi\|^2 = \sum_{\beta\in\mathcal{D}_N^3} \|J_\beta\psi\|^2 \tag{12.5}$$

and from (12.2),(12.3) and (12.4) we get

$$\langle \psi, (H-\mu^\alpha)\psi\rangle \geq \sum_{\beta\in\mathcal{D}_N^3} L[J_\beta\psi]. \tag{12.6}$$

Slightly abusing notation, we say $\beta = (\emptyset, \mathcal{C}_1, \mathcal{C}_2) \in \mathcal{D}^{at}$ if $\sharp\mathcal{C}_1 = Z_1$ and $\sharp\mathcal{C}_2 = Z_2$. From (12.6) we have

$$\sum_{\beta\in\mathcal{D}_N^3} L[J_\beta\psi] = \sum_{\beta\in\mathcal{D}^{at}} L[J_\beta\psi] + \sum_{\beta\in\mathcal{D}_N^3\setminus\mathcal{D}^{at}} L[J_\beta\psi]. \tag{12.7}$$

We start with estimating the second sum in the r.h.s. of (12.7).

For $\beta = (\mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2) \in \mathcal{D}_N^3$ we set

$$\begin{aligned} I_\beta := &\sum_{i\in\mathcal{C}_0\cup\mathcal{C}_1} \frac{-e^2 Z_2}{|x_i - X_2|} + \sum_{j\in\mathcal{C}_0\cup\mathcal{C}_2} \frac{-e^2 Z_1}{|x_j - X_1|} + \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} \frac{e^2}{|x_i - x_j|} \\ &+ \sum_{\substack{k\in\mathcal{C}_0\\ i\in\mathcal{C}_1\cup\mathcal{C}_2}} \frac{e^2}{|x_k - x_i|} + \frac{e^2 Z_1 Z_2}{|X_2 - X_1|} \end{aligned} \tag{12.8}$$

the sum of Coulomb interactions between particles belonging to different subsystems and let

$$H_\beta := H - I_\beta. \tag{12.9}$$

Then we can write

$$\langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle = \langle J_\beta\psi, (H_\beta-\mu^\alpha)J_\beta\psi\rangle + \langle J_\beta\psi, I_\beta J_\beta\psi\rangle. \tag{12.10}$$

12.1.1. *Non-neutral decompositions.* If β is a non-neutral cluster decomposition, i.e. $\beta \in \mathcal{D}_N^3 \setminus \mathcal{D}^{at}$, on the support of the function $J_\beta\psi$, the distances between a particle in subsystem 1 to a particle in subsystem 2 grows in $|D|$. The same is true for an electron in $\mathcal{C}_0$ and both of the nuclei.

Hence, since the interaction is small when the clusters are far apart, there exists $\epsilon_{|D|} > 0$ with $\epsilon_{|D|} \xrightarrow[|D|\to\infty]{} 0$ such that

$$\langle J_\beta\psi, I_\beta J_\beta\psi\rangle \geq -\epsilon_{|D|}\|J_\beta\psi\|^2. \tag{12.11}$$

As the next step, we find that for $\beta \in \mathcal{D}_N^3 \setminus \mathcal{D}^{at}$, for some $\delta > 0$ independent of ψ and $|D|$ we have

$$\langle J_\beta \psi, (H_\beta - \mu^\alpha) J_\beta \psi \rangle \geq \delta \|J_\beta \psi\|^2. \tag{12.12}$$

For $\mathcal{C}_0(\beta) = \emptyset$ the inequality (12.12) follows from Condition 1) in Theorem 9.2. If $\mathcal{C}_0(\beta) \neq \emptyset$, the inequality follows from the fact that for all irreducible representations of S_N, Hamiltonians of neutral atoms have discrete eigenvalues at the bottom of their spectrum (Theorem B.1). Removing an electron will increase the energy of the system, according to Theorem A.1. Combining (12.11) and (12.12) yields

$$\begin{aligned} L[J_\beta \psi] &\geq (\delta - \epsilon_{|D|}) \|J_\beta \psi\|^2 - C \left(|D|^{-\frac{3}{2}} \sum_{i=1}^N \|\chi^{(i)} J_\beta \psi\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}} \|J_\beta \psi\|^2 \right) \\ &\geq 0 \end{aligned} \tag{12.13}$$

choosing $|D|$ big enough. We can now begin to estimate the functionals $L[J_\beta \psi]$ for $\beta \in \mathcal{D}^{at}$.

12.1.2. *Neutral decompositions.* Let $\beta \in \mathcal{D}^{at}$, which implies $\sharp \mathcal{C}_1 = Z_1$ and $\sharp \mathcal{C}_2 = Z_2$. For this β and $\varphi, \psi \in \mathcal{H}^\alpha$ recall that the weighted bilinear form was defined as

$$\langle \varphi, \psi \rangle_1 := \langle \varphi, (\tilde{H}_\beta - \mu^\alpha) \psi \rangle \tag{12.14}$$

and the corresponding semi norm

$$\|\psi\|_1^2 := \langle \psi, \psi \rangle_1 \tag{12.15}$$

where $\tilde{H}_\beta$ was defined in (9.17). Let $\tilde{\mathcal{W}}_\beta^\alpha \subset \mathcal{H}^\alpha$ be the ground state space of $\tilde{H}_\beta^\alpha$ corresponding to μ^α. Note that $\tilde{\mathcal{W}}_\beta^\alpha \neq \emptyset$ by Theorem B.1. We project the function $\mathcal{U}_\beta J_\beta \psi$ onto the space $\tilde{\mathcal{W}}_\beta^\alpha$ with respect to the standard $L^2(\mathbb{R}^{3N})$-inner product where $\mathcal{U}_\beta$ was defined in (9.16). For some $\gamma_1 \in \mathbb{C}$ with $|\gamma_1| \leq 1$ and $\phi \in \tilde{\mathcal{W}}_\beta^\alpha$ with $\|\phi\| = 1$ we get

$$\mathcal{U}_\beta J_\beta \psi = \gamma_1 \phi + G. \tag{12.16}$$

As the next step we project G in the sense of the bilinear form $\langle \cdot, \cdot \rangle_1$ consecutively onto the functions

$$\phi_2 := (\tilde{H}_\beta - \mu^\alpha)^{-1} f_2 \phi \tag{12.17}$$

and

$$\phi_3 := (\tilde{H}_\beta - \mu^\alpha)^{-1} f_3 \phi. \tag{12.18}$$

We will prove in Lemma E.2 that the function ϕ, because of its rotational symmetry, is orthogonal to $f_2 \phi$ and $f_3 \phi$ with respect to the standard L^2-inner product, which ensures that the functions ϕ_2 and ϕ_3 are well defined. Furthermore we show in Corollary E.4 that ϕ, ϕ_2, ϕ_3 are mututally orthogonal with respect to the bilinear form $\langle \cdot, \cdot \rangle_1$. After this decomposition we have

$$J_\beta \psi = \mathcal{U}_\beta^* \big(\gamma_1 \phi + |D|^{-3} \gamma_2 \phi_2 + |D|^{-4} \gamma_3 \phi_3 + g \big), \tag{12.19}$$

where

$$\langle \phi, g \rangle = \langle g, \phi_2 \rangle_1 = \langle g, \phi_3 \rangle_1 = 0. \tag{12.20}$$

By definition of the functions ϕ, ϕ_2, ϕ_3, and g and their orthogonality with respect to $\langle \cdot, \cdot \rangle_1$ we have

$$\begin{aligned} \langle J_\beta \psi, (H_\beta - \mu^\alpha) J_\beta \psi \rangle &= \langle \mathcal{U}_\beta J_\beta \psi, (\mathcal{U}_\beta H_\beta \mathcal{U}_\beta^* - \mu^\alpha) \mathcal{U}_\beta J_\beta \psi \rangle \\ &= \langle \mathcal{U}_\beta J_\beta \psi, (\tilde{H}_\beta - \mu^\alpha) \mathcal{U}_\beta J_\beta \psi \rangle \\ &= \frac{|\gamma_2|^2}{|D|^6} \|\phi_2\|_1^2 + \frac{|\gamma_3|^2}{|D|^8} \|\phi_3\|_1^2 + \|g\|_1^2. \end{aligned} \tag{12.21}$$

Now we turn to the term with the intercluster interaction I_β. In Lemma D.5 we prove that for any $\delta > 0$ there exist $C > 0$ such that for $|D|$ sufficiently big

$$\begin{aligned}\langle J_\beta\psi, I_\beta J_\beta\psi\rangle \geq\ & 2|D|^{-6}\operatorname{Re}\gamma_1\overline{\gamma_2}\|\phi_2\|_1^2 + 2|D|^{-8}\operatorname{Re}\gamma_1\overline{\gamma_3}\|\phi_3\|_1^2 \\ & - C\frac{|\gamma_1|^2+|\gamma_2|^2+|\gamma_3|^2}{|D|^{10}} - \delta\|g\|^2.\end{aligned} \tag{12.22}$$

Summing (12.21) and (12.22) we arrive at

$$\begin{aligned}\langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle \geq\ & \frac{|\gamma_2|^2+2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^6}\|\phi_2\|_1^2 + \frac{|\gamma_3|^2+2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^8}\|\phi_3\|_1^2 \\ & - C\frac{|\gamma_1|^2+|\gamma_2|^2+|\gamma_3|^2}{|D|^{10}} - \delta\|g\|^2 + \|g\|_1^2.\end{aligned} \tag{12.23}$$

Let κ be the distance between ground state energy and the next higher eigenvalue of $\tilde{H}_\beta$. By Theorem B.1 $\kappa > 0$ and, since g is orthogonal to $\tilde{\mathcal{W}}_\beta^\alpha$, we have $\|g\|_1^2 = \langle g, (\tilde{H}_\beta - \mu^\alpha)g\rangle \geq \kappa\|g\|^2$. Taking $\delta < \frac{\kappa}{2}$ we get

$$\|g\|_1^2 - \delta\|g\|^2 \geq \frac{\kappa}{2}\|g\|^2. \tag{12.24}$$

Note that

$$|\gamma_2|^2 + 2\operatorname{Re}\gamma_1\overline{\gamma_2} = |\gamma_1+\gamma_2|^2 - |\gamma_1|^2 \tag{12.25}$$

and

$$|\gamma_3|^2 + 2\operatorname{Re}\gamma_1\overline{\gamma_3} = |\gamma_1+\gamma_3|^2 - |\gamma_1|^2. \tag{12.26}$$

Summing the bound for $\langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle$ yields

$$\begin{aligned}\langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle \geq\ & \frac{-|\gamma_1|^2+|\gamma_1+\gamma_2|^2}{|D|^6}\|\phi_2\|_1^2 + \frac{-|\gamma_1|^2+|\gamma_1+\gamma_3|^2}{|D|^8}\|\phi_3\|_1^2 \\ & - \frac{C(|\gamma_1|^2+|\gamma_2|^2+|\gamma_3|^2)}{|D|^{10}} + \frac{\kappa}{2}\|g\|^2.\end{aligned} \tag{12.27}$$

We now minimize the expression on the r.h.s. of (12.27) with respect to γ_2 and γ_3. We aim to show that for $|D|$ large enough, minimization in γ_2 yields

$$\frac{|\gamma_1+\gamma_2|^2}{|D|^6}\|\phi_2\|_1^2 - C\frac{|\gamma_2|^2}{|D|^{10}} \geq -\frac{4C|\gamma_1|^2}{|D|^{10}}. \tag{12.28}$$

Assume that $|\gamma_2| > 2|\gamma_1|$, then

$$\frac{|\gamma_1+\gamma_2|^2}{|D|^6}\|\phi_2\|_1^2 - \frac{C|\gamma_2|^2}{|D|^{10}} > \frac{\frac{1}{4}|\gamma_2|^2}{|D|^6}\|\phi_2\|_1^2 - \frac{C|\gamma_2|^2}{|D|^{10}} \tag{12.29}$$

which is positive for large $|D|$.

Whereas for $|\gamma_2| \leq 2|\gamma_1|$ we have

$$\frac{|\gamma_1+\gamma_2|^2}{|D|^6}\|\phi_2\|_1^2 - \frac{C|\gamma_2|^2}{|D|^{10}} \geq -\frac{4C|\gamma_1|^2}{|D|^{10}} \tag{12.30}$$

which is obviously smaller than the expression on the r.h.s. of (12.29). Minimizing similarly in γ_3, for $|D|$ large enough we get

$$\frac{|\gamma_1+\gamma_3|^2}{|D|^8}\|\phi_3\|_1^2 - \frac{C|\gamma_3|^2}{|D|^{10}} \geq -\frac{4C|\gamma_1|^2}{|D|^{10}}. \tag{12.31}$$

Plugging (12.30) and (12.31) into (12.27), taking into account that $|\gamma_1|^2 \leq \|J_\beta\psi\|^2$ we arrive at

$$\langle J_\beta\psi, (H-\mu^\alpha)J_\beta\psi\rangle \geq \left(-\frac{\|\phi_2\|_1^2}{|D|^6} - \frac{\|\phi_3\|_1^2}{|D|^8} - C|D|^{-10}\right)\|J_\beta\psi\|^2 + \frac{\kappa}{2}\|g\|^2. \tag{12.32}$$

Now we turn to the estimate of the term coming from the localization error, that is,

$$-C\left(|D|^{-\frac{3}{2}}\sum_{i=1}^{N}\|\chi^{(i)}J_\beta\psi\|^2+e^{-\frac{1}{64}|D|^{\frac{3}{4}}}\|J_\beta\psi\|^2\right). \tag{12.33}$$

The second term of this expression is exponentially small. For the first term we have

$$\|\chi^{(i)}J_\beta\psi\|^2\le 2\Big\|\chi^{(i)}\mathcal{U}_\beta^*\Big(\gamma_1\phi+\frac{\gamma_2}{|D|^3}\phi_2+\frac{\gamma_3}{|D|^4}\phi_3\Big)\Big\|^2+2\|\chi^{(i)}\mathcal{U}_\beta^*g\|^2. \tag{12.34}$$

The functions $\chi^{(i)}$ are supported in the region where $|x_i-X_l|\ge\frac{3}{32}|D|^{-\frac{3}{4}}$ for $l=1,2$. According to Corollary 10.4, ϕ,ϕ_2,and ϕ_3 are exponentially decaying, for normalized ψ we get

$$\|\chi^{(i)}J_\beta\psi\|^2\le 2\|\chi^{(i)}\mathcal{U}_\beta^*g\|^2+\mathcal{O}(e^{-|D|^{\frac{1}{2}}}). \tag{12.35}$$

Thus

$$\begin{aligned}&-C\left(|D|^{-\frac{3}{2}}\sum_{i=1}^{N}\|\chi^{(i)}J_\beta\psi\|^2+e^{-\frac{1}{64}|D|^{\frac{3}{4}}}\|J_\beta\psi\|^2\right)\\&\ge -C|D|^{-\frac{3}{2}}\sum_{i=1}^{N}\|\chi^{(i)}\mathcal{U}_\beta^*g\|^2+\mathcal{O}(e^{-|D|^{\frac{1}{2}}})\\&\ge -C|D|^{-\frac{3}{2}}N\|g\|^2+\mathcal{O}(e^{-|D|^{\frac{1}{2}}}).\end{aligned} \tag{12.36}$$

Substituting this into (12.4), together with the estimate for $\langle J_\beta\psi,(H-\mu^\alpha)J_\beta\psi\rangle$ in (12.32) we get

$$L[J_\beta\psi]\ge\left(-\frac{a_1}{|D|^6}-\frac{a_2}{|D|^8}-C|D|^{-10}\right)\|J_\beta\psi\|^2+\left(-CN|D|^{-\frac{3}{2}}+\frac{\kappa}{2}\right)\|g\|^2. \tag{12.37}$$

For $|D|$ sufficiently large, the last term is positive and we arrive at

$$L[J_\beta\psi]\ge\left(-\frac{a_1}{|D|^6}-\frac{a_2}{|D|^8}-C|D|^{-10}\right)\|J_\beta\psi\|^2. \tag{12.38}$$

This inequality is true for any $\beta\in\mathcal{D}^{at}$. Recall from (12.6) the bound

$$\langle\psi,(H-\mu^\alpha)\psi\rangle\ge\sum_{\beta\in\mathcal{D}_N^3}L[J_\beta\psi]. \tag{12.39}$$

By (12.13) for all $\beta\in\mathcal{D}_N^3\setminus\mathcal{D}^{at}$

$$L[J_\beta\psi]\ge 0. \tag{12.40}$$

Since $\sum_{\beta\in\mathcal{D}^{at}}\|J_\beta\psi\|^2\le\|\psi\|^2=1$, gathering (12.6), (12.13), and (12.38) we obtain

$$\langle\psi,(H-\mu^\alpha)\psi\rangle\ge-\frac{a_1}{|D|^6}-\frac{a_2}{|D|^8}-\frac{C}{|D|^{10}}. \tag{12.41}$$

12.2. **Estimate from above.** We aim to construct a trial function $\psi_0\in\mathcal{H}^\alpha$ with $\|\psi_0\|=1$ such that

$$\langle\psi_0,(H-\mu^\alpha)\psi_0\rangle\le-\frac{a_1}{|D|^6}-\frac{a_2}{|D|^8}-\frac{C}{|D|^{10}} \tag{12.42}$$

where a_1 and a_2 are defined in (9.21) and (9.22). Fix some $\beta\in\mathcal{D}^{at}$, $\tilde{\mathcal{W}}_\beta^\alpha$ denotes the ground state space of $\tilde{H}_\beta^\alpha$. By permutation symmetry of $\tilde{H}_\beta$ we have

$$\tilde{\mathcal{W}}_\beta^\alpha=\bigoplus_{\alpha'(\beta)\prec\alpha}P^{\alpha'(\beta)}\tilde{\mathcal{W}}_\beta^\alpha, \tag{12.43}$$

thus there is at least one $\alpha^*(\beta) \prec \alpha$ such that there exists $\phi \in P^{\alpha^*(\beta)}\tilde{\mathcal{W}}_\beta^\alpha$ that realises the maxima a_1 and a_2 with $\|\phi\| = 1$. For such a $\phi \in P^{\alpha^*(\beta)}\tilde{\mathcal{W}}_\beta^\alpha$ we set

$$\begin{aligned}\hat{\psi}_0 :&= \mathcal{U}_\beta^*\Big(\phi - \frac{(\tilde{H}_\beta - \mu^\alpha)^{-1} f_2 \phi}{|D|^3} - \frac{(\tilde{H}_\beta - \mu^\alpha)^{-1} f_3 \phi}{|D|^4}\Big) \\ &= \mathcal{U}_\beta^*(\phi - \frac{\phi_2}{|D|^3} - \frac{\phi_3}{|D|^4}),\end{aligned} \tag{12.44}$$

by definition of ϕ_2 and ϕ_3 in (12.17) and (12.18). With P^α being the projection onto $\mathcal{H}^\alpha$ and the cutoff function J_β defined in (11.17), we define the trial state as

$$\psi_0 := \frac{P^\alpha J_\beta \hat{\psi}_0}{\|P^\alpha J_\beta \hat{\psi}_0\|}. \tag{12.45}$$

12.2.1. As a first step, we will show

$$\langle \psi_0, (H - \mu^\alpha)\psi_0 \rangle = \frac{\langle J_\beta \hat{\psi}_0, (H - \mu^\alpha) J_\beta \hat{\psi}_0 \rangle}{\|J_\beta \hat{\psi}_0\|^2}. \tag{12.46}$$

Let $\chi^\alpha_{\pi^{-1}}$ denote the character of the element $\pi^{-1} \in S_N$ in the representation α. For $\mathcal{T}_\pi$ defined in (9.4), and $|\alpha|$ denoting the dimension of the irreducible representation α, by [18, p. 113] the projection operator onto $\mathcal{H}^\alpha$ is given by

$$P^\alpha = \frac{|\alpha|}{N!} \sum_{\pi \in S_N} \chi^\alpha_{\pi^{-1}} \mathcal{T}_\pi. \tag{12.47}$$

Following [50] we write the r.h.s. of (12.47) as two sums. In the first sum we collect the permutations which only permute particles within the subsystems of β. The second sum contains permutation which change at least one pair of particles belonging to different subsystems of β. We get

$$P^\alpha = \frac{|\alpha|}{N!} \sum_{\pi \in S_\beta} \chi^\alpha_{\pi^{-1}} \mathcal{T}_\pi + \frac{|\alpha|}{N!} \sum_{\pi \in S_N \setminus S_\beta} \chi^\alpha_{\pi^{-1}} \mathcal{T}_\pi. \tag{12.48}$$

For $\alpha'(\beta) \prec \alpha$ we set

$$\theta_{\alpha'(\beta)} := \frac{Z_1! Z_2!}{N!} \frac{|\alpha|}{|\alpha'(\beta)|}. \tag{12.49}$$

Note that for $\pi \in S_\beta$

$$\chi^\alpha_\pi = \sum_{\alpha'(\beta) \prec \alpha} \chi^{\alpha'(\beta)}_\pi \tag{12.50}$$

and

$$P^{\alpha'(\beta)} = \frac{|\alpha'(\beta)|}{Z_1! Z_2!} \sum_{\pi \in S_\beta} \chi^{\alpha'(\beta)}_{\pi^{-1}} \mathcal{T}_\pi. \tag{12.51}$$

Let us define

$$P_1^\alpha := \sum_{\alpha'(\beta) \prec \alpha} \theta_{\alpha'(\beta)} P^{\alpha'(\beta)} \tag{12.52}$$

and

$$P_2^\alpha := \frac{|\alpha|}{N!} \sum_{\pi \in S_N \setminus S_\beta} \chi^\alpha_{\pi^{-1}} \mathcal{T}_\pi. \tag{12.53}$$

Then following [50] we rewrite (12.48) as

$$P^\alpha = P_1^\alpha + P_2^\alpha. \tag{12.54}$$

To prove (12.46) we first compute $\|P^\alpha J_\beta \hat{\psi}_0\|^2$. Since $(P^\alpha)^2 = P^\alpha$, by (12.54) we have

$$\begin{aligned}\|P^\alpha J_\beta \hat{\psi}_0\|^2 &= \big\langle (P_1^\alpha + P_2^\alpha) J_\beta \hat{\psi}_0, J_\beta \hat{\psi}_0 \big\rangle \\ &= \sum_{\alpha'(\beta)\prec\alpha} \theta_{\alpha'(\beta)} \langle P^{\alpha'(\beta)} J_\beta \hat{\psi}_0, J_\beta \hat{\psi}_0\rangle + \frac{|\alpha|}{N!} \sum_{\pi\in S_N\setminus S_\beta} \chi^\alpha_{\pi^{-1}} \langle \mathcal{T}_\pi J_\beta \hat{\psi}_0, J_\beta \hat{\psi}_0\rangle.\end{aligned} \tag{12.55}$$

The function J_β is invariant under permutations in S_β, thus $J_\beta \hat{\psi}_0$ belongs to the same symmetry type $\alpha^*(\beta)$ as the function ϕ. The projectors $P^{\alpha'(\beta)}$ are mutually orthogonal for different $\alpha'(\beta)$. Hence for the first term on the r.h.s. of (12.55) we get

$$\sum_{\alpha'(\beta)\prec\alpha} \theta_{\alpha'(\beta)} \langle P^{\alpha'(\beta)} J_\beta \hat{\psi}_0, J_\beta \hat{\psi}_0\rangle = \theta_{\alpha^*(\beta)} \|J_\beta \hat{\psi}_0\|^2. \tag{12.56}$$

The last sum on the r.h.s. of (12.55) is zero, as the functions $\mathcal{T}_\pi J_\beta \hat{\psi}_0$ and $J_\beta \hat{\psi}_0$ are supported on different domains (for details see Appendix F). Thus

$$\|P^\alpha J_\beta \hat{\psi}_0\|^2 = \theta_{\alpha^*(\beta)} \|J_\beta \hat{\psi}_0\|^2. \tag{12.57}$$

Note that (12.49) implies $\theta_{\alpha^*(\beta)} \neq 0$, which yields, in particular, $P^\alpha J_\beta \hat{\psi}_0 \neq 0$.

As the next step we would like to show

$$\langle P^\alpha J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle = \theta_{\alpha^*(\beta)} \langle J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle. \tag{12.58}$$

To this end we split P^α as in (12.54) and get

$$\begin{aligned}&\langle P^\alpha J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle \\ &= \langle P_1^\alpha J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle + \langle P_2^\alpha J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle.\end{aligned} \tag{12.59}$$

Let us show that the second term on the r.h.s. of (12.59) is zero. Since for all $\pi \in S_\beta$, $\mathcal{T}_\pi J_\beta \hat{\psi}_0$ and $I_\beta J_\beta \hat{\psi}_0$ have disjoint support

$$\langle P_2^\alpha J_\beta \hat{\psi}_0, I_\beta J_\beta \hat{\psi}_0\rangle = 0. \tag{12.60}$$

Furthermore H_β is the sum of two operators

$$H_\beta = \mathcal{U}_\beta^* \tilde{H}^{Z_1}_{\mathcal{C}_1} \mathcal{U}_\beta + \mathcal{U}_\beta^* \tilde{H}^{Z_2}_{\mathcal{C}_2} \mathcal{U}_\beta. \tag{12.61}$$

The first operator acts only on particles in $\mathcal{C}_1$ and the second operator acts only on particles in $\mathcal{C}_2$. The localization function J_β is supported in the region, where particles in $\mathcal{C}_1$ are located near X_1 and particles in $\mathcal{C}_2$ are near X_2 with distances to the corresponding nucleus X_1 and X_2 much smaller than $|D| = |X_1 - X_2|$. We can apply Lemma F.1 to see

$$\langle P_2^\alpha J_\beta \hat{\psi}_0, \mathcal{U}_\beta^* \tilde{H}^{Z_1}_{\mathcal{C}_1} \mathcal{U}_\beta J_\beta \hat{\psi}_0\rangle = 0 \tag{12.62}$$

and

$$\langle P_2^\alpha J_\beta \hat{\psi}_0, \mathcal{U}_\beta^* \tilde{H}^{Z_2}_{\mathcal{C}_2} \mathcal{U}_\beta J_\beta \hat{\psi}_0\rangle = 0 \tag{12.63}$$

since the respective functions have disjoint support (see Appendix F). Equalities (12.60), (12.62) and (12.63) imply

$$\langle P_2^\alpha J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle = 0. \tag{12.64}$$

Now we turn to the first term on the r.h.s. of (12.59). The operators $(H_\beta - \mu^\alpha)$ and I_β are invariant under permutations in S_β, thus $(H-\mu^\alpha) J_\beta \hat{\psi}_0$ belongs to the representation $\alpha^*(\beta)$. By orthogonality of functions belonging to different irreducible representations, we get

$$\langle P_1^\alpha J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle = \theta_{\alpha^*(\beta)} \langle J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0\rangle. \tag{12.65}$$

This proves (12.46).

12.2.2. Our next goal is to estimate

$$\langle J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0 \rangle. \tag{12.66}$$

We substitute $H = H_\beta + I_\beta$ to get

$$\langle J_\beta \hat{\psi}_0, (H-\mu^\alpha) J_\beta \hat{\psi}_0 \rangle = \langle J_\beta \hat{\psi}_0, (H_\beta-\mu^\alpha) J_\beta \hat{\psi}_0 \rangle + \langle J_\beta \hat{\psi}_0, I_\beta J_\beta \hat{\psi}_0 \rangle. \tag{12.67}$$

For the first term on the r.h.s. of (12.67) we write

$$\begin{aligned} &\langle J_\beta \hat{\psi}_0, (H_\beta-\mu^\alpha) J_\beta \hat{\psi}_0 \rangle \\ &\quad = \langle \hat{\psi}_0, (H_\beta-\mu^\alpha) \hat{\psi}_0 \rangle - \langle \sqrt{1-J_\beta^2}\hat{\psi}_0, (H_\beta-\mu^\alpha)\sqrt{1-J_\beta^2}\hat{\psi}_0 \rangle + \mathcal{LE}_\beta[\hat{\psi}_0], \end{aligned} \tag{12.68}$$

where $\mathcal{LE}[\hat{\psi}_0]$ is the localization error coming from the partition of unity with cutoff functions J_β and $\sqrt{1-J_\beta^2}$. Similar to Theorem 11.3, this can be estimated as

$$|\mathcal{LE}_\beta[\hat{\psi}_0]| \leq C \left(|D|^{-\frac{3}{2}} \sum_{i=1}^N \|\chi^{(i)}\hat{\psi}_0\|^2 + e^{-\frac{1}{64}|D|^{\frac{3}{4}}} \|\hat{\psi}_0\|^2 \right). \tag{12.69}$$

By Proposition 10.3 the function $\hat{\psi}_0$ decays exponentially, which implies

$$|\mathcal{LE}_\beta[\hat{\psi}_0]| = \mathcal{O}(e^{-|D|^{\frac{1}{2}}}). \tag{12.70}$$

The operator $(H_\beta - \mu^\alpha)$ is semibounded from below, thus for some constant $C > 0$ we get

$$\langle \sqrt{1-J_\beta^2}\hat{\psi}_0, (H_\beta-\mu^\alpha)\sqrt{1-J_\beta^2}\hat{\psi}_0 \rangle \geq -C\|\sqrt{1-J_\beta^2}\hat{\psi}_0\|^2 \geq -Ce^{-|D|^{\frac{1}{2}}} \tag{12.71}$$

taking into account exponential decay of $\hat{\psi}_0$. This together with (12.68) yields

$$\langle J_\beta \hat{\psi}_0, (H_\beta-\mu^\alpha) J_\beta \hat{\psi}_0 \rangle \leq \langle \hat{\psi}_0, (H_\beta-\mu^\alpha) \hat{\psi}_0 \rangle + Ce^{-|D|^{\frac{1}{2}}}. \tag{12.72}$$

Once again, by exponential decay of $\hat{\psi}_0$

$$\|J_\beta \hat{\psi}_0\|^2 = \|\hat{\psi}_0\|^2 + \mathcal{O}(e^{-|D|^{\frac{1}{2}}}) \tag{12.73}$$

and since ϕ is orthogonal to ϕ_2 and ϕ_3, we get

$$\|\hat{\psi}_0\|^2 = 1 + \mathcal{O}(|D|^{-6}). \tag{12.74}$$

Combining (12.72), (12.73) and (12.74) yields

$$\begin{aligned} \|J_\beta \hat{\psi}_0\|^{-2} \langle J_\beta \hat{\psi}_0, (H_\beta-\mu^\alpha) J_\beta \hat{\psi}_0 \rangle &= \langle \hat{\psi}_0, (H_\beta-\mu^\alpha) \hat{\psi}_0 \rangle (1+\mathcal{O}(|D|^{-6}) \\ &= (\|\phi_2\|_1^2 + \|\phi_3\|_1^2)(1+\mathcal{O}(|D|^{-6})). \end{aligned} \tag{12.75}$$

Applying (12.75) in (12.46) we get

$$\langle \psi_0, (H-\mu^\alpha)\psi_0 \rangle = \big(\langle J_\beta \hat{\psi}_0, I_\beta J_\beta \hat{\psi}_0 \rangle + \|\phi_2\|_1^2 + \|\phi_3\|_1^2 \big)\big(1+\mathcal{O}(|D|^{-6})\big). \tag{12.76}$$

Similar to the estimates done in Lemma D.5, with simplifications coming from the fact that we have $\gamma_1 = 1$, $\gamma_2 = \gamma_3 = -1$ and $g = 0$, we obtain

$$\langle J_\beta \hat{\psi}_0, I_\beta J_\beta \hat{\psi}_0 \rangle = -2\|\phi_2\|_1^2 - 2\|\phi_3\|_1^2 + \mathcal{O}(|D|^{-10}) \tag{12.77}$$

which completes the proof of Theorem 9.2.

13. Proof of Theorem 9.4

The proof of Theorem 9.4 is very similar to the proof of Theorem 9.2.

We start with the estimate from below. Define cluster decompositions $\beta_M = \{\mathcal{C}_0, \cdots, \mathcal{C}_M\}$ into $M+1$ clusters, such that particles which are far from all nuclei belong to the subsystem $\mathcal{C}_0$. As the next step we define the cutoff functions J_{β_M} corresponding to the cluster decompositions β_M. The estimate of the localization error is not different from the diatomic case.

Similar to the proof of Theorem 9.2 one can show that if β_M is not a decomposition into M neutral atoms, for $\psi \in \mathcal{H}^\alpha$ we have the inequality

$$\langle J_{\beta_M}\psi, (H - \mu_M^\alpha) J_{\beta_M}\psi\rangle > 0.$$

Now we turn to the estimate of the quadratic form $\langle J_{\beta_M}\psi, (H-\mu^\alpha)J_{\beta_M}\psi\rangle$ for decompositions β_M corresponding to M neutral atoms (c.f. Section 12.1.2).

We defined $\tilde{H}_{\beta_M}, \tilde{\mathcal{W}}^\alpha_{\beta_M}$ and functions $f_2^{(k,l)}, f_3^{(k,l)}$ in equations (9.29)- (9.36). Let $\mathcal{U}_{\beta_M}$ be the shift operator defined analogous to $\mathcal{U}_\beta$ in (9.16). Similar to (12.19) we write

$$J_{\beta_M}\psi = \mathcal{U}^*_{\beta_M}\big(\gamma_1\phi + d^{-3}\gamma_2\phi_2 + d^{-4}\gamma_3\phi_3 + g\big). \tag{13.1}$$

where $\phi \in \mathcal{W}^\alpha_{\beta_M}$ and the functions ϕ_2, ϕ_3 are given by

$$\phi_2 = (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1}\sum_{k<l}|D_{k,l}|^{-3} f_2^{(k,l)}\phi \tag{13.2}$$

and

$$\phi_3 = (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1}\sum_{k<l}|D_{k,l}|^{-4} f_3^{(k,l)}\phi. \tag{13.3}$$

Note that by the same reasons as in the diatomic case we have

$$\langle\phi,\phi_2\rangle = \langle\phi,\phi_2\rangle_1 = \langle\phi,\phi_3\rangle = \langle\phi,\phi_3\rangle_1 = \langle\phi_2,\phi_3\rangle = \langle\phi_2,\phi_3\rangle_1 = 0. \tag{13.4}$$

With the above definitions we get the same epression as (12.21) for the expected value of $(\tilde{H}_{\beta_M} - \mu^\alpha_{\beta_M})$.

We now estimate the expectation value of the interaction I_{β_M} of particles belonging to dofferent clusters $\langle J_{\beta_M}\psi, I_{\beta_M}J_{\beta_M}\psi\rangle$. Our goal is to generalize the estimate (12.22), which is proven in Lemma D.5, to the case of M atoms. Let χ_{β_M} be the characteristic function of the support of J_{β_M} and let

$$I^o_{\beta_M} := I_{\beta_M}\chi_{\beta_M}. \tag{13.5}$$

Note that

$$\begin{aligned}
\langle J_{\beta_M}\psi, I_{\beta_M}J_{\beta_M}\psi\rangle &= \langle J_{\beta_M}\psi, I^o_{\beta_M}J_{\beta_M}\psi\rangle \\
&= |\gamma_1|^2\langle\mathcal{U}^*_{\beta_M}\phi, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi\rangle + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{d^3}\langle\mathcal{U}^*_{\beta_M}\phi_2, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi\rangle \\
&\quad + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{d^4}\langle\mathcal{U}^*_{\beta_M}\phi_3, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi\rangle + \frac{|\gamma_2|^2}{d^6}\langle\mathcal{U}^*_{\beta_M}\phi_2, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi_2\rangle \\
&\quad + \frac{2\operatorname{Re}\gamma_2\overline{\gamma_3}}{d^7}\langle\mathcal{U}^*_{\beta_M}\phi_3, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi_2\rangle + \frac{|\gamma_3|^2}{d^8}\langle\mathcal{U}^*_{\beta_M}\phi_3, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi_3\rangle \\
&\quad + 2\operatorname{Re}\gamma_1\langle\mathcal{U}^*_{\beta_M}g, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi\rangle + \frac{2\operatorname{Re}\gamma_2}{d^3}\langle\mathcal{U}^*_{\beta_M}g, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi_2\rangle \\
&\quad + \frac{2\operatorname{Re}\gamma_3}{d^4}\langle\mathcal{U}^*_{\beta_M}g, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}\phi_3\rangle + \langle\mathcal{U}_{\beta_M}g, I^o_{\beta_M}\mathcal{U}_{\beta_M}g\rangle \\
&= B_1^M + B_2^M + B_3^M + \langle\mathcal{U}^*_{\beta_M}g, I^o_{\beta_M}\mathcal{U}^*_{\beta_M}g\rangle,
\end{aligned} \tag{13.6}$$

where B_1^M contains the first three terms of the r.h.s. of (13.6), B_2^M the second triple and B_3^M third triple on the r.h.s. of (13.6). We define analogously to the diatomic case the functions $f_4^{(k,l)}, f_5^{(k,l)}$, see Appendix D. Let $x \in \mathbb{R}^{3N}$ and $|\cdot|$ denote the

standard norm in this space. On the support of J_{β_M} we have $|x| < C(D_0 d)^{\frac{3}{4}}$ with $D_0 = \min_{k,l} |D_{k,l}|$ and some constant C. We can expand $I^o_{\beta_M}$ as a Taylor series for large d arriving at

$$|I^o_{\beta_M} - \sum_{k\neq l} \frac{\mathcal{U}^*_{\beta_M} f_2^{(k,l)}}{2|D_{k,l}|^3 d^3} - \sum_{k\neq l} \frac{\mathcal{U}^*_{\beta_M} f_3^{(k,l)}}{2|D_{k,l}|^4 d^4} - \sum_{k\neq l} \frac{\mathcal{U}^*_{\beta_M} f_4^{(k,l)}}{2|D_{k,l}|^5 d^5} - \sum_{k\neq l} \frac{\mathcal{U}^*_{\beta_M} f_5^{(k,l)}}{2|D_{k,l}|^6 d^6}| \leq C \frac{|x|^6}{(D_0 d)^7}. \tag{13.7}$$

As the first step, we note that for B_1^M, similar to Proposition D.6 we have,

$$B_1^M \geq \frac{2\,\mathrm{Re}\,\gamma_1\overline{\gamma_2}}{d^6}\|\phi_2\|_1^2 + \frac{2\,\mathrm{Re}\,\gamma_1\overline{\gamma_3}}{d^8}\|\phi_3\|_1^2 - C\frac{|\gamma_1|^2 + |\gamma_2|^2 + |\gamma_3|^2}{d^{10}}. \tag{13.8}$$

To prove (13.8) we substitute (13.7) into the expression for $I^o_{\beta_M}$ in B_1 and follow the same steps as in the proof of Proposition D.6, replacing orthogonality relations from Lemma E.6 with the following proposition.

Proposition 13.1. *Let Condition 2') of Theorem 9.4 be fulfilled. Then for $n, m = 2, 3, 4, 5,\ n \neq m$ and all $k, l = 1, \cdots, M,\ k \neq l$ we have*

$$\langle \phi_m, f_n^{(k,l)}\phi\rangle = 0 \tag{13.9}$$

where

$$\phi_m = (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} \sum_{k\neq l} \frac{f_m^{(k,l)}\phi}{2|D_{k,l}|^{m+1}} \tag{13.10}$$

Proof. By Condition 2'), the state ϕ belongs to the irreducible representation of the $SO(3)$ group corresponding to the degree $\ell = 0$. The functions $f_n^{(k,l)}$ belong to the irreducible representation of the $SO(3)$ group corresponding to the degree $\ell = n$, see the proof of Lemma E.7. Consequently, ϕ_m and $f_n^{(k,l)}\phi$ are orthogonal as two functions belonging to different irreducible representations of the $SO(3)$ group. □

For B_2^M we have

$$\begin{aligned} B_2^M \geq & \sum_{\substack{k\neq l \\ l\neq n, n\neq k}} \frac{\langle (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} f_2^{(k,l)}\phi, f_2^{(l,n)} (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} f_2^{(n,k)}\phi\rangle}{8|D_{k,l}|^3 |D_{l,n}|^3 |D_{n,k}|^3} \\ & - C\frac{|\gamma_1|^2 + |\gamma_2|^2}{(D_0 d)^{10}}. \end{aligned} \tag{13.11}$$

To prove (13.11) we proceed similar to the proof of Proposition D.7 except the remark after (D.50), which says that for $M = 2$ we have $\langle \phi_2, f_2\phi_2\rangle = 0$. For $M \geq 3$ the argument of Lemma E.7 yields

$$\langle (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} f_2^{(m,m')}\phi, f_2^{(k,k')} (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} f_2^{(l,l')}\phi\rangle = 0 \tag{13.12}$$

only if at least one of the indices m, m', k, k', l, l' appears an even number of times. Consequently the terms with each of the indices m, m', k, k', l, l' coming twice contribute to the estimate of B_2^M.

The bound for B_3^M is not different from the one given in Proposition D.8 for $M = 2$.

To get the upper bound, analogous to the diatomic case let $\alpha^*_{\beta_M} \prec \alpha$ such that there is a function $\phi \in P^{\alpha^*_{\beta_M}} \tilde{\mathcal{W}}^\alpha_{\beta_M}$ with $\|\phi\| = 1$ that realises the maxima a_1^M and a_2^M. We set

$$\hat{\psi}_0 := \mathcal{U}^*_{\beta_M}\big(\phi - (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} \sum_{k<l} \frac{f_2^{(k,l)}}{|D_{k,l}|^3}\phi - (\tilde{H}_{\beta_M} - \mu_M^\alpha)^{-1} \sum_{k<l} \frac{f_3^{(k,l)}}{|D_{k,l}|^4}\phi\big) \tag{13.13}$$

and take as a trial function

$$\psi_0 := \frac{P^\alpha J_{\beta_M} \hat{\psi}_0}{\|P^\alpha J_{\beta_M} \hat{\psi}_0\|}, \tag{13.14}$$

and follow the same steps as in the proof of Theorem 9.2.

Appendix A. The HVZ theorem

In Appendix A and B we prove two fundamental facts regarding the spectra of a pseudo-relativistic Hamiltonian of an atom or positive ion, which are of crucial importance for Theorems 9.2 and 9.4.

In Appendix A we prove a HVZ-type theorem, which gives the location of the essential spectrum for an arbitrary type of permutational symmetry. In Appendix B we prove that Hamiltonians of pseudo-relativistic atoms and positive ions for any type of permutational symmetry have discrete eigenvalues at the bottom of the spectrum. Both results were announced earlier without proof by G. Zhislin in [49]. For the convenience of the reader, we give complete proofs in these appendices. In the nonrelativistic case both results are well-known. The first one, which is called HVZ theorem (see [13, p. 34]), was first proven without symmetry considerations in 1960 by G. Zhislin [48], and later generalized by Sigalov and Zhislin to the case of subspaces with fixed permutational symmetry [50]. The second one, which is known as Zhislin's theorem was proven in the same publications [48, 50].

For multiparticle Schrödinger operators with pseudo-relativistic kinetic energy the HVZ-type theorem was proven earlier in [22], where systems with finite particle masses and fixed total momentum were considered. The result needed for Theorems 9.2 and 9.4 is different from [22], because on one hand we have a particle with infinite mass, the nuclei, which makes the situation easier. On the other hand we need to include the permutational symmetry.

The method we use in the proof is a straightforward modification of the method by Sigalov and Zhislin, with necessary modifications related to the fact that the pseudo-relativistic kinetic energy operator is non-local, which also requires a different estimate of the localization error.

For any $k \in \mathbb{N}$ and $Ze^2 < \frac{2}{\pi}$ we set

$$H_k^Z := \sum_{i=1}^k T_i - \sum_{i=1}^k \frac{e^2 Z}{|x_i|} + \sum_{1 \le i < j \le k} \frac{e^2}{|x_i - x_j|} \tag{A.1}$$

acting on $L^2(\mathbb{R}^{3k})$, where T_i denotes the pseudo-relativistic kinetic energy operator for the i-th electron. Let α_k be an irreducible representation of the group of permutations of k electrons S_k. We set

$$\mu^{\alpha_k} := \inf \sigma(H_k^Z P^{\alpha_k}). \tag{A.2}$$

Denote by $\alpha'_{k-1} \prec \alpha_k$ an irreducible representation of S_{k-1} induced by α_k. We define

$$\mu_{k-1}^{\alpha_k} := \min_{\alpha'_{k-1} \prec \alpha_k} \inf \sigma(H_{k-1}^Z P^{\alpha'_{k-1}}). \tag{A.3}$$

Theorem A.1. *For subcritical nucleus charge $Ze^2 < \frac{2}{\pi}$ and for any irreducible representation α_k of S_k,*

$$\sigma_{ess}(H_k^Z P^{\alpha_k}) = [\mu_{k-1}^{\alpha_k}, +\infty).$$

Proof. The proof is split into two parts.

A.0.1. *"Easy part":* Let us first show that

$$\sigma_{ess}(H_k^Z P^{\alpha_k}) \supseteq [\mu_{k-1}^{\alpha_k}, +\infty). \tag{A.4}$$

To do so, for arbitrary $\lambda \geq \mu_{k-1}^{\alpha_k}$, we give the construction of a Weyl sequence $(\psi_m)_{m\in\mathbb{N}} \subset P^{\alpha_k} L^2(\mathbb{R}^{3k})$ with $\|\psi_m\| = 1$, $\psi_m \rightharpoonup 0$ and

$$\lim_{m\to\infty} \|(H_k^Z - \lambda)\psi_m\| = 0.$$

Let $\alpha_{k-1}^* \prec \alpha_k$ be an irreducible representation of S_{k-1} such that

$$\inf \sigma(H_{k-1}^Z P^{\alpha_{k-1}^*}) = \mu_{k-1}^{\alpha_k}.$$

Since $C_0^\infty(\mathbb{R}^{3(k-1)})$ is dense in the domain of $H_{k-1}^Z P^{\alpha_{k-1}^*}$, for any $\epsilon > 0$ there exists a function $\phi_\epsilon \in P^{\alpha_{k-1}^*} C_0^\infty(\mathbb{R}^{3(k-1)})$ with $\|\phi_\epsilon\| = 1$ such that

$$\|(H_{k-1}^Z - \mu_{k-1}^{\alpha_k})\phi_\epsilon\|^2 < \frac{\epsilon}{9}. \tag{A.5}$$

Let R_ϵ be such that

$$\operatorname{supp}(\phi_\epsilon) \subset \{x = (x_1, \cdots, x_{k-1}) \in \mathbb{R}^{3(k-1)} \big| |x_i| \leq R_\epsilon,\ i = 1, \cdots, k-1\}. \tag{A.6}$$

The spectrum of T_k is the positive real axis and $C_0^\infty(\mathbb{R}^3)$ is dense in the domain of T_k. Thus for any $\epsilon > 0$ there exists $f^{(\epsilon)} \in \mathbb{C}_0^\infty(\mathbb{R}^3)$ with $\|f^{(\epsilon)}\| = 1$ such that

$$\big\| \big[T_k - (\lambda - \mu_{k-1}^{\alpha_k})\big] f^{(\epsilon)} \big\|^2 \leq \frac{\epsilon}{9}.$$

Let us consider a decreasing sequence $\epsilon_m \to 0$ and the functions ϕ_{ϵ_m}, $f^{(\epsilon_m)}$ chosen accordingly as described above. For each of the ϵ_m we will pick a vector $A_m \in \mathbb{R}^3$ and define the shifted function

$$f_{A_m,\epsilon_m}(x_k) := f^{(\epsilon_m)}(x_k + A_m).$$

The sequence of shifts A_m is chosen such that $\operatorname{supp}(f_{A_m,\epsilon_m}) \cap B_{2R_{\epsilon_m}} = \emptyset$, and such that

$$\operatorname{supp}\big(f_{A_m,\epsilon_m}\big) \cap \Big(\bigcup_{l=1}^{m-1} \operatorname{supp}\big(f_{A_l,\epsilon_l}\big)\Big) = \emptyset.$$

Because the kinetic energy operator is translation invariant we get

$$\big\| \big[T_k - (\lambda - \mu_{k-1}^{\alpha_k})\big] f_{A_m,\epsilon_m} \big\|^2 \leq \frac{\epsilon_m}{9}. \tag{A.7}$$

We set

$$\varphi_m(x) := \phi_{\epsilon_m}(x_1, \cdots, x_{k-1}) f_{A_m,\epsilon_m}(x_k) \tag{A.8}$$

and let

$$\psi_m(x) := P^{\alpha_k} \varphi_m(x). \tag{A.9}$$

Similar to the proof in Section 12.2 we have

$$\|\psi_m\|^2 = \|P^{\alpha_k}\varphi_m\|^2 = \theta_{\alpha_{k-1}^*} \|\phi_{\epsilon_m} f_{A_m,\epsilon_m}\|^2, \tag{A.10}$$

where $\theta_{\alpha_{k-1}^*} > 0$ is a constant depending on α_{k-1}^* and α_k only (see Section 12.2). By choice of A_m, the functions ψ_m have disjoint support and thus $\psi_m \rightharpoonup 0$.

We will now estimate $\|(H_k^Z - \lambda)\psi_m\|$. The Hamiltonian H_k^Z commutes with the projection operator P^{α_k}, and since $\|P^{\alpha_k}\| \leq 1$ we get

$$\|(H_k^Z - \lambda)P^{\alpha_k}\varphi_m\|^2 = \|P^{\alpha_k}(H_k^Z - \lambda)\varphi_m\|^2 \leq \|(H_k^Z - \lambda)\varphi_m\|^2.$$

We split $(H_k^Z - \lambda)$ into three parts

$$(H_k^Z - \lambda) = (H_{k-1}^Z - \mu_{k-1}^{\alpha_k}) + \big(T_k - (\lambda - \mu_{k-1}^{\alpha_k})\big) + \big(\sum_{1\leq i<k} \frac{e^2}{|x_i - x_k|} - \frac{e^2 Z}{|x_k|}\big).$$

On the support of φ_m we have

$$\left| \sum_{1\le i<k} \frac{e^2}{|x_i - x_k|} - \frac{e^2 Z}{|x_k|} \right|^2 \le \frac{\epsilon_m}{9}. \tag{A.11}$$

Together with (A.5) and (A.7) this yields

$$\|(H_k^Z - \lambda)P^{\alpha_k}\varphi_m\|^2 \le \epsilon_m. \tag{A.12}$$

This shows that $\lambda \in \sigma_{ess}(H_k^Z P^{\alpha_k})$, and since $\lambda \in [\mu_{k-1}^{\alpha_k}, +\infty)$ was chosen arbitrarily this proves the inclusion.

A.0.2. *"Hard part":* We will show that

$$\sigma_{ess}(H_k^Z P^{\alpha_k}) \subseteq [\mu_{k-1}^{\alpha_k}, +\infty). \tag{A.13}$$

We prove this inclusion by induction in k. For $k = 1$, the hydrogen-like case, this is well-known. We fix an arbitrary $k \le Z$ and assume that for any $k' < k$ (A.13) is true. Take any $\lambda \in \sigma_{ess}(H_k^Z P^{\alpha_k})$ and a corresponding Weyl sequence $(\psi_l)_{l\in\mathbb{N}} \subset P^{\alpha_k} L^2(\mathbb{R}^{3k})$. Our aim is to show that

$$\lim_{l\to\infty} \langle \psi_l, H_k^Z \psi_l \rangle \ge \mu_{k-1}^{\alpha_k}.$$

By Weyl's criterion this implies (A.13).

Let $u_R \in C^\infty(\mathbb{R}^3; [0,1])$ such that

$$u_R(z) := \begin{cases} 1 & \text{if } |z| \le R \\ 0 & \text{if } |z| > 2R \end{cases} \tag{A.14}$$

and for any $\mathcal{C} \subseteq \{1, \cdots, k\}$ we define

$$F_{\mathcal{C}}(x) := \prod_{i\in\mathcal{C}} u_R(x_i) \prod_{j\notin\mathcal{C}} \sqrt{1 - u_R^2(x_j)}. \tag{A.15}$$

With this definition we have

$$\sum_{\mathcal{C}\subseteq\{1,\cdots,k\}} F_{\mathcal{C}}^2 \equiv 1. \tag{A.16}$$

Let $\mathcal{C}^* := \{1, \cdots, k\}$; observe that

$$\operatorname{supp}\left(F_{\mathcal{C}^*}\right) \subset \bigotimes_{i=1}^{k} B_{2R}^{(i)}.$$

We apply a weakened form of Theorem 11.3 to estimate the localization error and get

$$\begin{aligned} \langle \psi_l, H_k^Z \psi_l \rangle &= \langle F_{\mathcal{C}^*}\psi_l, H_k^Z F_{\mathcal{C}^*}\psi_l \rangle + \sum_{\mathcal{C}\neq\mathcal{C}^*} \langle F_{\mathcal{C}}\psi_l, H_k^Z F_{\mathcal{C}}\psi_l \rangle - \mathcal{LE} \\ &= \langle F_{\mathcal{C}^*}\psi_l, H_k^Z F_{\mathcal{C}^*}\psi_l \rangle + \sum_{\mathcal{C}\neq\mathcal{C}^*} \langle F_{\mathcal{C}}\psi_l, H_k^Z F_{\mathcal{C}}\psi_l \rangle + \mathcal{O}(R^{-2}). \end{aligned} \tag{A.17}$$

For the first term on the r.h.s. of (A.17) the definition of μ^{α_k}, see (A.2), implies

$$\begin{aligned} \langle F_{\mathcal{C}^*}\psi_l, H_k^Z F_{\mathcal{C}^*}\psi_l \rangle &\ge \mu^{\alpha_k} \|F_{\mathcal{C}^*}\psi_l\|^2 \\ &= \mu_{k-1}^{\alpha_k} \|F_{\mathcal{C}^*}\psi_l\|^2 + (\mu^{\alpha_k} - \mu_{k-1}^{\alpha_k}) \|F_{\mathcal{C}^*}\psi_l\|^2. \end{aligned} \tag{A.18}$$

Let

$$H_{\mathcal{C}}^Z = \sum_{i\in\mathcal{C}} T_i - \sum_{i\in\mathcal{C}} \frac{e^2 Z}{|x_i|} + \sum_{\substack{i,j\in\mathcal{C} \\ i<j}} \frac{e^2}{|x_i - x_j|}. \tag{A.19}$$

For each summand of the second term on the r.h.s. of (A.17) we write

$$\begin{aligned}\langle F_{\mathcal{C}}\psi_l, H_k^Z F_{\mathcal{C}}\psi_l\rangle &= \langle F_{\mathcal{C}}\psi_l, H_{\mathcal{C}}^Z F_{\mathcal{C}}\psi_l\rangle + \sum_{j\notin\mathcal{C}}\langle F_{\mathcal{C}}\psi_l, T_j F_{\mathcal{C}}\psi_l\rangle \\ &+ \sum_{j\notin\mathcal{C}}\big\langle F_{\mathcal{C}}\psi_l, \big(-\frac{e^2 Z}{|x_j|} + \sum_{i\neq j}\frac{e^2}{2|x_i - x_j|}\big)F_{\mathcal{C}}\psi_l\big\rangle.\end{aligned} \tag{A.20}$$

Each term in the second sum on the r.h.s. of (A.20) is non-negative. For the summands in the third term on the r.h.s. of (A.20), by construction of $F_{\mathcal{C}}$, there exists a constant $C > 0$ such that

$$\sum_{j\notin\mathcal{C}}\big\langle F_{\mathcal{C}}\psi_l, \big(-\frac{e^2 Z}{|x_j|} + \sum_{i\neq j}\frac{e^2}{2|x_i - x_j|}\big)F_{\mathcal{C}}\psi_l\big\rangle \geq -\frac{C}{R}\|F_{\mathcal{C}}\psi_l\|^2. \tag{A.21}$$

It is obvious that for any $\mathcal{C} \subseteq \{1, \cdots, k\}$ the function $F_{\mathcal{C}}$ is invariant under permutations in $S(\mathcal{C})$. This implies, that for $\psi \in P^{\alpha_k}L^2(\mathbb{R}^{3k})$ the function $F_{\mathcal{C}}\psi$ necessarily has a symmetry corresponding to an induced representation $\alpha'_{\mathcal{C}} \prec \alpha_k$ of $S(\mathcal{C})$. Thus for any $\mathcal{C} \neq \mathcal{C}^*$ we have

$$\langle F_{\mathcal{C}}\psi_l, H_{\mathcal{C}}^Z F_{\mathcal{C}}\psi_l\rangle \geq \min_{\alpha'_{\mathcal{C}}\prec\alpha_k} \inf \sigma(H_{\mathcal{C}}^Z P^{\alpha'_{\mathcal{C}}})\|F_{\mathcal{C}}\psi_l\|^2 \geq \mu_{k-1}^{\alpha_k}\|F_{\mathcal{C}}\psi_l\|^2 \tag{A.22}$$

by the induction assumption, since $H_{\mathcal{C}}^Z P^{\alpha'_{\mathcal{C}}}$ is unitarily equivalent to $H_{k'}^Z P^{\alpha_{k'}}$ for $k' = \sharp\mathcal{C}$ and some $\alpha_{k'} \prec \alpha_k$. Gathering (A.17), (A.18) and (A.20)-(A.22) we get that for some constant $C > 0$ independent of $l \in \mathbb{N}$ we have

$$\langle \psi_l, H_k^Z \psi_l\rangle \geq \mu_{k-1}^{\alpha_k}\underbrace{\sum_{\mathcal{C}}\|F_{\mathcal{C}}\psi_l\|^2}_{=1} + (\mu^{\alpha_k} - \mu_{k-1}^{\alpha_k})\|F_{\mathcal{C}^*}\psi_l\|^2 - \frac{C}{R}. \tag{A.23}$$

It remains to show that $\|F_{\mathcal{C}^*}\psi_l\|^2 \xrightarrow{l\to\infty} 0$. The operators $H_0 := \sum_{i=1}^k T_i$ and H_k^Z are semi-bounded from below, thus there exists a constant $c > 0$ such that $(H_0 + c)$ and $(H_k^Z + c)$ are positive operators. We write

$$F_{\mathcal{C}^*}\psi_l = F_{\mathcal{C}^*}(H_k^Z + c)^{-1}(H_k^Z + c)\psi_l. \tag{A.24}$$

Firstly we claim that the sequence $\big((H_k^Z + c)\psi_l\big)_{l\in\mathbb{N}}$ converges weakly to zero. Since $(\psi_l)_{l\in\mathbb{N}}$ is a Weyl sequence, $(H_k^Z - \lambda)\psi_l$ converges to zero in norm and

$$(H_k^Z + c)\psi_l = \underbrace{(H_k^Z - \lambda)\psi_l}_{\to 0} + \underbrace{(c + \lambda)\psi_l}_{\to 0}.$$

Our next goal is to show that the operator $F_{\mathcal{C}^*}(H_k^Z + c)^{-1}$ is compact. We write

$$F_{\mathcal{C}^*}(H_k^Z + c)^{-1} = F_{\mathcal{C}^*}(H_0 + c)^{-\frac{1}{2}}(H_0 + c)^{\frac{1}{2}}(H_k^Z + c)^{-\frac{1}{2}}(H_k^Z + c)^{-\frac{1}{2}}.$$

Since $(H_k^Z + c)^{-\frac{1}{2}}$ is the inverse of a strictly positive operator, it is bounded. To obtain a bound of $(H_0 + c)^{\frac{1}{2}}(H_k^Z + c)^{-\frac{1}{2}}$. Let V be the sum of Coulomb potentials in H_k^Z, such that

$$H_k^Z = H_0 + V.$$

Since V is relative H_0-bounded, there exist $1 > a > 0$ and $b > 0$ such that for all $\varphi \in \mathcal{D}(H_0) \cap \mathcal{D}(V)$ we have

$$|\langle\varphi, V\varphi\rangle| \leq a\langle\varphi, H_0\varphi\rangle + b\|\varphi\|^2.$$

By this inequality, for all $\varphi \in \mathcal{D}(H_0)$ we get

$$\begin{aligned}\langle\varphi, (H_0 + c)\varphi\rangle &= \langle\varphi, (H_0 + V + c)\varphi\rangle - \langle\varphi, V\varphi\rangle \\ &\leq \langle\varphi, (H_k^Z + c)\varphi\rangle + a\langle\varphi, H_0\varphi\rangle + b\|\varphi\|^2.\end{aligned}$$

Since $a < 1$, this is equivalent to

$$\langle \varphi, (H_0 + c)\varphi)\rangle \leq \frac{1}{1-a}\langle \varphi, (H_k^Z + c)\varphi\rangle + \frac{b-ac}{1-a}\|\varphi\|^2.$$

In particular, setting $\varphi = (H_k^Z + c)^{-\frac{1}{2}}\psi$ this yields

$$\begin{aligned}\|(H_0 + c)^{\frac{1}{2}}(H_k^Z + c)^{-\frac{1}{2}}\psi\|^2 &= \langle (H_k^Z + c)^{-\frac{1}{2}}\psi, (H_0 + c)(H_k^Z + c)^{-\frac{1}{2}}\psi\rangle \\ &\leq \frac{1}{1-a}\|\psi\|^2 + \frac{b-ac}{1-a}\|(H_k^Z + c)^{-\frac{1}{2}}\psi\|^2.\end{aligned}$$

Together with boundedness of $(H_k^Z + c)^{-\frac{1}{2}}$ this implies that $(H_0 + c)^{\frac{1}{2}}(H_k^Z + c)^{-\frac{1}{2}}$ is bounded. Finally note that the operator $F_{\mathcal{C}^*}(H_0 + c)^{-\frac{1}{2}}$ is compact, being a norm limit of Hilbert-Schmidt operators

$$B_n = F_{\mathcal{C}^*}(H_0 + c)^{-1}\chi(H_0 < n). \tag{A.25}$$

Thus

$$\|F_{\mathcal{C}^*}\psi_l\|^2 = \|F_{\mathcal{C}^*}(H_0 + c)^{-\frac{1}{2}}(H_0 + c)^{\frac{1}{2}}(H_k^Z + c)^{-\frac{1}{2}}(H_k^Z + c)^{\frac{1}{2}}\psi_l\|^2 \xrightarrow{l\to\infty} 0. \tag{A.26}$$

Recall from inequality (A.23) that

$$\langle \psi_l, H_k^Z\psi_l\rangle \geq \mu_{k-1}^{\alpha_k} + (\mu^{\alpha_k} - \mu_{k-1}^{\alpha_k})\|F_{\mathcal{C}^*}\psi_l\|^2 - \frac{C}{R}.$$

Picking R and l large yields $\lambda \geq \mu_{k-1}^{\alpha_k}$, where λ was an arbitrary value in the essential spectrum of $H_k^Z P^{\alpha_k}$. □

Appendix B. Existence of a ground state for atoms and positive ions

Let H_k^Z, S_k, and α_k be the same as in Appendix A and let $k \leq Z$.

Theorem B.1. *For any irreducible representation α_k of the group S_k, the operator $H_k^Z P^{\alpha_k}$ has a discrete eigenvalue at the bottom of its spectrum.*

Proof of Theorem B.1. We prove the theorem by induction in $k = 1, \cdots, Z$. For $k = 1$ we have

$$H_1^Z = \sqrt{p^2+1} - 1 - \frac{Ze^2}{|x|} \leq \frac{p^2}{2} - \frac{Ze^2}{|x|}.$$

The operator $\frac{p^2}{2} - \frac{Ze^2}{|x|}$ has an infinite number of negative eigenvalues, which yields the existence of a negative eigenvalue for H_1^Z. Note that for one electron we do not have restrictions regarding its symmetry.

For fixed but arbitrary $k \leq Z$, let us assume that for each irreducible representation α_{k-1} of the permutation group S_{k-1}, the operator $H_{k-1}^Z P^{\alpha_{k-1}}$ has a ground state.

We will construct a trial state $\psi_0 \in P^{\alpha_k} H^{1/2}(\mathbb{R}^{3k})$ for arbitrary irreducible representation α_k of S_k such that

$$\|\psi_0\|^{-2}\langle \psi_0, H_k^Z\psi_0\rangle < \inf \sigma_{ess}(H_k^Z P^{\alpha_k}).$$

Let $\alpha_{k-1}^* \prec \alpha_k$ be an irreducible representation of S_{k-1} such that

$$\inf \sigma(H_{k-1}^Z P^{\alpha_{k-1}^*}) = \min_{\alpha_{k-1}' \prec \alpha_k} \inf \sigma(H_{k-1}^Z P^{\alpha_{k-1}'}) =: \mu_{k-1}^{\alpha_k}. \tag{B.1}$$

By the induction assumption, there exists a state $\phi \in P^{\alpha_{k-1}^*} H^{1/2}(\mathbb{R}^{3(k-1)})$ with

$$\langle \phi, H_{k-1}^Z\phi\rangle = \mu_{k-1}^{\alpha_k}\|\phi\|^2. \tag{B.2}$$

Let $f \in C_0^\infty(\mathbb{R}^3)$ with $\|f\|_{L^2} = 1$ and $\mathrm{supp}(f) \subset \{x \in \mathbb{R}^3 | 1 \leq |x| \leq 2\}$, and let

$$f_R(z) := R^{-\frac{3}{2}} f(zR^{-1}), \tag{B.3}$$

so that $\|f_R\| = 1$. For $u \in C^\infty(\mathbb{R}^3; [0,1])$ with

$$u(z) := \begin{cases} 1 & \text{if } |z| \le \frac{1}{2} \\ 0 & \text{if } |z| \ge 1 \end{cases} \tag{B.4}$$

we define the cutoff function

$$\zeta_{R,Z}(x_1, \cdots, x_{k-1}) := \prod_{i=1}^{k-1} u\Big(x_i \cdot \frac{Z+1}{R}\Big). \tag{B.5}$$

This cutoff function localizes each particle $i = 1, \cdots, k-1$ in a ball of radius $\frac{R}{Z+1}$ and is invariant under permutations in S_{k-1}. We define

$$\hat{\psi}_0(x) := (\zeta_{R,Z}\phi)(x_1, \cdots, x_{k-1}) f_R(x_k) \tag{B.6}$$

and the trial state

$$\psi_0 := \frac{P^{\alpha_k}\hat{\psi}_0}{\|P^{\alpha_k}\hat{\psi}_0\|}. \tag{B.7}$$

Following the same argument as in Section 12.2, we have

$$\frac{\langle \hat{\psi}_0, H_k^Z P^{\alpha_k}\hat{\psi}_0\rangle}{\|P^{\alpha_k}\hat{\psi}_0\|^2} = \frac{\langle \hat{\psi}_0, H_k^Z \hat{\psi}_0\rangle}{\|\hat{\psi}_0\|^2}. \tag{B.8}$$

We split the Hamiltonian H_k^Z into three parts

$$H_k^Z = H_{k-1}^Z + T_k + \Big(\sum_{1\le i<k} \frac{e^2}{|x_i - x_k|} - \frac{e^2 Z}{|x_k|}\Big). \tag{B.9}$$

Using the exponential decay of the eigenfunction ϕ, similar to (12.72), we get

$$\langle \zeta_{R,Z}\phi, H_{k-1}^Z \zeta_{R,Z}\phi\rangle = \mu_{k-1}^{\alpha_k}\|\phi\|^2 + \mathcal{O}(e^{-cR}) \tag{B.10}$$

for some constant $c > 0$. Note that for $x_k \in \mathrm{supp}(f_R)$ we have $|x_k| = (1+\theta)R$ for some $\theta \in [0,1]$ and by choice of $\zeta_{R,Z}$, for $x \in \mathrm{supp}(\hat{\psi}_0)$ we get

$$\begin{aligned} \sum_{1\le i<k} \frac{e^2}{|x_i - x_k|} &\le \sum_{1\le i<k} \frac{e^2}{|x_k| - |x_i|} \\ &\le \frac{e^2(k-1)(Z+1)}{(Z+Z\theta+\theta)R} \end{aligned} \tag{B.11}$$

and

$$-\frac{e^2 Z}{|x_k|} = -\frac{e^2 Z}{(1+\theta)R}. \tag{B.12}$$

Using (B.11) and (B.12), and $k \le Z$, we arrive at

$$\begin{aligned} \sum_{1\le i<k} \frac{e^2}{|x_i - x_k|} - \frac{e^2 Z}{|x_k|} &\le -\frac{e^2(\theta+1+Z\theta)}{R(1+\theta)(Z+Z\theta+\theta)} \\ &\le -\frac{e^2}{R(Z+Z\theta+\theta)} - \frac{e^2 Z\theta}{R(1+\theta)(Z+Z\theta+\theta)}. \end{aligned} \tag{B.13}$$

The first term on the r.h.s. is increasing in θ and the second term is non-positive, which yields the bound

$$\langle \hat{\psi}_0, \Big(\sum_{1\le i<k} \frac{e^2}{|x_i - x_k|} - \frac{e^2 Z}{|x_k|}\Big)\hat{\psi}_0\rangle \le -\frac{e^2}{(2Z+1)R}\|\hat{\psi}_0\|^2. \tag{B.14}$$

Furthermore, for the particle k we have

$$\begin{aligned}\langle\psi_0, T_k\psi_0\rangle &= \|\zeta_{R,Z}\phi\|^2\langle f_R, T_k f_R\rangle \\ &\le \|\zeta_{R,Z}\phi\|^2\langle f_R, \frac{p_k^2}{2} f_R\rangle \\ &\le \frac{C}{R^2}\|\hat{\psi}_0\|^2.\end{aligned} \tag{B.15}$$

Collecting (B.8), (B.10), (B.14) and (B.15), we get

$$\frac{\langle\psi_0, H_k^Z P^{\alpha_k}\psi_0\rangle}{\|P^{\alpha_k}\psi_0\|^2} \le \mu_{k-1}^{\alpha_k} + \frac{C}{R^2} - \frac{e^2}{(2Z+1)R} < \mu_{k-1}^{\alpha_k} \tag{B.16}$$

for sufficiently large R. By Theorem A.1 we have

$$\mu_{k-1}^{\alpha_k} = \inf \sigma_{ess}(H_k^Z P^{\alpha_k}). \tag{B.17}$$

So (B.16) shows that the discrete spectrum of $H_k^Z P^{\alpha_k}$ below $\mu_{k-1}^{\alpha_k}$ is not empty, in particular a ground state of $H_k^Z P^{\alpha_k}$ exists. □

Appendix C. Proof of Lemma 10.5

To prove Lemma 10.5 we start with some auxiliary results.

Lemma C.1. *For a bounded function $\xi \in C^1(\mathbb{R}^3)$, $T = \sqrt{p^2+1} - 1$ and any function $\varphi \in H^{1/2}(\mathbb{R}^3)$ we have*

$$\big|\operatorname{Re}\langle\xi\varphi, [\xi, T]\varphi\rangle\big| \le \frac{1}{4\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)}{|x-y|^2}(\xi(x)-\xi(y))^2|\varphi(x)|\ |\varphi(y)|dydx \tag{C.1}$$

where K_2 is the modified Bessel function of order two.

Proof. Since T is self-adjoint we can write

$$\operatorname{Re}\langle\xi\varphi, [\xi, T]\varphi\rangle = \big(\frac{1}{2}\operatorname{Re}\langle\xi^2\varphi, T\varphi\rangle + \frac{1}{2}\operatorname{Re}\langle\varphi, T\xi^2\varphi\rangle - \langle\xi\varphi, T\xi\varphi\rangle\big). \tag{C.2}$$

For two functions $f, g \in H^{1/2}(\mathbb{R}^3)$ one has

$$\operatorname{Re}\langle f, Tg\rangle = \frac{1}{4\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)}{|x-y|^2}\operatorname{Re}\big((f(x)-f(y))\overline{(g(x)-g(y))}\big)dxdy. \tag{C.3}$$

To prove (C.3) notice that according to [24, Theorem 7.12] we have

$$\langle\varphi, T\varphi\rangle = \frac{1}{4\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)|\varphi(x)-\varphi(y)|^2}{|x-y|^2}dxdy \tag{C.4}$$

for $\varphi \in H^{1/2}(\mathbb{R}^3)$. If $f, g \in H^{1/2}(\mathbb{R}^3)$ then

$$\operatorname{Re}\langle f, Tg\rangle = \frac{1}{4}\Big(\langle f+g, T(f+g)\rangle - \langle f-g, T(f-g)\rangle\Big). \tag{C.5}$$

Using (C.4) in the r.h.s. of (C.5) we get

$$\begin{aligned}\operatorname{Re}\langle f, Tg\rangle = &\frac{1}{16\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)\big|(f+g)(x)-(f+g)(y)\big|^2}{|x-y|^2}dxdy \\ &- \frac{1}{16\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)\big|(f-g)(x)-(f-g)(y)\big|^2}{|x-y|^2}dxdy,\end{aligned} \tag{C.6}$$

and taking

$$\begin{aligned}&\frac{1}{4}\Big(\big|(f+g)(x)-(f+g)(y)\big|^2 - \big|(f-g)(x)-(f-g)(y)\big|^2\Big) \\ &= \operatorname{Re}\Big((f(x)-f(y))\overline{(g(x)-g(y))}\Big)\end{aligned}$$

into account, we arrive at (C.3).
Applying (C.3) for each of the three summands in (C.2) we have

$$\begin{aligned}
&\mathrm{Re}\langle \xi\varphi, [\xi, T]\varphi\rangle \\
&= \frac{1}{8\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)}{|x-y|^2}\,\mathrm{Re}\Big[\big(\xi^2\varphi(x)-\xi^2\varphi(y)\big)\overline{\big(\varphi(x)-\varphi(y)\big)}\Big]dydx \\
&+ \frac{1}{8\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)}{|x-y|^2}\,\mathrm{Re}\Big[\big(\varphi(x)-\varphi(y)\big)\overline{\big(\xi^2\varphi(x)-\xi^2\varphi(y)\big)}\Big]dydx \\
&- \frac{1}{4\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)}{|x-y|^2}\Big[\big(\xi\varphi(x)-\xi\varphi(y)\big)\overline{\big(\xi\varphi(x)-\xi\varphi(y)\big)}\Big]dydx.
\end{aligned} \tag{C.7}$$

By multiplying out the expressions in the square brackets in (C.7) one sees

$$\mathrm{Re}\langle \xi\varphi, [\xi, T]\varphi\rangle = -\frac{1}{4\pi^2}\int_{\mathbb{R}^6} \frac{K_2(|x-y|)}{|x-y|^2}\Big[\big(\xi(x)-\xi(y)\big)^2\,\mathrm{Re}\,\big(\overline{\varphi(x)}\varphi(y)\big)\Big]dydx \tag{C.8}$$

which yields (C.1), because the modified Bessel function $K_2(|\cdot|)$ is positive. □

For convenience of the reader we recall

Lemma 10.5. *Let the functions $F_{\nu,\epsilon}$, and ξ_ϵ be as in* (10.10) *and* (10.12)*. For fixed $R>0$ and arbitrary $\hat\delta>0$ there exists $\nu_0>0$ and $C_{\nu_0,R}<\infty$ such that for all $\nu\in(0,\nu_0)$, all $\epsilon\in(0,1)$, and all $\varphi\in H^{1/2}(\mathbb{R}(\mathcal{C}))$*

$$|\,\mathrm{Re}\langle \xi_\epsilon\varphi, [\xi_\epsilon, \tilde H^Z_{\mathcal{C}}]\varphi\rangle| \le \hat\delta\|\sum_{i=1}^k e^{F_{\nu,\epsilon}(|x_i|)}\varphi\|^2 + C_{\nu_0,R}\|\varphi\|^2. \tag{10.13}$$

Proof of Lemma 10.5. Since the Coulomb potential is a multiplicative operator, the commutator of ξ_ϵ with V is zero. Thus

$$\mathrm{Re}\langle \xi_\epsilon\varphi, [\xi_\epsilon, H^Z_{\mathcal{C}}]\varphi\rangle = \sum_{i=1}^k \mathrm{Re}\langle \xi_\epsilon\varphi, [\xi_\epsilon, T_i]\varphi\rangle. \tag{C.9}$$

Given $x=(x_1,\cdots,x_k)$, we set $y^i=(x_1,\cdots,x_{i-1},y_i,x_{i+1},\cdots,x_k)\in\mathbb{R}^{3k}$. By Lemma C.1 we get

$$\begin{aligned}
&|\,\mathrm{Re}\langle \xi_\epsilon\varphi, [\xi_\epsilon, H^Z_{\mathcal{C}}]\varphi\rangle| \\
&\le \sum_{i=1}^k \frac{1}{4\pi^2}\int_{R^{3k+3}} \frac{K_2(|x_i-y_i|)}{|x_i-y_i|^2}(\xi_\epsilon(x)-\xi_\epsilon(y^i))^2|\varphi(x)|\;|\varphi(y^i)|dy_idx \\
&=: \sum_{i=1}^k \mathcal{I}(i).
\end{aligned} \tag{C.10}$$

Let us estimate the term with $i=1$ in the r.h.s. of (C.10). The estimate for all other terms is similar. We begin by splitting the integrand into two parts, using the bound

$$\begin{aligned}
(\xi_\epsilon(x)-\xi_\epsilon(y^1))^2 &= \Big(\chi_R(x)\sum_{i=1}^k e^{F(|x_i|)} - \chi_R(y^1)\Big(e^{F(|y_1|)}+\sum_{i=2}^k e^{F(|x_i|)}\Big)\Big)^2 \\
&\le 2\Big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\Big)^2 + 2\Big(\chi_R(x)-\chi_R(y^1)\Big)^2\big(\sum_{i=2}^k e^{F(|x_i|)}\big)^2.
\end{aligned} \tag{C.11}$$

We define the integrals

$$\mathcal{I}_1(1) := \int_{R^{3k+3}} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dy_1dx \tag{C.12}$$

and

$$\mathcal{I}_2(1) := \int_{\mathbb{R}^{3k+3}} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x) - \chi_R(y^1)\big)^2\big(\sum_{i=2}^k e^{F(|x_i|)}\big)^2|\varphi(x)||\varphi(y^1)|dy_1dx. \tag{C.13}$$

In Lemma C.2 below we will prove the following estimate for $\mathcal{I}_1(1)$: For any $\hat{\delta} > 0$ there exists a constant $C_1(R) < \infty$ such that

$$\mathcal{I}_1(1) \le \frac{\hat{\delta}}{k}\|e^{F(|x_1|)}\varphi(x)\|^2 + C_1(R)\|\varphi\|^2. \tag{C.14}$$

In Lemma C.6 below we will prove that there exists a constant $C_2(R) < \infty$ such that

$$\mathcal{I}_2(1) \le C_2(R)\|\varphi\|^2. \tag{C.15}$$

So

$$\mathcal{I}(1) \le \mathcal{I}_1(1) + \mathcal{I}_2(1) \le \frac{\hat{\delta}}{k}\|e^{F(|x_1|)}\varphi(x)\|^2 + (C_1(R) + C_2(R))\|\varphi\|^2. \tag{C.16}$$

Together with the inequality

$$\|e^{F(|x_i|)}\varphi(x)\|^2 \le \|\sum_{l=1}^k e^{F(|x_l|)}\varphi\|^2 \quad \forall i \in \{1, \cdots, k\} \tag{C.17}$$

we get that for any $\hat{\delta}$ there exists a constant $C_R > 0$

$$\mathcal{I}(1) \le \frac{\hat{\delta}}{k}\|\sum_{i=1}^k e^{F(|x_i|)}\varphi\|^2 + C_R\|\varphi\|^2. \tag{C.18}$$

Using this, and similar bounds for $\mathcal{I}(i)$, for $i = 2, \cdots, k$ in (C.10), summation over $i = 1, \cdots, k$ proves Lemma 10.5. □

Lemma C.2. *For $\mathcal{I}_1(1)$ defined in* (C.12) *and any $\hat{\delta} > 0$ there exists a constant $C_1(R) < \infty$ such that for all $\nu < \nu_0$*

$$\mathcal{I}_1(1) \le \frac{\hat{\delta}}{k}\|e^{F(|x_1|)}\varphi(x)\|^2 + C_1(R)\|\varphi\|^2. \tag{C.19}$$

For the proof of Lemma C.2 we will need some preparation. We will split the area of integration into three regions. For this purpose, define

$$\mathcal{I}_1^\Omega(1) := \frac{1}{4\pi^2}\int_\Omega \frac{K_2(|x_1-y_1|)}{|x_1-y_1|^2}\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dy_1dx_1 \tag{C.20}$$

for a Borel set $\Omega \subset \mathbb{R}^6$. Note that for $\Omega = \mathbb{R}^6$ the integral $\mathcal{I}_1^\Omega(1)$ is symmetric under the permutation of x_1 with y_1. Consequently for $\Omega = \mathbb{R}^6$ it suffices to estimate this integral over the region $\{(x_1, y_1) \in \mathbb{R}^6 \mid |x_1| \le |y_1|\}$. We will cover this region by the three regions A_1, A_2, A_3 defined as

$$\begin{aligned} A_1 &:= \{(x_1, y_1) \in \mathbb{R}^6 |\ |x_1| \le |y_1|,\ |x_1 - y_1| > \tilde{R}\} \\ A_2 &:= \{(x_1, y_1) \in \mathbb{R}^6 |\ |x_1| \le |y_1| \le 2R + \tilde{R},\ |x_1 - y_1| \le \tilde{R}\} \\ A_3 &:= \{(x_1, y_1) \in \mathbb{R}^6 |\ 2R \le |x_1| \le |y_1|,\ |x_1 - y_1| \le \tilde{R}\}, \end{aligned} \tag{C.21}$$

where $\tilde{R}$ is a fixed positive parameter. The corresponding integrals are estimated separately in Lemmata C.3-C.5.

Lemma C.3 (bound in region A_1)**.** *We have*

$$\mathcal{I}_1^{A_1}(1) \leq \frac{\|K_2(|\cdot|)e^{\nu|\cdot|}\|^2_{L^1(\mathbb{R}^3)}}{4\pi^2\tilde{R}^2}\int_{\mathbb{R}^3}|\varphi(x)e^{F_{\nu,\epsilon}(|x_1|)}|^2dx_1. \tag{C.22}$$

Proof. The function $F(\cdot) = F_{\nu,\epsilon}(\cdot)$ is monotone increasing and since $|x_1| \leq |y_1|$ and $\chi_R \leq 1$, one has

$$\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2 \leq e^{2F(|y_1|)}. \tag{C.23}$$

Furthermore, using monotonicity of F and the triangle inequality

$$F(|x+y|) \leq F(|x|) + F(|y|)$$

we obtain

$$e^{F(|y_1|-|x_1-y_1|)} \leq e^{F(|x_1|)}.$$

Moreover, for $0 < r_1 < r_2$ we have $F(r_2) - F(r_1) \leq \nu(r_2 - r_1)$. Therefore

$$F(|y_1|) \leq F\big(|y_1| - |x_1 - y_1|\big) + \nu|x_1 - y_1|.$$

Gathering these bounds yields

$$e^{2F(|y_1|)} \leq e^{F(|y_1|)}e^{F(|x_1|)}e^{\nu|x_1-y_1|}. \tag{C.24}$$

Substituting (C.23) and (C.24) into $\mathcal{I}_1^{A_1}(1)$ we get

$$\begin{aligned}\mathcal{I}_1^{A_1}(1) &= \int_{A_1}\frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dx_1dy_1\\ &\leq \int_{A_1}\frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}e^{F(|x_1|)}e^{F(|y_1|)}e^{\nu|x_1-y_1|}|\varphi(x)||\varphi(y^1)|dx_1dy_1.\end{aligned} \tag{C.25}$$

Since for $(x_1, y_1) \in A_1$ one has $|x_1 - y_1| > \tilde{R}$, the r.h.s of (C.25) can be estimated as

$$\begin{aligned}&\int_{A_1}\frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}e^{F(|x_1|)}e^{F(|y_1|)}e^{\nu|x_1-y_1|}|\varphi(x)||\varphi(y^1)|dx_1dy_1\\ &\leq \frac{1}{4\pi^2\tilde{R}^2}\int_{\mathbb{R}^6}K_2(|x_1-y_1|)e^{\nu|x_1-y_1|}|\varphi(x)e^{F(|x_1|)}|\,|\varphi(y^1)e^{F(|y_1|)}|dx_1dy_1.\end{aligned} \tag{C.26}$$

Note that for ν sufficiently small $K_2(|\cdot|)e^{\nu|\cdot|} \in L^1(\mathbb{R}^3)$ which, together with Young's inequality for the y_1 integration implies

$$\mathcal{I}_1^{A_1}(1) \leq \frac{\|K_2(|\cdot|)e^{\nu|\cdot|}\|^2_{L^1(\mathbb{R}^3)}}{4\pi^2\tilde{R}^2}\int_{\mathbb{R}^3}|\varphi(x)e^{F(|x_1|)}|^2dx_1.$$

□

Lemma C.4 (bound in region A_2)**.** *Let $R, \tilde{R}, \nu > 0$ be fixed. For all $\epsilon \in (0,1)$ we have*

$$\mathcal{I}_1^{A_2}(1) \leq (\|\chi_R\|_{C^1} + \nu)^2e^{2\nu(2R+\tilde{R})}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}\int_{\mathbb{R}^3}|\varphi(x)|^2dx_1 \tag{C.27}$$

Proof. Using the mean value theorem we have

$$\begin{aligned}&\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2\\ &\leq |x_1-y_1|^2\Big(\Big(\|\chi_R\|_{C^1} + \chi_R((x^*, x_2, \cdots, x_k)F'(|x^*|)\Big)e^{F(|x^*|)}\Big)^2\end{aligned} \tag{C.28}$$

for some x^* with $|x_1| \leq |x^*| \leq |y_1|$. Since $\|F'\|_\infty \leq \nu$, we obtain

$$\begin{aligned}&|x_1-y_1|^2\Big(\Big(\|\chi_R\|_{C^1} + \chi_R((x^*, x_2, \cdots, x_k)F'(|x^*|)\Big)e^{F(|x^*|)}\Big)^2\\ &\leq |x_1-y_1|^2(\|\chi_R\|_{C^1} + \nu)^2e^{2F(|y_1|)}.\end{aligned} \tag{C.29}$$

Using (C.28) and (C.29) we arrive at

$$\begin{aligned}\mathcal{I}_1^{A_1}(1) &= \int_{A_2} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dx_1dy_1 \\ &\leq \frac{(\|\chi_R\|_{C^1}+\nu)^2}{4\pi^2}\int_{A_2} K_2(|x_1-y_1|)e^{2F(|y_1|)}|\varphi(x)||\varphi(y^1)|dx_1dy_1.\end{aligned} \tag{C.30}$$

Due to monotonicity of F and since $|y_1| \leq 2R+\tilde{R}$, the r.h.s. of (C.30) can be bounded as

$$\begin{aligned}&\frac{(\|\chi_R\|_{C^1}+\nu)^2}{4\pi^2}e^{2F(2R+\tilde{R})}\int_{A_2} K_2(|x_1-y_1|)|\varphi(x)||\varphi(y^1)|dx_1dy_1 \\ &\leq \frac{(\|\chi_R\|_{C^1}+\nu)^2}{4\pi^2}e^{2\nu(2R+\tilde{R})}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}\int_{\mathbb{R}^3}|\varphi(x)|^2dx_1.\end{aligned}$$

In the last step, we applied Young's inequality and the fact that $F(r) \leq \nu r$ for all $r > 0$. This concludes the proof of the lemma. □

Lemma C.5 (bound in region A_3)**.** *Let $R, \tilde{R}, \nu > 0$ be fixed. For all $\epsilon \in (0,1)$ we have*

$$\mathcal{I}_1^{A_3}(1) \leq \frac{\nu^2 e^{\nu\tilde{R}}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}}{4\pi^2}\int_{\mathbb{R}^3}|\varphi e^{F_{\nu,\epsilon}(|x_1|)}|^2dx_1. \tag{C.31}$$

Proof. Since $\chi_R(x) = \chi_R(y^1)$ in the region A_3 and $\chi_R \leq 1$, we get

$$\begin{aligned}\mathcal{I}_1^{A_3}(1) &= \int_{A_3} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x)e^{F(|x_1|)} - \chi_R(y^1)e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dx_1dy_1 \\ &\leq \int_{A_3} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(e^{F(|x_1|)} - e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dx_1dy_1.\end{aligned}$$

By mean value theorem and monotonicity of F, we obtain

$$\begin{aligned}&\int_{A_3} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(e^{F(|x_1|)} - e^{F(|y_1|)}\big)^2|\varphi(x)||\varphi(y^1)|dx_1dy_1 \\ &\leq \frac{\nu^2}{4\pi^2}\int_{A_3} K_2(|x_1-y_1|)e^{2F(|y_1|)}|\varphi(x)|\varphi(y^1)|dx_1dy_1.\end{aligned} \tag{C.32}$$

Using (C.24) this implies

$$\begin{aligned}&\frac{\nu^2}{4\pi^2}\int_{A_3} K_2(|x_1-y_1|)e^{2F(|y_1|)}|\varphi(x)|\varphi(y^1)|dx_1dy_1 \\ &\leq \frac{\nu^2}{4\pi^2}\int_{A_3} K_2(|x_1-y_1|)|e^{F(|x_1|)}\varphi(x)|\,|e^{F(|y_1|)}\varphi(y^1)|\,e^{\nu|x_1-y_1|}dx_1dy_1.\end{aligned} \tag{C.33}$$

Applying Young's inequality together with the bound $e^{\nu|x_1-y_1|} \leq e^{\nu\tilde{R}}$ yields

$$\begin{aligned}&\frac{\nu^2}{4\pi^2}\int_{A_3} K_2(|x_1-y_1|)|e^{F(|x_1|)}\varphi(x)|\,|e^{F(|y_1|)}\varphi(y^1)|\,e^{\nu|x_1-y_1|}dx_1dy_1 \\ &\leq \frac{\nu^2 e^{\nu\tilde{R}}}{4\pi^2}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}\int_{\mathbb{R}^3}|\varphi(x)e^{F(|x_1|)}|^2dx_1.\end{aligned} \tag{C.34}$$

□

Proof of Lemma C.2. By the definition of $\mathcal{I}_1(1)$ we have

$$\mathcal{I}_1(1) \leq 2\int_{\mathbb{R}^{3K-3}}\Big(\mathcal{I}_1^{A_1}(1) + \mathcal{I}_1^{A_2}(1) + \mathcal{I}_1^{A_3}(1)\Big)dx_2\cdots dx_k.$$

Summing (C.22), (C.27), and (C.31) and integrating against variables $x_2, \cdots, x_k$ we arrive at

$$\begin{aligned}\mathcal{I}_1(1) \leq & \Big(\frac{\|K_2(|\cdot|)e^{\nu|\cdot|}\|^2_{L^1(\mathbb{R}^3)}}{2\pi^2\tilde{R}^2} + \frac{\nu^2 e^{\nu\tilde{R}}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}}{2\pi^2}\Big)\|\varphi(x)e^{F(|x_1|)}\|^2 \\ & + \frac{(\|\chi_R\|_{C^1}+\nu)^2 e^{2\nu(2R+\tilde{R})}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}}{2\pi^2}\|\varphi\|^2.\end{aligned} \tag{C.35}$$

Given $\hat{\delta} > 0$ we choose $\tilde{R}$ sufficiently large and ν_0 sufficiently small such that for all $\nu \in (0, \nu_0)$ one has

$$\Big(\frac{\|K_2(|\cdot|)e^{\nu|\cdot|}\|^2_{L^1(\mathbb{R}^3)}}{2\pi^2\tilde{R}^2} + \frac{\nu^2 e^{\nu\tilde{R}}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}}{2\pi^2}\Big) < \frac{\hat{\delta}}{k}. \tag{C.36}$$

For this fixed $\tilde{R}$, the choice

$$C_1(R) := \frac{(\|\chi_R\|_{C^1}+\nu_0)^2 e^{2\nu_0(2R+\tilde{R})}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}}{2\pi^2} \tag{C.37}$$

completes the proof of Lemma C.2. □

As the next step we will estimate the integral $\mathcal{I}_2(1)$ defined in (C.13).

Lemma C.6 (bound of the integral $\mathcal{I}_2(1)$)**.** *There exists a constant $C_2(R,\nu) > 0$ such that*

$$\mathcal{I}_2(1) \leq C_2(R,\nu)\|\varphi\|^2.$$

Proof. Recall

$$\mathcal{I}_2(1) = \int_{\mathbb{R}^{3k+3}} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x)-\chi_R(y^1)\big)^2\big(\sum_{l=2}^k e^{F(|x_l|)}\big)^2|\varphi(x)||\varphi(y^1)|dy_1dx. \tag{C.38}$$

By definition of χ_R, we have

$$\chi_R(x)-\chi_R(y^1) = \mathbf{1}_{|x_2|<2R}\cdots\mathbf{1}_{|x_k|<2R}(\chi_R(x)-\chi_R(y^1)).$$

Using $F(|x_l|) \leq 2\nu R$ for $l = 2, \cdots, k$ and the mean value theorem, we get

$$\begin{aligned}(\chi_R(x)-\chi_R(y^1))^2\big(e^{F(|x_2|)}+\cdots+e^{F(|x_k|)}\big)^2 &\leq (\chi_R(x)-\chi_R(y^1))^2k^2e^{4\nu R} \\ &\leq \|\chi_R\|^2_{C^1}k^2e^{4\nu R}|x_1-y_1|^2.\end{aligned}$$

Together with Young's inequality this yields

$$\begin{aligned}&\int_{\mathbb{R}^{3k+3}} \frac{K_2(|x_1-y_1|)}{4\pi^2|x_1-y_1|^2}\big(\chi_R(x)-\chi_R(y^1)\big)^2\big(\sum_{i=2}^k e^{F(|x_i|)}\big)^2|\varphi(x)||\varphi(y^1)|dy_1dx \\ &\leq \frac{\|\chi_R\|^2_{C^1}(k-1)^2e^{4\nu R}}{4\pi^2}\int_{\mathbb{R}^{3k+3}} K_2(|x_1-y_1|)|\varphi(x)||\varphi(y^1)|dy_1dx_1\cdots dx_k \\ &\leq \frac{\|\chi_R\|^2_{C^1}(k-1)^2e^{4\nu R}}{4\pi^2}\|K_2(|\cdot|)\|_{L^1(\mathbb{R}^3)}\|\varphi\|^2.\end{aligned}$$

□

Appendix D. Intercluster interaction in diatomic molecules

In this part we estimate the term $\langle J_\beta\psi, I_\beta J_\beta\psi\rangle$ which is an important part in the proof of Theorem 9.2. For these estimates we will use orthogonality relations, which will be proven in Appendix E.

Denote by $P_n(z), n \in \mathbb{N}, z \in \mathbb{R}$ the n-th degree Legendre polynomial, these polynomials are generated by $(1-2zt+t^2)^{-\frac{1}{2}}$ (see [1, 22.9.12]), more explicitly for $-1 < z < 1$ and $|t| < 1$ we have

$$\frac{1}{\sqrt{1-2zt+t^2}} = \sum_{n=0}^{\infty} P_n(z)t^n. \tag{D.1}$$

Consequently, for $D, h \in \mathbb{R}^3$ with $h < D$ we get

$$\frac{1}{|D-h|} = \sum_{n=0}^{\infty} P_n\left(\frac{h}{|h|}\cdot\frac{D}{|D|}\right)\frac{|h|^n}{|D|^{n+1}}. \tag{D.2}$$

In particular for $n = 2, 3, 4$ we have

$$P_2(z) = \frac{1}{2}(3z^2-1), \quad P_3(z) = \frac{1}{2}(5z^3-3z), \quad P_4(z) = \frac{1}{8}(35z^4-30z^2+3). \tag{D.3}$$

Let β be a decomposition into two clusters $\mathcal{C}_1$ and $\mathcal{C}_2$ with $\sharp\mathcal{C}_1 = Z_1$ and $\sharp\mathcal{C}_2 = Z_2$. The intercluster interaction is given by

$$I_\beta(x) = -\sum_{i\in\mathcal{C}_1}\frac{e^2Z_2}{|x_i-X_2|} - \sum_{j\in\mathcal{C}_2}\frac{e^2Z_1}{|x_j-X_1|} + \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}}\frac{e^2}{|x_i-x_j|} + \frac{e^2Z_1Z_2}{|D|}. \tag{D.4}$$

For $i_k \in \mathcal{C}_1$ we define

$$\mathcal{F}_n^{(1)}(x) := \sum_{i\in\mathcal{C}_1}|x_i|^nP_n\left(\frac{x_i}{|x_i|}\cdot\frac{D}{|D|}\right), \tag{D.5}$$

$$\mathcal{F}_n^{(2)}(x) := \sum_{j\in\mathcal{C}_2}|x_j|^nP_n\left(\frac{-x_j}{|x_j|}\cdot\frac{D}{|D|}\right) \tag{D.6}$$

and

$$\mathcal{F}_n^{(3)}(x) := \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}}|x_i-x_j|^nP_n\left(\frac{x_i-x_j}{|x_i-x_j|}\cdot\frac{D}{|D|}\right). \tag{D.7}$$

Let

$$f_n(x) := -e^2Z_2\mathcal{F}_n^{(1)}(x) - e^2Z_1\mathcal{F}_n^{(2)}(x) + e^2\mathcal{F}_n^{(3)}(x). \tag{D.8}$$

Note that for $n = 2, 3$ the functions defined in (D.8) are the same as f_2 and f_3 in (9.18) and (9.19). Observe that $x \in \mathrm{supp}(J_\beta)$ implies $|x_i - X_1| << |D|$ for $i \in \mathcal{C}_1$ and $|x_j - X_2| << |D|$ for $j \in \mathcal{C}_2$ and the Taylor series of I_β converges. This yields

$$I_\beta(x) = \sum_{n=0}^{\infty}\frac{\mathcal{U}_\beta^*f_n(x)}{|D|^{n+1}} + \frac{e^2Z_1Z_2}{|D|} \qquad \forall x \in \mathrm{supp}(J_\beta) \tag{D.9}$$

where $\mathcal{U}_\beta$ is defined in (9.16).

Lemma D.1. *For any decomposition $\beta \in \mathcal{D}^{at}$*

$$\frac{\mathcal{U}_\beta^*f_0(x)}{|D|} + \frac{e^2Z_1Z_2}{|D|} = 0 \quad \textit{and} \quad f_1(x) = 0. \tag{D.10}$$

Proof. For $\beta \in \mathcal{D}^{at}$ we have $\sharp\mathcal{C}_1 = Z_1$ and $\sharp\mathcal{C}_2 = Z_2$. Since $P_0(z) = 0$, by (D.5) - (D.7) we get

$$\mathcal{F}_0^{(1)}(x) = Z_1, \quad \mathcal{F}_0^{(2)}(x) = Z_2, \text{ and } \mathcal{F}_0^{(3)}(x) = Z_1 Z_2. \tag{D.11}$$

By definition of f_0 in (D.8) this implies

$$\mathcal{U}_\beta^* f_0(x) = \mathcal{U}_\beta^*(-e^2 Z_2 Z_1) = -e^2 Z_2 Z_1 \tag{D.12}$$

which proves the first part of the lemma. Since $P_1(z) = z$, writing $e_D := \frac{D}{|D|}$ we have

$$\begin{aligned} \mathcal{F}_1^{(1)}(x) &= \sum_{i\in\mathcal{C}_1} x_i \cdot e_D, \\ \mathcal{F}_1^{(2)}(x) &= \sum_{j\in\mathcal{C}_2} -x_j \cdot e_D \end{aligned} \tag{D.13}$$

and

$$\mathcal{F}_1^{(3)}(x) = \sum_{\substack{i\in\mathcal{C}_1 \\ j\in\mathcal{C}_2}} (x_i - x_j) \cdot e_D. \tag{D.14}$$

By definition

$$\mathcal{U}_\beta^* f_1(x) = \mathcal{U}_\beta^*\Big(-\sum_{i\in\mathcal{C}_1} e^2 Z_2 (x_i \cdot e_D) - \sum_{j\in\mathcal{C}_2} e^2 Z_1 (-x_j \cdot e_D) + \sum_{\substack{i\in\mathcal{C}_1 \\ j\in\mathcal{C}_2}} e^2 [(x_i - x_j)\cdot e_D] \Big) = 0. \tag{D.15}$$

□

In the next lemma we will establish a bound of the remainder in the Taylor expansion of I_β. Let us define the potential

$$I_\beta^o(x) := (I_\beta \chi_{J_\beta})(x) \tag{D.16}$$

where $\chi_{J_\beta}(x)$ is the characteristic function of the support of J_β.

Lemma D.2. *Let $\beta \in \mathcal{D}^{at}$, then for any $k \geq 2$ there exists a constant $0 < C < \infty$ such that for $x \in \operatorname{supp}(I_\beta^o)$ we have*

$$\Big| I_\beta^o(x) - \sum_{n=2}^{k-1} \frac{\mathcal{U}_\beta^* f_n(x)}{|D|^{n+1}} \Big| \leq C \frac{\big(d_\beta(x)\big)^k}{|D|^{k+1}} \tag{D.17}$$

where

$$d_\beta(x) := \Big(\sum_{l=1,2} \sum_{i\in\mathcal{C}_l} |x_i - X_l|^2 \Big)^{\frac{1}{2}}. \tag{D.18}$$

Remark D.3. *Notice that $d_\beta(\cdot)$ characterizes how far away the particles in $\mathcal{C}_1$ and $\mathcal{C}_2$ are from their respective nucleus. This norm does not depend on the distance $|D|$ between the nuclei. In particular*

$$\mathcal{U}_\beta d_\beta(\cdot) = \|\cdot\|. \tag{D.19}$$

Proof. Note that for $k = 2$ the sum on the l.h.s of (D.17) is the empty sum which, by convention, is zero. The k-th summand of the Taylor expansion of I_β^o is

$$\begin{aligned} &\mathcal{U}_\beta^*\Big(-e^2 Z_2 \sum_{i\in\mathcal{C}_1} P_k\left(\frac{x_i}{|x_i|}\cdot e_D\right) \frac{|x_i|^k}{|D|^{k+1}} - e^2 Z_1 \sum_{j\in\mathcal{C}_2} P_k\left(\frac{-x_j}{|x_j|}\cdot e_D\right) \frac{|x_j|^k}{|D|^{k+1}} \\ &+ e^2 \sum_{\substack{i\in\mathcal{C}_1 \\ j\in\mathcal{C}_2}} P_k\left(\frac{x_i - x_j}{|x_i - x_j|}\cdot e_D\right) \frac{|x_i - x_j|^k}{|D|^{k+1}} \Big). \end{aligned} \tag{D.20}$$

We apply the Taylor theorem with a remainder in Lagrange form. Since the Legendre polynomials take values between -1 and 1 on the interval $[-1,1]$, the Lagrange form remainders are bounded above by one. Consequently

$$\Big|I^o_\beta - \sum_{n=2}^{k-1} \frac{\mathcal{U}^*_\beta f_n}{|D|^{n+1}}\Big| \le \sum_{i\in\mathcal{C}_1} \frac{e^2 Z_2 |x_i - X_1|^k}{|D|^{k+1}} + \sum_{j\in\mathcal{C}_2} \frac{e^2 Z_1 |x_j - X_2|^k}{|D|^{k+1}} + \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} \frac{e^2 |(x_i - X_1) - (x_j - X_2)|^k}{|D|^{k+1}}$$

and there exists a constant C such that

$$\Big|I^o_\beta(x) - \sum_{n=2}^{k-1} \frac{\mathcal{U}^*_\beta f_n(x)}{|D|^{n+1}}\Big| \le C \frac{\big(d_\beta(x)\big)^k}{|D|^{k+1}} \quad \forall x \in \mathrm{supp}(I^o_\beta). \tag{D.21}$$

□

Corollary D.4. *Let $\beta \in \mathcal{D}^{at}$ and $\varphi_1,\ \varphi_2 \in L^2(\mathbb{R}^{3N})$ such that there exists $b > 0$ and A_0 with*

$$\|\varphi_2(\cdot)e^{b|\cdot|}\|^2 \le A_0 \|\varphi_2\|^2. \tag{D.22}$$

Then for any $k \ge 2$ there exists a constant $C_k(b, A_0) < \infty$ such that

$$\Big|\Big\langle \mathcal{U}^*_\beta \varphi_1, \Big(I^o_\beta - \sum_{n=2}^{k-1} \frac{\mathcal{U}^*_\beta f_n}{|D|^{n+1}}\Big)\mathcal{U}^*_\beta \varphi_2\Big\rangle\Big| \le C_k |D|^{-(k+1)} \|\varphi_1\|\ \|\varphi_2\|. \tag{D.23}$$

Proof. To prove (D.23) we apply Lemma D.2 to get

$$\Big|\Big\langle \mathcal{U}^*_\beta \varphi_1, \Big(I^o_\beta - \sum_{n=2}^{k-1} \frac{\mathcal{U}^*_\beta f_n}{|D|^{n+1}}\Big)\mathcal{U}^*_\beta \varphi_2\Big\rangle\Big| \le C\big|\langle \mathcal{U}^*_\beta \varphi_1, \frac{\big(d_\beta(\cdot)\big)^k}{|D|^{k+1}} \mathcal{U}^*_\beta \varphi_2\rangle\big| \tag{D.24}$$

and by (D.19) we arrive at

$$C\big|\langle \mathcal{U}^*_\beta \varphi_1, \frac{\big(d_\beta(\cdot)\big)^k}{|D|^{k+1}} \mathcal{U}^*_\beta \varphi_2\rangle\big| = C\big|\langle \varphi_1, \frac{\|\cdot\|^k}{|D|^{k+1}} \mathcal{U}_\beta \mathcal{U}^*_\beta \varphi_2\rangle\big|. \tag{D.25}$$

Now (D.23) follows, using the Cauchy-Schwarz inequality and the exponential decay of φ_2 from assumption (D.22). □

To simplify the notation in the remainder of the section, we set

$$\tilde\phi := \mathcal{U}^*_\beta \phi, \quad \tilde\phi_2 := \mathcal{U}^*_\beta \phi_2, \quad \tilde\phi_3 := \mathcal{U}^*_\beta \phi_3 \text{ and } \tilde g := \mathcal{U}^*_\beta g. \tag{D.26}$$

Lemma D.5. *Let $\phi, \phi_2, \phi_3, g \in \mathcal{H}^\alpha$ and $\gamma_1, \gamma_2, \gamma_3 \in \mathbb{C}$ be as defined in* (12.16)-(12.19)*. For any fixed $\delta > 0$, there exist $C > 0, D_0 > 0$ such that for $|D| > D_0$*

$$\begin{aligned}\langle J_\beta\psi, I_\beta J_\beta \psi\rangle \ge{}& 2|D|^{-6}\,\mathrm{Re}\,\gamma_1\overline{\gamma_2}\|\phi_2\|_1^2 + 2|D|^{-8}\,\mathrm{Re}\,\gamma_1\overline{\gamma_3}\|\phi_3\|_1^2 \\ &- C\frac{|\gamma_1|^2 + |\gamma_2|^2 + |\gamma_3|^2}{|D|^{10}} - \delta\|g\|^2.\end{aligned} \tag{D.27}$$

Proof. Note that by definition of I^o_β in (D.16) one has

$$\langle J_\beta\psi, I_\beta J_\beta\psi\rangle = \langle J_\beta\psi, I^o_\beta J_\beta\psi\rangle, \tag{D.28}$$

and, according to (12.19),

$$\begin{aligned}J_\beta\psi &= \mathcal{U}^*_\beta\big(\gamma_1\phi + |D|^{-3}\gamma_2\phi_2 + |D|^{-4}\gamma_3\phi_3 + g\big) \\ &= \gamma_1\tilde\phi + |D|^{-3}\gamma_2\tilde\phi_2 + |D|^{-4}\gamma_3\tilde\phi_3 + \tilde g.\end{aligned} \tag{D.29}$$

Using this we can split the expression on the r.h.s of (D.28) into the terms

$$\begin{aligned}
&\langle J_\beta\psi, I_\beta^o J_\beta\psi\rangle \\
&= |\gamma_1|^2\langle\tilde\phi, I_\beta^o\tilde\phi\rangle + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^3}\langle\tilde\phi_2, I_\beta^o\tilde\phi\rangle + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^4}\langle\tilde\phi_3, I_\beta^o\tilde\phi\rangle \\
&\quad + 2\operatorname{Re}\gamma_1\langle\tilde g, I_\beta^o\tilde\phi\rangle + \frac{|\gamma_2|^2}{|D|^6}\langle\tilde\phi_2, I_\beta^o\tilde\phi_2\rangle + \frac{2\operatorname{Re}\gamma_2\overline{\gamma_3}}{|D|^7}\langle\tilde\phi_3, I_\beta^o\tilde\phi_2\rangle \\
&\quad + \frac{2\operatorname{Re}\gamma_2}{|D|^3}\langle\tilde g, I_\beta^o\tilde\phi_2\rangle + \frac{|\gamma_3|^2}{|D|^8}\langle\tilde\phi_3, I_\beta^o\tilde\phi_3\rangle + \frac{2\operatorname{Re}\gamma_3}{|D|^4}\langle\tilde g, I_\beta^o\tilde\phi_3\rangle + \langle\tilde g, I_\beta^o\tilde g\rangle \\
&= B_1 + B_2 + B_3 + \langle\tilde g, I_\beta^o\tilde g\rangle,
\end{aligned} \tag{D.30}$$

where

$$B_1 := |\gamma_1|^2\langle\tilde\phi, I_\beta^o\tilde\phi\rangle + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^3}\langle\tilde\phi_2, I_\beta^o\tilde\phi\rangle + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^4}\langle\tilde\phi_3, I_\beta^o\tilde\phi\rangle \tag{D.31}$$

$$B_2 := \frac{|\gamma_2|^2}{|D|^6}\langle\tilde\phi_2, I_\beta^o\tilde\phi_2\rangle + \frac{2\operatorname{Re}\gamma_2\overline{\gamma_3}}{|D|^7}\langle\tilde\phi_3, I_\beta^o\tilde\phi_2\rangle + \frac{|\gamma_3|^2}{|D|^8}\langle\tilde\phi_3, I_\beta^o\tilde\phi_3\rangle \tag{D.32}$$

and

$$B_3 := 2\operatorname{Re}\gamma_1\langle\tilde g, I_\beta^o\tilde\phi\rangle + \frac{2\operatorname{Re}\gamma_2}{|D|^3}\langle\tilde g, I_\beta^o\tilde\phi_2\rangle + \frac{2\operatorname{Re}\gamma_3}{|D|^4}\langle\tilde g, I_\beta^o\tilde\phi_3\rangle. \tag{D.33}$$

In Propositions D.6, D.7 and D.8 we bound these three terms separately. We obtain

$$B_1 \geq \frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^6}\|\phi_2\|_1^2 + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^8}\|\phi_3\|_1^2 - C\frac{|\gamma_1|^2+|\gamma_2|^2+|\gamma_3|^2}{|D|^{10}}, \tag{D.34}$$

$$B_2 \geq -C\frac{|\gamma_2|^2+|\gamma_3|^2}{|D|^{10}} \tag{D.35}$$

and we show that for any $\delta > 0$ there exist $C, D_0 > 0$ such that for all $|D| > D_0$ we have

$$B_3 \geq -C\frac{|\gamma_1|^2}{|D|^{10}} - C\frac{|\gamma_2|^2}{|D|^{12}} - C\frac{|\gamma_3|^2}{|D|^{14}} - \frac{\delta}{2}\|g\|^2. \tag{D.36}$$

For the term $\langle\tilde g, I_\beta^o\tilde g\rangle$ in (D.30) we use the fact that on the support of J_β, the distance between particles belonging to different subsystems grows proportionally to $|D|$. Thus for any $\delta > 0$ we can choose $D_0 > 0$ such that for $|D| > D_0$

$$\langle\tilde g, I_\beta^o\tilde g\rangle \geq -\frac{\delta}{2}\|\tilde g\|^2 = -\frac{\delta}{2}\|g\|^2. \tag{D.37}$$

Collecting the estimates (D.34) - (D.37) proves the lemma. □

Proposition D.6 (Estimate of B_1)**.** *We have*

$$B_1 \geq \frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^6}\|\phi_2\|_1^2 + \frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^8}\|\phi_3\|_1^2 - C\frac{|\gamma_1|^2+|\gamma_2|^2+|\gamma_3|^2}{|D|^{10}}. \tag{D.38}$$

Proof. By Condition 2) of Theorem 9.2, for each $\alpha' \prec \alpha$, the functions in $P^{\alpha'}\tilde{\mathcal{W}}_\beta^\alpha$ transform according to the $\ell = 0$ degree irreducible representation of $SO(3)$. In particular the one electron densities are spherically symmetric with respect to their associated nucleus, see [3]. Due to mutual orthogonality of the spaces $P^{\alpha'}\tilde{\mathcal{W}}_\beta^\alpha$ for different α', all functions in $\tilde{\mathcal{W}}_\beta^\alpha$ have this property. Applying Newton's theorem ([24, Theorem 9.7]) we get

$$|\gamma_1|^2\langle\tilde\phi, I_\beta^o\tilde\phi\rangle = 0. \tag{D.39}$$

For the second term of (D.31), by Lemma D.1 and Lemma D.2 we get

$$\langle\tilde{\phi}_2, I^o_\beta\tilde{\phi}\rangle \geq |D|^{-3}\langle\tilde{\phi}_2, (\mathcal{U}^*_\beta f_2)\tilde{\phi}\rangle + |D|^{-4}\langle\tilde{\phi}_2, (\mathcal{U}^*_\beta f_3)\tilde{\phi}\rangle + |D|^{-5}\langle\tilde{\phi}_2, (\mathcal{U}^*_\beta f_4)\tilde{\phi}\rangle + |D|^{-6}\langle\tilde{\phi}_2, (\mathcal{U}^*_\beta f_5)\tilde{\phi}\rangle - C\Big|\Big\langle\tilde{\phi}_2, \frac{\big(d_\beta(\cdot)\big)^6}{|D|^7}\tilde{\phi}\Big\rangle\Big|. \tag{D.40}$$

Notice that for $l = 2, 3, 4, 5$ we have

$$\langle\tilde{\phi}_2, (\mathcal{U}^*_\beta f_l)\tilde{\phi}\rangle = \langle\mathcal{U}^*_\beta\phi_2, (\mathcal{U}^*_\beta f_l)(\mathcal{U}^*_\beta\phi)\rangle = \langle\phi_2, f_l\phi\rangle. \tag{D.41}$$

We will use the following orthogonality relations between ϕ_2 and $f_l\phi$, $l = 3, 4, 5$ from Lemma E.6:

$$\langle\phi_2, f_3\phi\rangle = \langle\phi_2, f_4\phi\rangle = \langle\phi_2, f_5\phi\rangle = 0. \tag{D.42}$$

This implies

$$\langle\tilde{\phi}_2, I^o_\beta\tilde{\phi}\rangle \geq |D|^{-3}\langle\phi_2, f_2\phi\rangle - C\Big|\Big\langle\tilde{\phi}_2, \frac{\big(d_\beta(\cdot)\big)^6}{|D|^7}\tilde{\phi}\Big\rangle\Big|. \tag{D.43}$$

Note that by Remark D.3 and due to exponential decay of the function ϕ we have

$$C\Big|\langle\tilde{\phi}_2, \frac{\big(d_\beta(\cdot)\big)^6}{|D|^7}\tilde{\phi}\rangle\Big| \leq C|D|^{-7}\|\phi_2\|\|\phi\|. \tag{D.44}$$

By definition of the semi-norm, see (12.15), $\langle\phi_2, f_2\phi\rangle = \|\phi_2\|_1^2$ and since $2\operatorname{Re}\gamma_1\overline{\gamma_2} \leq |\gamma_1|^2 + |\gamma_2|^2$ we get

$$\frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^3}\langle\tilde{\phi}_2, I^o_\beta\tilde{\phi}\rangle \geq \frac{2\operatorname{Re}\gamma_1\overline{\gamma_2}}{|D|^6}\|\phi_2\|_1^2 - C\frac{|\gamma_1|^2 + |\gamma_2|^2}{|D|^{10}}. \tag{D.45}$$

Now we estimate the last term in (D.31). Since ϕ decays exponentially we can apply Corollary D.4 with $k = 5$ and proceeding as in (D.41) yields

$$\langle\tilde{\phi}_3, I^o_\beta\tilde{\phi}\rangle \geq |D|^{-3}\langle\phi_3, f_2\phi\rangle + |D|^{-4}\langle\phi_3, f_3\phi\rangle + |D|^{-5}\langle\phi_3, f_4\phi\rangle - C|D|^{-6}\|\phi_2\|\ \|\phi\|. \tag{D.46}$$

According to Lemma E.6 the first and third summand of (D.46) are zero and we get

$$\frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^4}\langle\tilde{\phi}_3, I^o_\beta\tilde{\phi}\rangle \geq \frac{2\operatorname{Re}\gamma_1\overline{\gamma_3}}{|D|^8}\|\phi_3\|_1^2 - C\frac{|\gamma_1|^2 + |\gamma_3|^2}{|D|^{10}}. \tag{D.47}$$

□

Proposition D.7 (Estimate of B_2)**.** *There exists a constant $C > 0$ such that*

$$B_2 \geq -C\frac{|\gamma_2|^2 + |\gamma_3|^2}{|D|^{10}} \tag{D.48}$$

Proof. Recall that

$$B_2 = \frac{|\gamma_2|^2}{|D|^6}\langle\tilde{\phi}_2, I^o_\beta\tilde{\phi}_2\rangle + \frac{2\operatorname{Re}\gamma_2\overline{\gamma_3}}{|D|^7}\langle\tilde{\phi}_3, I^o_\beta\tilde{\phi}_2\rangle + \frac{|\gamma_3|^2}{|D|^8}\langle\tilde{\phi}_3, I^o_\beta\tilde{\phi}_3\rangle. \tag{D.49}$$

For the first term on the r.h.s. of (D.49), since ϕ_2 decays exponentially (see Corollary 10.4), we can use Corollary D.4 with $k = 3$ and the analogous to (D.41) we get

$$\langle\tilde{\phi}_2, I^o_\beta\tilde{\phi}_2\rangle \geq |D|^{-3}\langle\phi_2, f_2\phi_2\rangle - C|D|^{-4}\|\phi_2\|^2. \tag{D.50}$$

By Lemma E.7 we have $\langle\phi_2, f_2\phi_2\rangle = 0$ which implies

$$\frac{|\gamma_2|^2}{|D|^6}\langle\tilde{\phi}_2, I^o_\beta\tilde{\phi}_2\rangle \geq -C\frac{|\gamma_2|^2}{|D|^{10}}. \tag{D.51}$$

To bound the second and third term on the r.h.s. of (D.49) we apply Corollary D.4 with $k = 2$ to get

$$\frac{2\operatorname{Re}\gamma_2\overline{\gamma_3}}{|D|^7}\langle\tilde{\phi_3}, I_\beta^o\tilde{\phi_2}\rangle \geq -C\frac{|\gamma_2|^2+|\gamma_3|^2}{|D|^{10}} \tag{D.52}$$

and

$$\frac{|\gamma_3|^2}{|D|^8}\langle\tilde{\phi_3}, I_\beta^o\tilde{\phi_3}\rangle \geq -C\frac{|\gamma_3|^2}{|D|^{11}}. \tag{D.53}$$

□

Proposition D.8 (Estimate of B_3)**.** *For any fixed $\delta > 0$ there exist $C > 0$ and $D_0 > 0$ such that for $|D| > D_0$ we have*

$$B_3 \geq -C\frac{|\gamma_1|^2}{|D|^{10}} - C\frac{|\gamma_2|^2}{|D|^{12}} - C\frac{|\gamma_3|^2}{|D|^{14}} - \frac{\delta}{2}\|g\|^2. \tag{D.54}$$

Proof. Recall

$$B_3 = 2\operatorname{Re}\gamma_1\langle\tilde{g}, I_\beta^o\tilde{\phi}\rangle + \frac{2\operatorname{Re}\gamma_2}{|D|^3}\langle\tilde{g}, I_\beta^o\tilde{\phi_2}\rangle + \frac{2\operatorname{Re}\gamma_3}{|D|^4}\langle\tilde{g}, I_\beta^o\tilde{\phi_3}\rangle. \tag{D.55}$$

For the first term, by Corollary D.4 with $k = 4$ and the analogous to (D.41) we get

$$2\operatorname{Re}\gamma_1\langle\tilde{g}, I_\beta^o\tilde{\phi}\rangle \geq 2\operatorname{Re}\gamma_1|D|^{-3}\langle g, f_2\phi\rangle + 2\operatorname{Re}\gamma_1|D|^{-4}\langle g, f_3\phi\rangle - C|\gamma_1||D|^{-5}\|g\|\|\phi\| \tag{D.56}$$

where by definition of g we have

$$\langle g, f_2\phi\rangle = \langle g, \phi_2\rangle_1 = 0 \tag{D.57}$$

and

$$\langle g, f_3\phi\rangle = \langle g, \phi_3\rangle_1 = 0. \tag{D.58}$$

This implies

$$2\operatorname{Re}\gamma_1\langle\tilde{g}, I_\beta^o\tilde{\phi}\rangle \geq -C|\gamma_1||D|^{-5}\|g\|\|\phi\|. \tag{D.59}$$

By Corollary D.4 with $k = 2$ we get

$$\frac{2\operatorname{Re}\gamma_2}{|D|^3}\langle\tilde{g}, I_\beta^o\tilde{\phi_2}\rangle \geq -C|\gamma_2||D|^{-6}\|g\|\|\phi_2\| \tag{D.60}$$

and

$$\frac{2\operatorname{Re}\gamma_3}{|D|^4}\langle\tilde{g}, I_\beta^o\tilde{\phi_3}\rangle \geq -C|\gamma_3||D|^{-7}\|g\|\|\phi_3\|. \tag{D.61}$$

Applying Young's inequality for products in (D.59)-(D.61) yields the result. □

Appendix E. Orthogonality relations

In this subsection we prove several orthogonality relations, which follow from the symmetry properties of functions in $\tilde{\mathcal{W}}_\beta^\alpha$. Let $\mathcal{P}^{(i)} : L^2(\mathbb{R}^{3N}) \to L^2(\mathbb{R}^{3N})$ such that

$$(\mathcal{P}^{(i)}\varphi)(x) := \varphi(x_1, \cdots, x_{i-1}, -x_i, x_{i+1}, \cdots, x_N) \tag{E.1}$$

and define

$$\mathcal{P}_{\mathcal{C}_1} := \prod_{i\in\mathcal{C}_1}\mathcal{P}^{(i)}.$$

As usual we say that a function $\varphi \in L^2(\mathbb{R}^{3N})$ is $\mathcal{P}_{\mathcal{C}_1}$-*symmetric* iff $\mathcal{P}_{\mathcal{C}_1}\varphi = \varphi$. A function $\varphi \in L^2(\mathbb{R}^{3N})$ is called $\mathcal{P}_{\mathcal{C}_1}$-*antisymmetric* iff $\mathcal{P}_{\mathcal{C}_1}\varphi = -\varphi$. Similarly, we define the operator $\mathcal{P}_{\mathcal{C}_2}$ and set

$$\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2} := \mathcal{P}_{\mathcal{C}_1}\mathcal{P}_{\mathcal{C}_2}.$$

Lemma E.1. *Let $\alpha' \prec \alpha$ an irreducible representation of S_β be such that $P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta \neq \emptyset$. For $\mathcal{P}_\bullet = \mathcal{P}_{\mathcal{C}_1}, \mathcal{P}_{\mathcal{C}_2}, \mathcal{P}_{\mathcal{C}_1\mathcal{C}_2}$ we have*

$$\begin{aligned} \textit{either: all functions } \phi \in P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta \textit{ are } \mathcal{P}_\bullet\textit{-symmetric} \\ \textit{or: all functions } \phi \in P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta \textit{ are } \mathcal{P}_\bullet\textit{-antisymmetric.} \end{aligned}$$

Proof. From the definition of $\tilde{H}_\beta$ it is apparent that $\mathcal{P}_\bullet \tilde{H}_\beta \mathcal{P}_\bullet = \tilde{H}_\beta$. Consequently the $\mathcal{P}_\bullet$-symmetric and the $\mathcal{P}_\bullet$-antisymmetric functions are invariant subspaces of $\tilde{H}_\beta$. By Condition 2) we have $\dim(P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta) = \dim \alpha'$ and since α' is irreducible it can not contain nontrivial invariant subspaces, so either all functions in $P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta$ are $\mathcal{P}_\bullet$-symmetric or all functions in $P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta$ are $\mathcal{P}_\bullet$-antisymmetric. □

Lemma E.2. *For any $\phi \in \tilde{\mathcal{W}}^\alpha_\beta$ we have*

$$\langle \phi, f_2\phi\rangle = \langle \phi, f_3\phi\rangle = 0 \tag{E.2}$$

Proof. Recall the definitions

$$f_2(x) = \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} -e^2\big(3(x_i\cdot e_D)(x_j\cdot e_D) - x_i\cdot x_j\big) \tag{E.3}$$

and

$$\begin{aligned} f_3(x) = \sum_{\substack{i\in\mathcal{C}_1\\ j\in\mathcal{C}_2}} \frac{e^2}{2}\Big(&3(x_i - x_j)\cdot e_D\big[2(x_i\cdot x_j) - 5(x_i\cdot e_D)(x_j\cdot e_D)\big] \\ &+ 3|x_i|^2(x_j\cdot e_D) - 3|x_j|^2(x_i\cdot e_D)\Big). \end{aligned} \tag{E.4}$$

It is easy to see that f_2 is $\mathcal{P}_{\mathcal{C}_1}$-antisymmetric and $\mathcal{P}_{\mathcal{C}_2}$-antisymmetric. Note that f_2 is invariant under permutations in S_β which preserve the cluster decomposition β. Hence multiplication by f_2 commutes with the projection $P^{\alpha'}$. Since the spaces $P^{\alpha'}\tilde{\mathcal{W}}^\alpha_\beta$ are mutually orthogonal for different α', for all $\phi \in \tilde{\mathcal{W}}^\alpha_\beta$ we have

$$\langle \phi, f_2\phi\rangle = \sum_{\alpha'\prec\alpha} \langle P^{\alpha'}\phi, f_2 P^{\alpha'}\phi\rangle. \tag{E.5}$$

Since $|P^{\alpha'}\phi|^2$ is $\mathcal{P}_{\mathcal{C}_1}$-symmetric and f_2 is $\mathcal{P}_{\mathcal{C}_1}$-antisymmetric we get

$$\langle P^{\alpha'}\phi, f_2 P^{\alpha'}\phi\rangle = 0. \tag{E.6}$$

Similarly, from the explicit expression of f_3 in (E.4) follows

$$\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2} f_3 = -f_3 \tag{E.7}$$

which yields

$$\langle \phi, f_3\phi\rangle = \sum_{\alpha'\prec\alpha} \langle P^{\alpha'}\phi, f_3 P^{\alpha'}\phi\rangle = \sum_{\alpha'\prec\alpha} \langle P^{\alpha'}\phi, (\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2} f_3) P^{\alpha'}\phi\rangle = 0. \tag{E.8}$$

□

Corollary E.3. *For any $\phi \in \tilde{\mathcal{W}}^\alpha_\beta$ the functions*

$$\phi_k := (\tilde{H}^\alpha_\beta - \mu^\alpha)^{-1} f_k \phi, \quad k = 2, 3 \tag{E.9}$$

are well defined.

Proof. This is an immediate consequence of Lemma E.2, since it states that $f_k\phi$ is orthogonal to ϕ. □

Corollary E.4. *For any $\phi \in \tilde{\mathcal{W}}^\alpha_\beta$ we have*

$$\langle \phi, \phi_2\rangle_1 = \langle \phi, \phi_3\rangle_1 = 0. \tag{E.10}$$

Lemma E.5. *For any $\phi \in \tilde{\mathcal{W}}_\beta^\alpha$ we have*

$$\langle \phi_2, f_2\phi_2 \rangle = 0. \tag{E.11}$$

Proof. By the same argument used in Lemma E.2, since f_2 appears three times in the expression

$$\langle \phi_2, f_2\phi_2 \rangle = \langle (\tilde{H}_\beta - \mu^\alpha)^{-1} f_2\phi, f_2(\tilde{H}_\beta - \mu^\alpha)^{-1} f_2\phi \rangle \tag{E.12}$$

applying $\mathcal{P}_{\mathcal{C}_1}$ results in a change of sign which yields the result. □

Lemma E.6. *For any $\phi \in \tilde{\mathcal{W}}_\beta^\alpha$ and with $\phi_2,\ \phi_3$ defined in Corollary E.3 we have*

$$\begin{aligned} &i) \quad \langle \phi_2, \phi_3 \rangle = \langle \phi_2, f_3\phi \rangle = \langle \phi_2, f_5\phi \rangle = 0 \\ &ii) \quad \langle \phi_3, f_2\phi \rangle = \langle \phi_3, f_4\phi \rangle = 0. \end{aligned} \tag{E.13}$$

Proof. Notice that the Legendre polynomials fulfill

$$P_n(-z) = (-1)^n P_n(z). \tag{E.14}$$

In particular for $h, D \in \mathbb{R}^3$ we get

$$P_n\left(\frac{-h}{|h|} \cdot \frac{D}{|D|}\right) = (-1)^n P_n\left(\frac{h}{|h|} \cdot \frac{D}{|D|}\right) \tag{E.15}$$

and thus

$$\big(\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2} f_n\big)(x) = (-1)^n f_n(x). \tag{E.16}$$

Hence

$$\langle f_2\phi, f_3\phi \rangle = \big\langle \big(\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2} f_2\big)\phi, \big(\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2} f_3\big)\phi \big\rangle = -\langle f_2\phi, f_3\phi \rangle = 0. \tag{E.17}$$

Analogously

$$\langle f_2\phi, f_5\phi \rangle = \langle f_3\phi, f_4\phi \rangle = 0. \tag{E.18}$$

Since $\tilde{H}_\beta$ commutes with $\mathcal{P}_{\mathcal{C}_1\mathcal{C}_2}$, so do $(\tilde{H}_\beta - \mu^\alpha)$ and $(\tilde{H}_\beta - \mu^\alpha)^{-1}$. Hence by the same argument we also get

$$\langle \phi_2, \phi_3 \rangle = \langle \phi_2, f_3\phi \rangle = \langle \phi_2, f_5\phi \rangle = \langle \phi_3, f_2\phi \rangle = \langle \phi_3, f_4\phi \rangle = 0. \tag{E.19}$$

□

In the next lemma we will use the $SO(3)$ symmetry of the system.

Lemma E.7. *For any $\phi \in \tilde{\mathcal{W}}_\beta^\alpha$ and $\phi_2 = (\tilde{H}_\beta - \mu^\alpha)^{-1} f_2\phi$ we have*

$$\langle \phi_2, f_4\phi \rangle = 0. \tag{E.20}$$

Proof. As a first step we notice that the functions f_2 and f_4 are the sums of Legendre polynomials of degrees 2 and 4 respectively. For the Legendre polynomials P_k of order k and the spherical harmonics Y_ℓ^m we have

$$P_\ell(\cos\theta) = \sqrt{\frac{4\pi}{(2\ell+1)}} Y_\ell^0(\theta, \varphi).$$

Note that in (D.5)-(D.7), leading to the definition of f_n in (D.8), for $\mathcal{F}_n^{(1)}$ we have $\cos\theta = \frac{x_i}{|x_i|} \cdot \frac{D}{|D|}$, for $\mathcal{F}_n^{(2)}$ we have $\cos\theta = \frac{-x_j}{|x_j|} \cdot \frac{D}{|D|}$ and in $\mathcal{F}_n^{(3)}$ we have $\cos\theta = \frac{x_i - x_j}{|x_i - x_j|} \cdot \frac{D}{|D|}$ respectively. Consequently the Legendre polynomials of order ℓ are transformed according to the irreducible representation of degree ℓ under the actions of the $SO(3)$ group, see [18].

By Condition 2) of Theorem 9.2, the state ϕ belongs to the irreducible representation of degree $\ell = 0$ of the group $SO(3)$. Thus the products $f_2\phi$ and $f_4\phi$ are transformed according to the representations of degree $\ell = 2$ and $\ell = 4$ respectively.

By rotational invariance of the operator $\tilde{H}_\beta$, the function $(\tilde{H}_\beta - \mu^\alpha)^{-1} f_2\phi$ has the same symmetry a $f_2\phi$, namely it transforms according to the irreducible representation of degree $\ell = 2$.

But functions belonging to two different irreducible representations are orthogonal. This proves the lemma. □

Appendix F. Remark on actions of the permutation group

Let $g \in L^2(\mathbb{R}^{3(m+n)})$ be a function depending on position vectors of $(m+n)$ particles. Let A be an operator on $L^2(\mathbb{R}^{3m})$ and $g \in \mathcal{D}(A \otimes \mathbb{1}^{3n})$, so that $A \otimes \mathbb{1}^{3n}$ acts on g as a function of the first m positon vectors.

Lemma F.1. *Assume that for some $R > 0$ we have* $\mathrm{supp}(g) \subset \{\xi \in \mathbb{R}^{3(m+n)}, |\xi_i| < R\ i = 1, \cdots, m, |\xi_j| \geq 2R\ j \geq m+1\}$. *Let S_{m+n} be the permutation group of $(m+n)$ particles and $\pi \in S_{m+n}$ such that $\pi \notin S_m \otimes S_n$. In other words π exchanges at least one of the first m particles with a particle labelled by $j \geq m+1$. Then*

$$\mathrm{supp}\left((A \otimes \mathbb{1}^{3n})g\right) \cap \mathrm{supp}\left(\mathcal{T}_\pi g\right) = \emptyset \tag{F.1}$$

where $\mathcal{T}_\pi g(\xi) = g(\xi_{\pi^{-1}(1)}, \cdots, \xi_{\pi^{-1}(m+n)})$.

Proof. For local operators A this relation was first used by Sigalov and Zhislin to prove existence of an eigenvalue of atoms with arbitrary types of rotational and permutational symmetry [50]. If the operator is local, (F.1) can be rewritten as

$$\mathrm{supp}(g) \cap \mathrm{supp}\left(\mathcal{T}_\pi g\right) = \emptyset. \tag{F.2}$$

If A is a non-local operator, (F.1) is still true, because for at least one particle $i_0 \geq m+1$ we have

$$|\xi_{i_0}| > 2R \quad \text{on } \mathrm{supp}\left((A \otimes \mathbb{1}^{3n})g\right) \tag{F.3}$$

and

$$|\xi_{i_0}| < R \quad \text{on } \mathrm{supp}\left(\mathcal{T}_\pi g\right). \tag{F.4}$$

□

Appendix G. Spin, Symmetries and Representation Theory

In the following section of this thesis we will give a brief introduction to the discovery of the electron spin followed by a description of the mathematical implementation into the quantum mechanical theory.

We give a summary of some of the important results of representation theory in an attempt to shed some light on this algebraic topic, that might lie within unfamiliar territory for mathematical physicists.

It is beyond the scope of this section to rigorously deduce the theory of group representations. The aim of the discussion is to convey some of the underlying concepts so that the applications of representation theory in the context of quantum mechanics can at least heuristically be understood without a profound knowledge of group theory.

G.1. **Quantum numbers.** Quantum numbers come from the formal description and enumeration of observed atomic spectra. Physicists attempted to give a formal description of the energy levels of an atom. The result were formulas that replicated the different energy levels in the observed atomic spectrum. Initially physicists used two integer variables and the charge of the atomic nucleus to desctribe an atomic spectrum. More explicitly, introducing the fine structure constant

$$\alpha = \frac{e^2}{\hbar c} \tag{G.1}$$

Sommerfeld came up with a formula for the energy levels of the relativistic hydrogen energy spectrum

$$E_{n,l} = -\operatorname{Ry}\left[\frac{1}{n^2} + \frac{\alpha^2}{n^3}\left(\frac{1}{l+1} - \frac{3}{4n}\right)\right] + \mathcal{O}(\alpha^4) \tag{G.2}$$

using two integer valued variables $n = 1, 2, \cdots$ and $l = 0, 1, \cdots, n-1$ and the Rydberg constant Ry.

These first formulas could only reproduce the rough structure of the spectrum. It was then observed, that the spectral lines would split under the influence of an external magnetic field. An effect known as the Zeeman effect. To accomodate for this spectral line splitting, more variables needed to be introduced. This is the origin of the quantum numbers as they are known today. These numbers did not come with a physical interpretation right away. Only later have these variables been associated to physical properties.

These quantum numbers are a 4-tuple consisting of the principal quantum number n, indicating the electron shell attributed to the electron which is directly linked to the distance to the nucleus; the azimuthal quantum number ℓ linked to the angular momentum of the electron; the magnetic quantum number m_ℓ which describes the orientation of the orbital of an electron relative to a fixed direction; and the spin projection quantum number m_s describing the spin of a quantum state.

The famous exclusion principle due to Pauli was formulated with the help of these quantum numbers. Pauli attempted to formalize the spectrum of alkali metals when he found that certain energy levels that were theoretically possible did not show on the measured spectral lines. He observed that for two electrons with angular momentum $\ell = 0$ there are four possible states if they have different principal quantum number (i.e. are in different electron shells) but only one possible state when their principal numbers agree. His explanation was that each state characterized by the four quantum numbers (n, ℓ, m_ℓ, s) can be occupied by at most one electron.

With the discovery that the wave-particle duality does not only apply for photons, but also for electrons, quantum mechanics changed significantly, shifting from the use of quantum numbers to utilizing a wave function $\psi \in L^2(\mathbb{R}^{3N}, \mathbb{C}^{2N})$ depending on position and spin of electrons to describe a quantum mechanical state for N electrons. Consequently the Pauli exclusion principle needed to be reformulated. In this setting the Pauli exclusion principle can be implemented as a restriction of the state space to the subspace $\bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ of functions that are antisymmetric with respect to any transposition of two particle variables. Before we discuss this specific antisymmetry restriction we will take a short detour to the general symmetric properties of the physical system and the implications thereof for the quantum mechanical system.

G.2. **Symmetries of the physical system.** The symmetries of a physical system should always be replicated by the theoretical model that is used to study the system or predict natural phenomena. By neglecting the intrinsic symmetries we could possibly artificially create dependences where there should not be any, which could lead to false conclusions and predictions. For example, the attraction of two massive objects should only depend on their relative distance and their masses and not on the point of view of the observer. We want similar properties for the quantum mechanical model of an atom. The specific requirements are discussed in this section.

We will look at the quantum mechanical model of a time-independent N electron wave function ψ depending on N particle variables (to be specified later) that describe every relevant property of the electron. The energy associated with a state

is modelled by a Hamiltonian operator H. More precisely, the expected value of the energy of a given state modelled by a state function ψ is given by the quadratic form

$$\langle \psi, H\psi \rangle. \tag{G.3}$$

The first requirement with respect to symmetry of the system is, that the point from which the system is observed should not have an impact on the energy of the system. So the system should be invariant under rotations of the space. The second requirement is that electrons should be viewed as identical particles, so interchanging electrons with identical properties should not alter the result.

Let us focus on the Hamiltonian of N electrons of charge $-e$ with position vectors $x_1, \cdots, x_N \in \mathbb{R}^3$ interacting with a nucleus of charge eZ located at the origin

$$\begin{aligned} H :=& \sum_{i=1}^{N} \left(-\Delta_i - \frac{e^2 Z}{|x_i|} \right) + \sum_{1 \leq i < j \leq N} \frac{e^2}{|x_i - x_j|} \\ =& -\Delta + V \end{aligned} \tag{G.4}$$

acting on $L^2(\mathbb{R}^{3N})$ where V is the collection of Coulomb interactions. This Hamiltonian depends solely on the electron position (relative to the nucleus) and does not depend on their spin. Let us now find a mathematically rigorous way to express the symmetries of the physical system in this mathematical model.

Exchange of particles is modelled by the permutation group and rotations of the system are modelled by the group of rotations in $\mathbb{R}^3$. However a quantum mechanical state is not described by a collection of elements that we can permute, but wave functions depending on the particle variables. We need to *represent* a permutation $\pi \in S_N$ of a set of N elements by an operator $\mathcal{T}_\pi$ acting on the space L^2 which transforms the wave function in the same way as the "physical" exchange of electrons according to the permutation π would. Analoguously we need to formalize how a rotation of the physical system should translate to the quantum mechanical model.

The definition of such operators is rather straightforward. We realize a rotation of a vector in $\mathbb{R}^3$ by a unitary matrix transform with determinant 1, that is $R \in SO(3)$. To implement the rotation for the state function, we apply the rotation matrix R to each of the position variables $x_1, \cdots, x_N \in \mathbb{R}^3$. More explicitly, we set

$$\mathcal{T}_R \psi(x_1, \cdots, x_N) := \psi(Rx_1, \cdots, Rx_N). \tag{G.5}$$

By defining the composition $\mathcal{T}_R \mathcal{T}_S := \mathcal{T}_{SR}$ we ensure that the collection of these operators form a group, which is isomorphic to $SO(3)$, the group of rotations in $\mathbb{R}^3$.

We want to proceed similarly for the permutation group S_N. Let $\pi \in S_N$ be such a permutation. We define the operator

$$\mathcal{T}_\pi \psi(x_1, \cdots, x_N) := \psi(x_{\pi^{-1}(1)}, \cdots, x_{\pi^{-1}(N)}) \tag{G.6}$$

and the composition $\mathcal{T}_\pi \mathcal{T}_\nu := \mathcal{T}_{\pi \circ \nu}$.

"The symmetry of a body is described by giving the set of all those transformations which preserve the distance between all pairs of points of the body and bring the body into coicidence with itself. Any such transformation is called a *symmetry transformation*. It is clear that this set forms a group, the *symmetry group* of the body." (see [18, p. 32])

This is a very general definition and it applies not only to physical objects, but also abstract objects such as geometric shapes, metric spaces and even operators in a canonical way. To determine the symmetry group of the Hamiltonian defined in (G.4) note that the Laplacian in three dimensions is invariant under rotations. Permutation invariance of $-\Delta$ is obvious, since the Laplacian in $\mathbb{R}^{3N}$ is a sum of

Laplacians in $\mathbb{R}^3$. The Coulomb potential depends only on relative distances which are left unchanged by permutations and rotations. So as desired, the symmetry group of the Hamiltonian defined in (G.4) is the product group $SO(3) \times S_N$. Furthermore, elements of the group of rotations $SO(3)$ and the permutation group S_N commute.

This symmetric property of the Hamiltonian implies that for an eigenfunction $\phi \in L^2(\mathbb{R}^{3N})$ of H corresponding to the eigenvalue E and a rotation $R \in SO(3)$

$$H(\mathcal{T}_R\phi) = \mathcal{T}_R H\phi = \mathcal{T}_R E\phi = E(T_R\phi). \tag{G.7}$$

So for an eigenfunction ϕ of H and any $R \in SO(3)$, the functions $\mathcal{T}_R\phi$ are eigenfunctions of H corresponding to the same energy E. The same is true if we replace the rotation R in (G.7) by a permutation $\pi \in S_N$, or even a combination of the two transformations.

In general, the functions $\mathcal{T}_g\phi$ are linearly independent for different $g \in SO(3) \times S_N$. Physically speaking, the states resulting one from another by the operations of the symmetry group of H are not fundamentally different states. Hence linear dependence does not do a good job at identifying this equivalence.

In the case of quantum numbers, this matter is clearer. Except for the so-called principal quantum number, the quantum numbers are directly linked to the representations of the to aforementioned groups.

Or as Weyl [46] writes: "Two groups, the group of rotations in 3-dimensional space and the permutation group, play here the principal rôle, for the laws governing the possible electron configurations grouped about the stationary nucleus of an atom or ion are spherically symmetric with respect to the nucleus, and since the various electrons of which the atom or ion is composed are identical, these possible configurations are invariant under a permutation of the individual electrons. The investigation of groups first becomes a connected and complete theory on *the theory of the representation of groups by linear transformations*, and is exactly this mathematically most important part which is necessary for an adequate description of the quantum mechanical relations."

G.3. **Representation theory.** In this chapter we will discuss a very narrow field of representation theory. We will focus purely on the permutation group and representations on the space of $L^2(\mathbb{R}^{3N})$ functions. Compared to the representations of the permutation group, the representation theory for the rotation group is rather straightforward.

In mathematical physics, when dealing with N fermions, we usually consider the space $\bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ which is built by the antisymmetric tensor product of copies of the space $L^2(\mathbb{R}^3, \mathbb{C}^2)$. This means that for N electrons we are dealing with 2^N possible configurations of spin. Such an array of information complicates the notation and carries information which is basically not relevant for the investigation of a Hamiltonian of the form (G.4). The spin of the electrons affects the energy only indirectly (if no magnetic field is present). Neither kinetic energy nor the Coulomb energy depend on the spin of an electron, so we are tempted to neglect the spin of an electron. However experimental data suggests that two electrons spinning in the same direction never occupy the same orbit, whereas two electrons with opposite spin can. The latter spin configuration would thus allow a lower energy, whereas in the first case one of the electrons would need to enter a higher orbit resulting in higher total energy.

So if we want to "forget" about the spin of the electrons we need to accomodate for this phenomenon in another way. Requiring full antisymmetry of the wave function, restricting the search for solutions to the Schrödinger equation to the space $\bigwedge_1^N L^2(\mathbb{R}^3)$ would only allow solutions which correspond to states where all

electrons spin in the same direction, thus overestimate the energy. Allowing any function in $L^2(\mathbb{R}^{3N})$ would predict too low energies, since an arbitrary number of electrons could occupy a low orbit. The sweet spot is somewhere inbetween. Recall that according to Sigalov Zhislin [50] all eigenvalues of the atomic Hamiltonian for 4 electrons corresponding to realizable symmetries are embedded in the continuous spectrum if we neglect electron-electron interaction.

In the case where we neglect the Pauli principle and search for solutions of the spin independent Schrödinger equation on $L^2(\mathbb{R}^{3N})$, the predicted ground states would suggest that atoms would become smaller with increasing atomic number and volume would not increase proportionally to the number of particles. The Pauli exclusion principle is a necessary condition for stability of matter of the second kind, see [25, Part IV]. Consider an operator of the form (G.4), extended canonically to the space $\bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$. If a function $\psi \in \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ is an eigenfunction corresponding to the eigenvalue λ, then component wise for every spin configuration $s \in \{-\frac{1}{2}, \frac{1}{2}\}^N$ we get

$$(H - \lambda)\psi(x, s) = 0. \tag{G.8}$$

Meaning that an eigenfunction $\psi \in \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ gives rise to a family of eigenfunctions in $L^2(\mathbb{R}^{3N})$.

Let $\psi \in \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ be given in the form of a 2^N-dimensional vector valued function

$$\psi(x, s) = \big(\psi_k(x)\big)_{k=1}^{2^N}. \tag{G.9}$$

Define the family of maps $\mathcal{L}_k : \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2) \to L^2(\mathbb{R}^{3N})$ with

$$\mathcal{L}_k\big(\psi(x, s)\big) := \psi_k(x). \tag{G.10}$$

The maps $\mathcal{L}_k$ are not surjective. In fact, the vector space spanned by the ranges

$$\operatorname{span}\{\operatorname{Ran}(\mathcal{L}_k), k = 1 \cdots 2^N\}$$

of this map is a strict subspace of functions in $L^2(\mathbb{R}^{3N})$ with certain symmetric properties, that we will specify later.

Note that for any function $\psi \in \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ which is a solution to some wave equation in coordinate variables, we can find a pair of functions $f : \{-\frac{1}{2}, \frac{1}{2}\}^N \to \mathbb{C}$ and $\phi \in L^2(\mathbb{R}^{3N})$ with

$$\psi(x_1, s_1, \cdots, x_N, s_N) = f(s_1, \cdots, s_N)\phi(x_1, \cdots, x_N). \tag{G.11}$$

Another way of asking the question: "Which state is physically relevant?" is: "From which class of functions $\phi \in L^2(\mathbb{R}^{3N})$ can we construct a function $\psi \in \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ by choosing a suitable function f depending on the spin variables alone".

At first glance it looks as if we are restricted to completely antisymmetric functions in $L^2(\mathbb{R}^{3N})$. But note that for a symmetric function $\phi(x_1, x_2) \in L^2(\mathbb{R}^6)$, we can choose the antisymmetric spin function $f(s_1, s_2) := s_1 - s_2$ and get that $f\phi \in L^2(\mathbb{R}^3, \mathbb{C}^2) \wedge L^2(\mathbb{R}^3, \mathbb{C}^2)$.

Another example with three electrons illustrates that there are functions in $L^2(\mathbb{R}^{3N})$ from which we can not construct a function in the space $\bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$.

Example G.1. *The permutation group S_3 is generated by the three transpositions* $(12), (23)$ *and* (13). *We claim that a function $\phi \in L^2(\mathbb{R}^9)$ which is symmetric under these three transpositions is not in the image of $\mathcal{L}_k$ for any k, where $\mathcal{L}_k$ is defined in* (G.10). *In other words, there exists no function $f : \{-\frac{1}{2}, \frac{1}{2}\}^3 \to \mathbb{C}$ such that $\phi f \in \bigwedge_1^3 L^2(R^3, \mathbb{C}^2)$. Since ϕ is symmetric, the antisymmetry condition on the*

product state $f\phi$ imposes restrictions to the function f. We have

$$\begin{aligned}\mathcal{T}_{(12)}\big(f(s_1,s_2,s_3)\phi(x_1,x_2,x_3)\big) &= f(s_2,s_1,s_3)\phi(x_2,x_1,x_3)\\ &= f(s_2,s_1,s_3)\phi(x_1,x_2,x_3)\\ &\overset{!}{=} -f(s_1,s_2,s_3)\phi(x_1,x_2,x_3),\end{aligned} \tag{G.12}$$

where the last equality sign "$\overset{!}{=}$" should be read as "is required to be equal to". This implies

$$f(s_1,s_2,s_3) = -f(s_2,s_1,s_3). \tag{G.13}$$

Analogously we get

$$f(s_1,s_2,s_3) \overset{!}{=} -f(s_1,s_3,s_2) \tag{G.14}$$

and

$$f(s_1,s_2,s_3) \overset{!}{=} -f(s_3,s_2,s_1). \tag{G.15}$$

Since the spin is a binary variable, at least two of the spin variables have the same value and thus (G.13)-(G.15) *imply $f \equiv 0$.*

Now that we have established a heuristic view on the matter, let us introduce some formal definitions and notation. The way in which a function $\phi \in L^2(\mathbb{R}^{3N})$ transforms under certain symmetry operation $\mathcal{T}_\pi$ for $\pi \in S_N$ is given by the collection of functions

$$K^\phi := \{\mathcal{T}_\pi\phi | \pi \in S_N\}. \tag{G.16}$$

We will fix some order of these $N!$ elements, where the first element corresponds to the identity permutation such that

$$K^\phi = \{\phi_1, \phi_2, \cdots, \phi_{N!}\} \tag{G.17}$$

with $\phi_1 = \phi$. We will use this ordered set to span a vector space of $N!$-tuples $(a_1, \cdots, a_{N!})$ corresponding to the linear combination $\sum_{i=1}^{N!} a_i\phi_i$. By the group structure of S_N, the space $V^\phi := \mathrm{span}(K^\phi)$ is an invariant subspace of the operators $\mathcal{T}_\pi$ for any $\pi \in S_N$. In fact, by using the ordered set K^ϕ one can express any permutation acting on an element in V^ϕ as a matrix transformation with respect to the ordered set K^ϕ. In this notation, each permutation is represented by a matrix with only ones and zeroes, in each row and column there is exactly one nonzero entry. The collection of matrices $\{D^{\alpha_R}(\pi) | \pi \in S_N\}$ is called the *regular representation* α_R and we already know this type of matrices as permutation matrices (of size $N!$). But this is just the starting point. It might very well happen, that the set K^ϕ is not linearly independent. In fact, the dimension of V^ϕ depends on the symmetric properties of ϕ. It is easy to see that if ϕ is fully symmetric (or fully antisymmetric) with respect to permutations, then the vector space spanned by $\mathcal{T}_\pi\phi$ has only one dimension. Let us choose a basis of V^ϕ, we get a new representation α_B if we express the matrices with respect to the basis B^ϕ. Obviously this process decreases the size of the transformation matrices. The matrices of the regular representation have $N!$ rows and columns.

To simplify these matrices further, we can try to block-diagonalize the matrices. We search for subspaces of V^ϕ which are invariant under all permutations. By changing the basis, we can then simultaneously block-diagonalize the $N!$ matrices $D^{\alpha_B}(\pi)$ (this is possible since the subspace is required to be $\mathcal{T}_\pi$-invariant for *all* permutations $\pi \in S_N$) and get

$$T_1^{-1} D^{\alpha_B}(\pi) T_1 = \begin{pmatrix} D^{\alpha_u}(\pi) & 0 \\ 0 & D^{\alpha_d}(\pi) \end{pmatrix} \tag{G.18}$$

where the blocks $D^{\alpha_u}(\pi)$ have same size for each $\pi \in S_N$. This process is called *reducing* of a representation and we repeat this process for the block-matrices $D^{\alpha_u}(\pi)$

and $D^{\alpha_d}(\pi)$ until there are no further invariant subspaces. For any element $\pi \in S_N$, the matrix $D^{\alpha_B}(\pi)$ can be brought to the form

$$T^{-1}D^{\alpha_B}(\pi)T = \begin{pmatrix} D^{\alpha_1}(\pi) & 0 & \dots & 0 \\ 0 & D^{\alpha_2}(\pi) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & D^{\alpha_k}(\pi) \end{pmatrix} \tag{G.19}$$

where each collection of $N!$ matrices $D^{\alpha_i}(\pi)$ is a representation α_i on some invariant subspace of V^ϕ. Note that *simultaneous* block-diagonalization implies that the transformation matrix T must be the same for all $N!$ permutations.

For an arbitrary matrix representation α of the group S_N, we denote by n_α the size/dimension of the matrix $D^\alpha(\pi)$. We call n_α the *degree* or *dimension of the representation* α. If the invariant subspace corresponding to any representation α contains no proper invariant subspace, then the representation α is called an *irreducible representation*. In particular, the representations $\alpha_1, \cdots, \alpha_k$ are irreducible representations. Note that a representation is always a collection of (in this case $N!$) matrices, one for each element of the group S_N.

We call two representations α and α' *equivalent*, if there exists a matrix T such that for all $\pi \in S_N$ we have $T^{-1}D^\alpha(\pi)T = D^{\alpha'}(\pi)$. As usual we can choose any element of the equivalence class $\{\alpha\}$ to label this class of representations. Now that we have potentially the same name for representations that act on different spaces we need to clarify that the invariant subspace corresponding to α refers, if not stated explicitly, to the collection of invariant subspaces corresponding to representations in the equivalence class of α. The invariant subspace of V^ϕ corresponding to irreducible representation α_i is a strict subspace of the invariant subspace of L^2 corresponding to the irreducible representation α_i.

In the case of L^2 the invariant subspace corresponding to the regular representation α_R is all of L^2. This abuse of notation is justifiable in the sense that we use this label only to describe the symmetric property. In our construction of representations, they are very closely linked to the vector spaces that they act on. This is not an necessary framework for the representations but a helpful framework for us. Since we are dealing with the finite groups S_N, each irreducible representation α is finite-dimensional and we can always find an equivalent irreducible representation $\alpha_u \in \{\alpha\}$ for which all matrices $D^{\alpha_u}(\pi)$ are unitary matrices. Such a representation is called a *unitary representation*.

The representations obtained in the way described above are a priori different for each function ϕ that we use as a starting point. However, it is clear that if we start with any fully symmetric function, the permutations will span a one-dimensional vector space and the matrix representation on this space is the identity in one dimension, called the *trivial representation* or *identity representation*. Thus the construction will yield an equivalent representation if we start with any symmetric function. The invariant subspace (of L^2) corresponding to the identity representation thus contains all fully symmetric functions of L^2.

By starting with a fully antisymmetric function we will obtain a one-dimensional vector space, where the matrix representation is given by the sign of the permutation $\mathrm{sgn}(\pi)$, called the *alternating representation* or *sign representation*.

An irreducible representation describes how a permutation acts on a specific space of functions, thus it describes nothing less than the "type of symmetry" that is shared among the functions in the corresponding space. For the case $N = 2$ it is clear that each function can be written as the sum of a fully symmetric function

and a fully antisymmetric function, thus the method discussed above will yield at most two nonequivalent irreducible representations.

Intuitively, if the number of particles increases, so should the complexity of possible symmetric properties, or the number of "symmetry types" needed to accurately describe these symmetric properties.

The goal is to express any function ϕ as a sum of functions ϕ^{α_i} which transform according to the irreducible representation α_i. From this sum we can then deduce the symmetric properties of the initial function ϕ by looking at the irreducible representations that are needed to express ϕ. It is fundamental, that a *finite* group has *finitely* many nonequivalent irreducible representations. This result is discussed in the next section. To avoid getting lost in vague explanations and fishy definitions it is necessary to dive a bit more into representation theory. The concepts are not complicated, but require some notation and explanation.

G.4. **Projection onto irreducible representations.** An irreducible representation is an abstract object, that can have many faces. A more algebraic view on irreducible representations will be discussed in Chapter G.5. This chapter however is dedicated to an alternative interpretation, with the upside that mathematical physicists are more familiar with the objects involved and it can help us to understand the matter. The downside is that it does not reveal a lot of the algebraic structures beneath it, and thus the vast theory of representations and groups remains somewhat unused.

For many of the applications in mathematical physics we can think of irreducible representations as their corresponding invariant subspaces of L^2 which contain functions with certain symmetric properties and that decompose the space L^2 into orthogonal subspaces. We will construct projection operators that are based on a linear combination of the operators $\mathcal{T}_\pi$. The coefficients in this linear combination depend on the matrices in the representation, in particular on the traces of the matrices. The *character* χ^α_π of the group element π in the representation α is defined as the trace of the matrix $D^{(\alpha)}(\pi)$. Let n_α denote the dimension of α. The projection operator onto the invariant subspace corresponding to the irreducible representation α is given by

$$P^\alpha = \frac{n_\alpha}{N!} \sum_{\pi \in S_N} \chi^\alpha_{\pi^{-1}} \mathcal{T}_\pi. \tag{G.20}$$

For a more general group G, the prefactor $N! = \sharp(S_N)$ is replaced by $\sharp(G)$, the number of elements in the group G. Note that if we consider unitary representations $\chi^\alpha_{\pi^{-1}} = \overline{\chi}^\alpha_\pi$ where the overbar denotes the complex conjugate.

For a set of nonequivalent irreducible representations $\{\alpha_1, \cdots, \alpha_k\}$ of S_N we have

$$P^{\alpha_p} P^{\alpha_q} = \delta_{pq} P^{\alpha_q} \text{ for } p, q \in \{1, \cdots, k\} \tag{G.21}$$

and if $\{\alpha_1, \cdots, \alpha_k\}$ includes every irreducible representation of the group S_N, then

$$\sum_{i=1}^{k} P^{\alpha_k} = 1. \tag{G.22}$$

Equation (G.21) is due to the orthogonality relations for the representation matrices

$$\sum_{\pi \in S_N} D^{\alpha_p}_{ji}(\pi) \overline{D^{\alpha_q}_{lm}}(\pi) = \frac{N!}{\chi^{\alpha_p}_{id}} \delta_{jl} \delta_{im} \delta_{pq}. \tag{G.23}$$

By clever multiplication of matrix components $D^\alpha_{ij}(\pi)$ and summations, taking advantage of certain group properties, relation(G.23) yields all the desired properties for the projection operators defined in (G.20), see for example [18, p. 111ff]. Note that this is an orthogonal decomposition with respect to permutation invariant

subspaces. Since the Hamiltonian H is also permutation invariant, the eigenspaces of the Hamiltonian are direct sums of invariant subspaces corresponding to the irreducible representations $\{\alpha_1, \ldots, \alpha_k\}$. Thus we gain complete information about the spectrum of H if we study the operator on each of these invariant subspaces separately.

We will now proceed to discuss which of the irreducible representations in the set $\{\alpha_1, \cdots, \alpha_k\}$ correspond to physically relevant states. This is done in the next section by introducing Young diagrams.

G.5. **Young diagrams.** In this subsection we try to give a geometric meaning to the abstract notions of irreducible representations. We categorize every possible irreducible representation with the help of Young diagrams and describe how to evaluate these objects in order to attribute a symmetric property, or symmetry type to each irreducible representation.

We will start this subsections by one of the fundamental results of representation theory, [18, §3-17].

Theorem G.2. *The number of nonequivalent irreducible representations of a (finite) group is equal to the number of classes in the group.*

For the permutation group, the cycle type of a permutation π determines the conjugacy class to which it belongs. Furthermore, the number of distinct cycle types in S_N is equal to the number of partitions of N. In fact there is a one-to-one relation between partitions of N and the irreducible representations of S_N. This relation is best illustrated using Young diagrams.

A Young diagram is a graphical expression of a partition $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_k)$ with $\lambda_1 \geq \cdots \geq \lambda_k$, using an arrangement of empty boxes, where one puts λ_j boxes in the jth row. For example, the Young diagram associated to the partition $(5, 4, 2, 2)$ is

$$\begin{array}{|c|c|c|c|c|}\hline & & & & \\ \hline & & & \\ \cline{1-4} & \\ \cline{1-2} & \\ \cline{1-2} \end{array} \tag{G.24}$$

A Young diagram corresponding to the partition λ of N in which we label the boxes with numbers from 1 to N is called a *Young tableau*. If we label the boxes in a strictly increasing order with respect to rows and columns the tableau is called a *standard Young tableau* Λ. For example

$$\begin{array}{|c|c|c|c|c|}\hline 1 & 2 & 5 & 10 & 11 \\ \hline 3 & 4 & 6 & 12 \\ \cline{1-4} 7 & 9 \\ \cline{1-2} 8 & 13 \\ \cline{1-2} \end{array} \tag{G.25}$$

is a standard Young tableau, whereas

$$\begin{array}{|c|c|c|c|c|}\hline 2 & 1 & 5 & 10 & 11 \\ \hline 3 & 4 & 6 & 12 \\ \cline{1-4} 7 & 9 \\ \cline{1-2} 8 & 13 \\ \cline{1-2} \end{array} \tag{G.26}$$

is *not* a standard Young tableau since the labelling decreases in the first row from 2 to 1.

In fact each *standard* Young tableau Λ with shape λ will give us a technique to construct one of the basis vectors of the irreducible representation λ. For the tableau Λ we define the *row subgroup* R_Λ of permutations that permute only elements within the same row and the *column subgroup* C_Λ which permutes only the elements within the same column. We will then symmetrize with respect to permutations in the

row subgroup R_Λ and antisymmetrize with respect to permutations in the column subgroup C_Λ. The corresponding operation is given by an element of $\mathbb{C}[S_N]$, the algebra over the group S_N, which contains linear combinations of group elements $\pi \in S_N$. Specifically, for a fixed Young tableau Λ we define the Λ-symmetrizer as

$$p_\Lambda := \sum_{\pi \in R_\Lambda} \pi. \tag{G.27}$$

We can realize the element $p_\Lambda \in \mathbb{C}[S_N]$ as an operator on L^2 by representing π via the linear operators $\mathcal{T}_\pi$. The operator that symmetrizes an L^2-function with respect to the group S_N will also be called Λ-symmetrizer and is given by

$$P_\Lambda := \sum_{\pi \in R_\Lambda} \mathcal{T}_\pi \tag{G.28}$$

acting on L^2. Similarly we define the Λ-antisymmetrizer

$$q_\Lambda = \sum_{\pi \in C_\Lambda} \operatorname{sgn}(\pi)\pi \tag{G.29}$$

and the corresponding operator

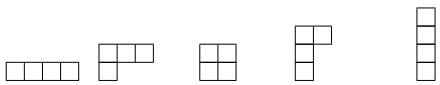
$$Q_\Lambda = \sum_{\pi \in C_\Lambda} \operatorname{sgn}(\pi)\mathcal{T}_\pi. \tag{G.30}$$

The element $y = pq$ is a generator of a left ideal in the algebra of permutations. If the Young tableau is *standard*, then the left ideal is *minimal*, meaning that it contains no proper sub left ideal, which means that the corresponding representation is irreducible. The *Young operator* corresponding to the tableau Λ is given by the product of the Λ-symmetrizer and the Λ-antisymmetrizer

$$Y_\Lambda = Q_\Lambda P_\Lambda. \tag{G.31}$$

The image of this operator is a subspace of L^2 containing functions with symmetric properties corresponding to the Young tableau Λ. That is, antisymmetric with respect to the column subgroup C_Λ and symmetric with respect to the row subgroup R_Λ. The symmetry classes (types) categorized in this way have the property that "the symmetry of the tensors (functions) belonging to it cannot be further increased by the addition of further symmetry conditions - such an additional condition either reproduces all the tensors of the class or reduces them all to 0." [46, p. 359] Which is excatly the categorization that we want.

We will illustrate this process with the help of the group S_4 and the partition of 4. The Young diagrams corresponding to the partitions of 4 are



$$(4) \quad (3,1) \quad (2,2) \quad (2,1,1) \quad (1,1,1,1) \tag{G.32}$$

The diagram (4) can only be filled in one way, that is $\Lambda_1^{(4)} = \boxed{1}\boxed{2}\boxed{3}\boxed{4}$. So we symmetrize with respect to the row subgroup, which in this case is the group S_4. The $\Lambda_1^{(4)}$-symmetrizer is

$$P_{\Lambda_1^{(4)}} = \sum_{\pi \in S_N} \mathcal{T}_\pi. \tag{G.33}$$

The $\Lambda_1^{(4)}$-antisymmetrizer is the identity operator and thus $Y_{\Lambda_1^{(4)}} = P_{\Lambda_1^{(4)}}$. Obviously a function in the image of $Y_{\Lambda_1^{(4)}}$ is completely symmetric as desired. For the

alternating representation with diagram $(1,1,1,1) = (1)^4$, there is only one standard tableau $\Lambda_1^{(1)^4}$. Thus the row subgroup only contains the identity permutation ($R_{\Lambda_1^{(1)^4}} = \{\mathrm{Id}\}$) and the $\Lambda_1^{(1)^4}$-antisymmetrizer is

$$q_{\Lambda_1^{(1)^4}} = \sum_{\pi \in S_4} \mathrm{sgn}(\pi)\pi \tag{G.34}$$

with corresponding operator

$$Q_{\Lambda_1^{(1)^4}} := \sum_{\pi \in S_4} \mathrm{sgn}(\pi)\mathcal{T}_\pi. \tag{G.35}$$

For the shape $(2,1,1) = (2)(1)^2$ there are three standard Young tableaux

$$\Lambda_1^{(2)(1)^2} := \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \qquad \Lambda_2^{(2)(1)^2} := \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 \\ \cline{1-1} 4 \\ \cline{1-1}\end{array} \qquad \Lambda_3^{(2)(1)^2} := \begin{array}{|c|c|}\hline 1 & 4 \\ \hline 2 \\ \cline{1-1} 3 \\ \cline{1-1}\end{array}. \tag{G.36}$$

We will express permutations in cycle form, i.e. (134) refers to the permutation with $1 \to 3$, $3 \to 4$ and $4 \to 1$. The tableau $\Lambda_1^{(2)(1)^2}$ yields the operators

$$P_{\Lambda_1^{(2)(1)^2}} = \mathrm{Id} + \mathcal{T}_{(12)} \tag{G.37}$$

$$Q_{\Lambda_1^{(2)(1)^2}} = \mathrm{Id} - \mathcal{T}_{(13)} - \mathcal{T}_{(14)} - \mathcal{T}_{(34)} + \mathcal{T}_{(134)} + \mathcal{T}_{(143)} \tag{G.38}$$

and the Young operator is

$$Y_{\Lambda_1^{(2)(1)^2}} = \left(\mathrm{Id} + \mathcal{T}_{(12)}\right)\left(\mathrm{Id} - \mathcal{T}_{(13)} - \mathcal{T}_{(14)} - \mathcal{T}_{(34)} + \mathcal{T}_{(134)} + \mathcal{T}_{(143)}\right). \tag{G.39}$$

The construction of the Young operators $Y_{\Lambda_2^{(2)(1)^2}}$ and $Y_{\Lambda_3^{(2)(1)^2}}$ are very similar. The diagram $(2,2)$ will yield antisymmetrizers of a slightly different form. The diagram $\begin{array}{|c|c|}\hline & \\ \hline & \\ \hline\end{array}$ can be filled in two standard ways:

$$\Lambda_1^{(2)^2} = \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 3 & 4 \\ \hline\end{array} \qquad \Lambda_2^{(2)^2} = \begin{array}{|c|c|}\hline 1 & 3 \\ \hline 2 & 4 \\ \hline\end{array} \tag{G.40}$$

and we get

$$P_{\Lambda_1^{(2)^2}} = \mathrm{Id} + \mathcal{T}_{(12)} + \mathcal{T}_{(34)} + \mathcal{T}_{(12)(34)} = (\mathrm{Id} + \mathcal{T}_{(12)})(\mathrm{Id} + \mathcal{T}_{(34)}). \tag{G.41}$$

From the factorized form on the right hand side of (G.41) we can better see the group structure of the row subgroup: $R_{\Lambda_1^{(2)^2}} = S_{\{1,2\}} \times S_{\{3,4\}}$. This direct product group is isomorphic to $S_2 \times S_2$. Similarly

$$Q_{\Lambda_1^{(2)^2}} = \mathrm{Id} - \mathcal{T}_{(13)} - \mathcal{T}_{(24)} + \mathcal{T}_{(13)(24)} = (\mathrm{Id} - \mathcal{T}_{(13)})(\mathrm{Id} - \mathcal{T}_{(24)}) \tag{G.42}$$

and $C_{\Lambda_1^{(2)^2}} = S_{\{1,3\}} \times S_{\{2,4\}}$. The Young operator is thus

$$Y_{\Lambda_1^{(2)^2}} = (\mathrm{Id} - \mathcal{T}_{(13)})(\mathrm{Id} - \mathcal{T}_{(24)})(\mathrm{Id} + \mathcal{T}_{(12)})(\mathrm{Id} + \mathcal{T}_{(34)}). \tag{G.43}$$

For the second tableau we get the operators

$$P_{\Lambda_1^{(2)^2}} = (\mathrm{Id} + \mathcal{T}_{(13)})(\mathrm{Id} + \mathcal{T}_{(24)}) \tag{G.44}$$

$$Q_{\Lambda_1^{(2)^2}} = (\mathrm{Id} + \mathcal{T}_{(13)})(\mathrm{Id} + \mathcal{T}_{(24)}) \tag{G.45}$$

and

$$Y_{\Lambda_2^{(2)^2}} = (\mathrm{Id} + \mathcal{T}_{(13)})(\mathrm{Id} + \mathcal{T}_{(24)})(\mathrm{Id} - \mathcal{T}_{(12)})(\mathrm{Id} - \mathcal{T}_{(34)}). \tag{G.46}$$

This construction shows naturally the symmetry type that is associated with a certain Young diagram.

Remark G.3. *Note that the matrix representation on the bases obtained in this way are not the unitary matrices that we discussed on the previous chapter.*

Heuristically speaking, the construction via Young tableaux gives rise to irreducible representations where the basis vectors are in the spotlight, whereas the construction discussed in the previous chapter via matrix reduction puts emphasise on the matrices of the representation. Both methods have their legitimization and there is no strictly superior choice.

We now try to build the bridge back to the physical application of the insight we have gained. As we have seen in the example of three electrons (example G.1), a function that is physically relevant can be symmetric with respect to at most two particles. Permutation symmetry of $\phi \in L^2(\mathbb{R}^{3N})$ with respect to a triplet of particles would, by the same argument as in example G.1, immediatly yield that the function can not arise from a projection of $\psi \in \bigwedge_1^N L^2(\mathbb{R}^3, \mathbb{C}^2)$ onto $L^2(\mathbb{R}^{3N})$ via the maps $\mathcal{L}_k$ defined in (G.10).

This means that if we study the energy of the Hamiltonian H we only consider those functions that belong to spaces categorized by a Young diagram with at most two columns.

Let $\mathcal{A}$ be the collection of nonequivalent irreducible representations that correspond to a Young diagram with at most two columns, then the subspace of $L^2(\mathbb{R}^{3N})$ which contains all states that respect the Pauli exclusion principle is given by

$$\mathcal{H}_{\text{Fermi}} := \bigoplus_{\alpha \in \mathcal{A}} P^\alpha L^2(\mathbb{R}^{3N}). \tag{G.47}$$

Appendix H. Prospect: generalization and application of the method

We will discuss some mathematical and physical problems that are related to the results of Theorem 9.2 and Theorem 9.4. This includes for one part the discussion of the conditions that we require in the proof of the main result. We will also examine the applicability of the methodology to other models.

Let us briefly recall the assumptions of Theorem 9.2 without restating the respective definitions:
Condition 1) For all $\beta \in \mathcal{D}_N^2 \setminus \mathcal{D}^{at}$

$$\mu_\beta^\alpha > \mu^\alpha.$$

Condition 2) For each $\beta \in \mathcal{D}^{at}$ and each irreducible representation α_β^* of the group S_β with $\alpha_\beta^* \prec \alpha$ such that $P^{\alpha_\beta^*}\tilde{\mathcal{W}}_\beta^\alpha \neq \emptyset$,

$$\dim(P^{\alpha_\beta^*}\tilde{\mathcal{W}}_\beta^\alpha) = \dim \alpha_\beta^*.$$

Condition 1) implies that the lowest energy for a non-interacting system occurs when the electrons are allocated neutrally. Note that in order to obtain an upper bound for the van der Waals energy we do not need this condition. There are no mathematically rigorous results that we know of that confirm or contradict this assumption. But we can compare the assumption to physical data and observations.

While there exist many examples in nature where ionic bounds are formed, this observation does not contradict the physical validity of the assumption. These ionizations are caused by the interaction between the two atoms at relatively small distances. We need to compare the combined energy of an *isolated* Na^+ ion and an *isolated* Cl^- ion to the combined energy of the respective *isolated* atoms Na and Cl. This comparison can be done by looking at the ionization energies and electron affinities of the two elements. The electron affinity of chloride (Cl) is measured at around $348kJ/mol$, whereas the ionization energy for sodium (Na) is approximately $496kJ/mol$, see e.g. [23]. To find possible cases that contradict the condition, we have to compare the ionization energy to the electron affinity for different elements. If there exists a tuple of elements where the ionization energy for one element is bigger than the electron affinity of the other, this combination of

elements would not satisfy condition 1). As it turns out chloride is the element with the highest ionization energy at around $348kJ/mol$. The element with the lowest electron affinity is caesium (Cs) with a value of approximately $375kJ/mol$ which is still above the ionization energy of chloride. So for any collection of interacting atoms this seems to be a reasonable assumption.

Since our investigation focuses on a pseudo-relativistic model, the interesting cases are larger atoms or even molecules interacting via van der Waals-London dispersion. In the setting of molecules we have not done such a comparison of experimental data and it is probably not reasonable to do so.

Condition 2) is used to prove that the intercluster interaction of two atomic ground states is zero. We use it to obtain spherical symmetry of the one-electron densities of the atomic ground state around the respective nucleus. By Newton's theorem spherically distributed charges can be viewed as point charges located at the center of the distribution for all points outside the support of the distribution.

Mathematically the spherical symmetry is true for the ground states of non-relativistic hydrogen. For other atoms there exist, at least to our knowledge, no rigorous results on the rotational symmetry of the ground states.

In the orbital model that we have discussed briefly in Section G.1 and that is still used in quantum chemistry the one-electron densities are modelled by orbitals. Mathematically these orbitals are described by the spherical harmonics Y_ℓ^m, in particular, the s orbital is described by the spherical harmonic of degree $\ell = 0$, the p orbital by the spherical harmonics of degree $\ell = 1$ and so on. In this model, the condition would be fulfilled for atoms up to atomic number $Z = 4$. For atoms with bigger atomic number, only some of the one-electron densities would be spherically symmetric, i.e. for those electrons that are in an s orbital.

However, as the name suggests these spherical harmonics Y_ℓ^m are harmonic functions, in particular the mean value property is true for Y_ℓ^m for any $\ell \in \mathbb{N}$ and $|m| \leq \ell$. This mean value property of harmonic functions could possibly be used to obtain a similar "Newtonesque theorem" that would allow us to view the electrons as point charges located at the respective nucleus. Or at least to gain some information about interaction of multipoles for states with one-electron densities described by spherical harmonics of finite degree. We have so far not investigated this topic in a rigorous way since it has only recently come to our attention.

In the next section we will evaluate another possible generalization of the result. We aim to use the same underlying idea to obtain a similar result for another quantum Hamiltonian.

Appendix I. The Brown-Ravenhall operator

In this chapter we discuss the possibility to apply the developped method to another pseudo-relativistic model, the Brown-Ravenhall operator. The main difficulty will again be the nonlocality of the operator.

I.1. **Introduction to the model.** The N-electron Brown-Ravenhall Hamiltonian for M nuclei of charge eZ_l at fixed positions X_l interacting with N electrons of charge $-e$ and position vector $x_j \in \mathbb{R}^3$ is given by

$$H^{BR} = \Lambda_+^N \Big(\sum_{j=1}^N \Big(D_j - \sum_{l=1}^M \frac{e^2 Z_l}{|x_j - X_l|} \Big) + \sum_{1 \leq i < j \leq N} \frac{e^2}{|x_i - x_j|} + \sum_{1 \leq k < l \leq M} \frac{e^2 Z_k Z_l}{|X_k - X_l|} \Big) \Lambda_+^N. \tag{I.1}$$

Here D_j is the Dirac operator given by

$$D_j = -i\boldsymbol{\alpha} \cdot \nabla_j + \beta \tag{I.2}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ and β are the four Dirac matrices,

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}, \quad k = 1, 2, 3, \tag{I.3}$$

with σ_k denoting the kth Pauli matrix

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{I.4}$$

and

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{I.5}$$

The operator H^{BR} defined in (I.2) is well-defined as the semibounded closed form on $\Lambda_+^N H^{1/2}(\mathbb{R}^{3N}, \mathbb{C}^{4N})$ for

$$e^2 Z_l < 2\left(\frac{\pi}{2} + \frac{2}{\pi}\right)^{-1} \qquad \text{for all } l = 1, \cdots, M. \tag{I.6}$$

The operator Λ_+^N is a projection on the positive spectral subspaces of the Dirac operator in each direction, namely

$$\Lambda_+^N = \bigoplus_{i=1}^{N} \Lambda_+^{(i)} \tag{I.7}$$

where $\Lambda_+^{(i)}$ the projection onto the positive spectral subspace of D_i. This projection is called a *Casimir projection*, in [5, (2.3.3)] the one electron projection is given by

$$\Lambda_+ = \frac{1}{2} I_4 + \mathcal{F}^* \frac{\boldsymbol{\alpha} \cdot \mathbf{p} + \beta}{2\sqrt{|\mathbf{p}|^2 + 1}} \mathcal{F}, \tag{I.8}$$

where $\mathcal{F}$ is the Fourier transform. More important for the use in this thesis will be the following integral representation for a function $f \in C_0^1(\mathbb{R}^3, \mathbb{C}^4)$.

$$\begin{aligned}(\Lambda_+ f)(x) = \frac{f(x)}{2} + \frac{1}{4\pi^2} \int_{\mathbb{R}^3} \left(\beta \frac{K_1(|x-y|)}{|x-y|} + \frac{i\boldsymbol{\alpha} \cdot (x-y)}{|x-y|^2} K_0(|x-y|) \right) f(y) dy \\ + \frac{i}{2\pi^2} PV \int_{\mathbb{R}^3} \frac{\boldsymbol{\alpha} \cdot (x-y)}{|x-y|^3} K_1(|x-y|) f(y) dy\end{aligned} \tag{I.9}$$

where the prefix PV indicates an integral in the *Cauchy principle value* sense, since due to the singularity it is a priori not well-defined in the Lebesgue sense. The functions K_0 and K_1 are modified Bessel functions of the second kind.

We will study the case $M = 2$ and unlike in the main result, we do not decompose the state space with respect to subspaces with certain symmetry types. For $D := X_2 - X_1$ we define

$$E_{|D|} := \inf \sigma(H^{BR}). \tag{I.10}$$

We adapt the notations for cluster decompositions from the Herbst case. Similarly, for a two cluster decomposition $\beta = (\mathcal{C}_1, \mathcal{C}_2) \in \mathcal{D}_N^2$ we define the intercluster interaction

$$I_\beta := \sum_{i \in \mathcal{C}_1} \frac{-e^2 Z_2}{|x_i - X_2|} + \sum_{j \in \mathcal{C}_2} \frac{-e^2 Z_1}{|x_j - X_1|} + \sum_{\substack{i \in \mathcal{C}_1 \\ j \in \mathcal{C}_2}} \frac{e^2}{|x_i - x_j|} + \frac{e^2 Z_1 Z_2}{|X_2 - X_1|} \tag{I.11}$$

and $I_\beta^{BR} := \Lambda_+^N I_\beta \Lambda_+^N$. The cluster Hamiltonian is defined as

$$H_\beta^{BR} := H^{BR} - I_\beta^{BR}. \tag{I.12}$$

We define the lowest energy for a given decomposition $\beta \in \mathcal{D}_N^2$ as

$$\mu_\beta := \inf \sigma(H_\beta^{BR}) \tag{I.13}$$

and

$$\mu := \min_{\beta \in \mathcal{D}_N^2} \mu_\beta. \tag{I.14}$$

Recall that we denote by $\mathcal{D}^{at} \subset \mathcal{D}_N^2$ the atomic decompositions, i.e. $\beta \in \mathcal{D}_N^2$ such that $\sharp \mathcal{C}_1(\beta) = Z_1$ and $\sharp \mathcal{C}_2(\beta) = Z_2$.

Theorem I.1 (The vdW interaction in the Brown-Ravenhall model)**.** *Let the following conditions hold:*
1) For all $\beta \in \mathcal{D}_N^2 \setminus \mathcal{D}^{at}$

$$\mu_\beta > \mu. \tag{I.15}$$

2) The ground state space of H_β^{BR} transforms according to the irreducible representation of weight $\ell = 0$.

Then there exists $c > 0$ such that

$$E_{|D|} - \mu = -\frac{c}{|D|^6} + \mathcal{O}(|D|^{-7}). \tag{I.16}$$

Remark I.2. *Condition 2) implies spherically symmetric one particle densities for the ground states of H_β^{BR} and is analogue to the second condition, if we do not decompose the state space with respect to different symmetry types.*

The modification to the localization estimate is the key point in proving the above theorem. We will give a rather detailed proof for the localization error estimate, while using the existing results and estimates from [32], whenever they fit the purpose.

The idea for the proof of the theorem is on a fundamental level very similar to the proof of Theorem 9.2, but we will deviate from the proof for the sake of briefness. We did not find a reason that would argue against using the more accurate method. In fact, the presented method is rather close to the original idea that lead to the refinement presented in the main part.

I.2. **Localization error estimate modified.** The aim of this discussion is to modify a localization error estimate established in [32], to yield exponentially small localization error for exponentially decaying wave functions.

We use mostly the notations and results of [32]. By slight modifications to the proofs in the spirit of the localization estimate for the Herbst operator, we obtain a slightly stronger result. With this we can use the decaying nature of the wave function far from the nucleus to estimate the intercluster interaction energy.

For the estimate of the localization error we will need the following result from [42]. We state it here, without proof, for convenience of the reader.

Theorem I.3. *Let $K : \mathbb{R}^n \to \mathbb{C}$ be measurable such that*

$$|K(x)| \leq B|x|^{-n}, \quad x \neq 0, \tag{I.17}$$

$$\int_{|x| \geq 2|y|} |K(x-y) - K(x)| dy \leq B, \quad 0 < |y|, \tag{I.18}$$

and

$$\int_{R_1 < |x| < R_2} K(x) dx = 0, \quad \text{for all} \quad 0 < R_1 < R_2 < \infty. \tag{I.19}$$

For an arbitrary $f \in L^p(\mathbb{R}^n)$, $1 < p < \infty$, let

$$T_\epsilon(f)(x) = \int_{|y| \geq \epsilon} f(x-y) K(y) dy, \quad \epsilon > 0. \tag{I.20}$$

Then

$$\|T_\epsilon(f)\|_p \leq A_p \|f\|_p \tag{I.21}$$

with A_p independent of f and ϵ. Furthermore the limit $\epsilon \to 0$ exists and the same bound is true.

We define the cutoff function $\chi \in C_0^2(\mathbb{R}^3, [0,1])$ such that

$$\chi(x) = \begin{cases} 1, \text{ for all } |x| \leq R \\ 0, \text{ for all } |x| > 2R \end{cases} \tag{I.22}$$

and smooth inbetween. Define the set $\Omega := \{x \in \mathbb{R}^3 \big| R - \delta R < |x| < 2R + \delta R\}$ and the corresponding indicator function $\mathbf{1}_\Omega$.

Lemma I.4. *Let $\chi \in C_0^2(\mathbb{R}^3, [0,1])$ be defined as above. Then for any function $f \in L^2(\mathbb{R}^3, \mathbb{C}^4)$*

$$\|[\chi, \Lambda_+]f\|_{H_1} \leq C\left((R^{-1} + R^{-2})\|\mathbf{1}_\Omega f\|_2 + e^{-\delta R}\|f\|_2\right). \tag{I.23}$$

Proof. Note that

$$\|[\chi, \Lambda_+]f\|_{H_1} = \|[\chi, \Lambda_+]f\|_2 + \|\nabla[\chi, \Lambda_+]f\|_2. \tag{I.24}$$

We bound both summands separately. For the first summand on the right hand side, we use the integral representation of $[\chi, \Lambda_+]$ to get

$$\begin{aligned} &[\chi, \Lambda_+]f(x) = \\ &\quad \frac{1}{4\pi^2}\int_{\mathbb{R}^3}\left(\beta\frac{K_1(|x-y|)}{|x-y|} + \frac{i\boldsymbol{\alpha}\cdot(x-y)}{|x-y|^2}K_0(|x-y|)\right)\big(\chi(x) - \chi(y)\big)f(y)dy \\ &\quad + \frac{i}{2\pi^2}PV\int_{\mathbb{R}^3}\frac{\boldsymbol{\alpha}\cdot(x-y)K_1(|x-y|)}{|x-y|^3}\big(\chi(x) - \chi(y)\big)f(y)dy. \end{aligned} \tag{I.25}$$

Define the function $K \in L^1(\mathbb{R}^3)$ with

$$\frac{K(x)}{|x|} := \beta\frac{K_1(|x|)}{|x|} + \frac{i\boldsymbol{\alpha}\cdot x}{|x|^2}K_0(|x|) + \frac{2i\boldsymbol{\alpha}\cdot x}{|x|^3}K_1(|x|). \tag{I.26}$$

By this definition from (I.25) and split $\mathbb{R}^3$ into two regions to get

$$\begin{aligned} [\chi, \Lambda_+]f(x) =& \frac{1}{4\pi^2}\int_{|x-y|\leq\delta R}\frac{K(x-y)}{|x-y|}\big(\chi(x) - \chi(y)\big)f(y)dy \\ &+ \frac{1}{4\pi^2}\int_{|x-y|\geq\delta R}\frac{K(x-y)}{|x-y|}\big(\chi(x) - \chi(y)\big)f(y)dy. \end{aligned} \tag{I.27}$$

To bound the absolute value of the first integral, note that in the region $|x-y| \leq \delta R$ we have

$$\chi(x) - \chi(y) = \mathbf{1}_\Omega(y)\big(\chi(x) - \chi(y)\big). \tag{I.28}$$

By the mean value theorem

$$|\chi(x) - \chi(y)| \leq \|\nabla\chi\|_\infty |x-y|. \tag{I.29}$$

For the first integral on the right hand side of (I.27) this yields

$$\left|\int_{|x-y|\leq\delta R}\frac{K(x-y)}{|x-y|}\big(\chi(x) - \chi(y)\big)f(y)dy\right| \leq \|\nabla\chi\|_\infty\int_{\mathbb{R}^3}|K(x-y)\mathbf{1}_\Omega(y)f(y)|\,dy. \tag{I.30}$$

Since $K \in L^1(\mathbb{R}^3)$, we can use Young's inequality for convolutions to get

$$\left\|\int_{|x-y|\leq\delta R}\frac{K(x-y)}{|x-y|}\big(\chi(x) - \chi(y)\big)f(y)dy\right\|_2 \leq \|\nabla\chi\|_\infty\|K\|_1\|\mathbf{1}_\Omega f\|_2. \tag{I.31}$$

For the second integral on the right hand side of (I.27) use exponential decay of the modified Bessel functions of the second kind. Since $\chi \leq 1$ and $f \in L^2$ we get

$$\left\| \int_{|x-y|\geq \delta R} \frac{K(x-y)}{|x-y|} \big(\chi(x)-\chi(y)\big) f(y) dy \right\|_2 \leq C e^{-\delta R} \|f\|_2. \tag{I.32}$$

This finishes the first part of the proof and we have

$$\|[\chi, \Lambda_+] f\|_2 \leq C\big(\|\nabla \chi\|_\infty \|\mathbf{1}_\Omega f\|_2 + e^{-\delta R}\|f\|_2\big) \tag{I.33}$$

Now let us bound $\|\nabla[\chi, \Lambda_+] f\|_2$. We estimate this norm in three parts, define the functions

$$\begin{aligned} I_1(x) &:= \frac{1}{4\pi^2} \int_{\mathbb{R}^3} \frac{\beta K_1(|x-y|)}{|x-y|} \big(\chi(x)-\chi(y)\big) f(y) dy \\ I_2(x) &:= \frac{i}{4\pi^2} \int_{\mathbb{R}^3} \frac{\boldsymbol{\alpha}\cdot(x-y) K_0(|x-y|)}{|x-y|^2} \big(\chi(x)-\chi(y)\big) f(y) dy \\ I_3(x) &:= \frac{i}{2\pi^2} \int_{\mathbb{R}^3} \frac{\boldsymbol{\alpha}\cdot(x-y) K_1(|x-y|)}{|x-y|^3} \big(\chi(x)-\chi(y)\big) f(y) dy. \end{aligned} \tag{I.34}$$

Note that

$$\|\nabla[\chi, \Lambda_+] f\|_2 \leq C(\|\nabla I_1\|_2 + \|\nabla I_2\|_2 + \|\nabla I_3\|_2). \tag{I.35}$$

the first two terms can be bounded by a similar approach as in the estimate of $\|[\chi, \Lambda_+] f\|_2$. We have

$$\begin{aligned} \left|\frac{\partial I_1(x)}{\partial x_j}\right| \leq{}& \frac{1}{4\pi^2} \int_{\mathbb{R}^3} \left| \beta \frac{(x_j-y_j) K_0(|x-y|)}{|x-y|^2} \big(\chi(x)-\chi(y)\big) f(y) \right| dy \\ &+ \frac{1}{4\pi^2} \int_{\mathbb{R}^3} \left| \beta \frac{2(x_j-y_j) K_1(|x-y|)}{|x-y|^3} \big(\chi(x)-\chi(y)\big) f(y) \right| dy \\ &+ \frac{1}{4\pi^2} \int_{\mathbb{R}^3} \left| \beta \frac{K_0(|x-y|)}{|x-y|} \partial_j \chi(x) f(y) \right| dy. \end{aligned} \tag{I.36}$$

Again we look at two regions separately. For $|x-y| \leq \delta R$ we have

$$|\chi(x)-\chi(y)| = |\chi(x)-\chi(y)| \mathbf{1}_\Omega(y) \leq \|\nabla\chi\|_\infty |x-y| \mathbf{1}_\Omega(y) \tag{I.37}$$

and

$$|\partial_j \chi(x)| = \mathbf{1}_\Omega(y) |\partial_{x_j}\chi(x)| \leq \|\nabla\chi\|_\infty \mathbf{1}_\Omega(y). \tag{I.38}$$

Using Young's convolution inequality yields the desired bound. Since the modified Bessel functions of second kind decay exponentially, the integral where $|x-y| \geq \delta R$ are exponentially small in δR. Since $|\partial_j \chi| \sim R^{-1}$ we get

$$\|\nabla I_1\|_2 \leq C R^{-1} \|\mathbf{1}_\Omega f\|_2 + C(1+R^{-1}) e^{-\delta R} \|f\|_2. \tag{I.39}$$

In the second part, the singularities of the kernels are "well-behaved" and analogously we get

$$\|\nabla I_2\|_2 \leq C R^{-1} \|\mathbf{1}_\Omega f\|_2 + C(1+R^{-1}) e^{-\delta R} \|f\|_2. \tag{I.40}$$

The third part has a kernel with a more severe singularity. The partial derivative of $I_3(x)$ in the jth coordinate is given by

$$\begin{aligned} &\frac{i}{2\pi^2} PV \int_{\mathbb{R}^3} \left(\alpha_j \frac{K_1(|x-y|)}{|x-y|^3} - \frac{\boldsymbol{\alpha}\cdot(x-y)(x_j-y_j) K_0(|x-y|)}{|x-y|^4} \right. \\ &\qquad \left. - \frac{4\boldsymbol{\alpha}\cdot(x-y)(x_j-y_j) K_1(|x-y|)}{|x-y|^5} \right) \big(\chi(x)-\chi(y)\big) f(y) dy \\ &+ \frac{i}{2\pi^2} PV \int_{\mathbb{R}^3} \frac{\boldsymbol{\alpha}\cdot(x-y)(x_j-y_j) K_0(|x-y|)}{|x-y|^4} \partial_j \chi(x) f(y) dy \end{aligned} \tag{I.41}$$

where the integrals are to be understood in the Cauchy principle value sense. The singularity of the kernel

$$\frac{(\boldsymbol{\alpha}\cdot x)x_j K_0(|x|)}{|x|^4} \tag{I.42}$$

around the origin can once again be cancelled out by taking the mean value estimate of $\chi(x)-\chi(y)$. We proceed as above to get the desired upper bound on the respective L^2-norm. Exponential decay of the Bessel functions yields exponential small bounds for the region $|x-y|\geq \delta R$. It remains to bound the integral in the region $|x-y|\leq \delta R$, where once again we have $\chi(x)-\chi(y)=\big(\chi(x)-\chi(y)\big)\mathbf{1}_\Omega(y)$. We use Taylor expansion of $\chi(x)$ around y to get

$$\chi(x)-\chi(y)=(x_k-y_k)\partial_k\chi(x)-\frac{1}{2}(x_k-y_k)(x_l-y_l)\partial^2_{kl}\chi(\xi) \tag{I.43}$$

for some vector ξ on the line segment connecting x and y. Substituting (I.43) into (I.41) for the region $|x-y|\leq\delta R$ we get

$$\begin{aligned}
&\frac{i\partial_{kl}\chi(x)}{2\pi^2}PV\int_{|x-y|\leq\delta R}\left(\alpha_j-\frac{4\boldsymbol{\alpha}\cdot(x-y)(x_j-y_j)}{|x-y|}\right)\frac{K_1(|x-y|)(x_k-y_k)}{|x-y|^3}(\mathbf{1}_\Omega f)(y)dy\\
&+\frac{i\partial_j\chi(x)}{2\pi^2}PV\int_{|x-y|\leq\delta R}\frac{\boldsymbol{\alpha}\cdot(x-y)K_1(|x-y|)}{|x-y|^3}(\mathbf{1}_\Omega f)(y)dy\\
&-\frac{i}{4\pi^2}\int_{|x-y|\leq\delta R}\left(\alpha_j-\frac{4\boldsymbol{\alpha}\cdot(x-y)(x_j-y_j)}{|x-y|^2}\right)\frac{K_1(|x-y|)}{|x-y|^3}\\
&\quad\times(x_k-y_k)(x_l-y_l)\partial^2_{kl}\chi(\xi)(\mathbf{1}_\Omega f)(y)dy.
\end{aligned} \tag{I.44}$$

The last integral in (I.44) is a convolution with a L^1 function and can thus be bounded from above by

$$C\|\partial^2\chi\|_\infty\|\mathbf{1}_\Omega f\|_2. \tag{I.45}$$

For the first two integrals in (I.44) we will use Theorem I.3. Obviously $\mathbf{1}_\Omega f$ is a L^2 function. Condition (I.17) is true since for non-negative $r\in\mathbb{R}$ we have

$$rK_0(r)<\frac{1}{2}\quad\text{and}\quad rK_1(r)\leq 1. \tag{I.46}$$

Condition (I.19) is true since the kernel is a product of a spherical symmetric function and an antisymmetric function. To verify that the kernel also satisfies (I.18) note that, according to [42], the condition

$$|\nabla K(x)|\leq B|x|^{-4} \tag{I.47}$$

implies (I.18). Note that since

$$K_0'(r)=-K_1(r) \tag{I.48}$$

and

$$K_1'(r)=-K_0(r)-\frac{K_1(r)}{r} \tag{I.49}$$

inequalities (I.46) give an easy way to check that condition (I.47) is fulfilled. Hence the L^2 norm of the (I.44) is bounded above by

$$C\left(\|\partial^2\chi\|_\infty+\|\nabla\chi\|_\infty\right)\|\mathbf{1}_\Omega f\|_2. \tag{I.50}$$

Note that $\|\partial^2\chi\|_\infty\sim R^{-2}$ and $\|\nabla\chi\|_\infty\sim R^{-1}$. Exponential decay of the Bessel functions yields

$$\|\nabla I_3\|_2\leq C(R^{-1}+R^{-2})\|\mathbf{1}_\Omega f\|_2+Ce^{-\delta R}\|f\|_2. \tag{I.51}$$

□

Note that this result implies in particular, that for any $f \in L^2$

$$\|[\chi, \Lambda_+]f\|_{H^{1/2}} \leq C\left((R^{-1} + R^{-2})\|\mathbf{1}_\Omega f\|_2 + e^{-\delta R}\|f\|_2\right) \tag{I.52}$$

The following lemma is the analogue to the localization estimate [32, Lemma 6]. We use the refined estimate on the localization error to obtain a suitable bound. In the following it will be necessary to distinguish the operators with and without projections Λ_+^N, let $\mathcal{H}^{BR}$ be such that

$$\Lambda_+^N \mathcal{H}^{BR} \Lambda_+^N = H^{BR}$$

for H^{BR} as in (I.1). Similarly let $\mathcal{H}_\beta^{BR}$ be such that

$$\Lambda_+^N \mathcal{H}_\beta^{BR} \Lambda_+^N = H_\beta^{BR}$$

with H_β^{BR} defined in (I.12). Let $\mathbf{1}_\Omega^{(j)}$ be the function $\mathbf{1}_\Omega$ in the jth electron direction extended by identity to the whole space.

Lemma I.5 (Localization error Brown-Ravenhall)**.** *Let $\{J_\beta\}_{\beta \in \mathcal{D}_N^3}$ be a partition of unity that is constructed from one-electron cutoff functions satisfying the conditions of Lemma I.4. Then there exist $C < \infty$ and $a > 0$ such that for any $\psi \in \Lambda_+^N H^{1/2}(\mathbb{R}^{3N}, \mathbb{C}^{4N})$ we have*

$$\begin{aligned} &\left| \langle \psi, (H^{BR} - \mu)\psi - \sum_{\beta \in \mathcal{D}_N^3} \langle \Lambda_+^N J_\beta \psi, (H^{BR} - \mu)\Lambda_+^N J_\beta \psi \rangle \right| \\ &\leq C \sum_{\beta \in \mathcal{D}_N^3} \left((|D|^{-1} + |D|^{-2}) \sum_{i=1}^N \|\mathbf{1}_\Omega^{(i)} J_\beta \psi\| + e^{-a|D|}\|J_\beta \psi\|^2\right). \end{aligned} \tag{I.53}$$

Proof. Note that following the proof in [32] we have

$$\begin{aligned} &\langle \psi, (H^{BR} - \mu)\psi \rangle - \sum_{\beta \in \mathcal{D}_N^3} \langle \Lambda_+^N J_\beta \psi, (H^{BR} - \mu)\Lambda_+^N J_\beta \psi \rangle \\ &= \sum_{\beta \in \mathcal{D}_N^3} \langle (\mathcal{H}^{BR} - \mu)[J_\beta, \Lambda_+^N]\psi, J_\beta \Lambda_+^N \psi \rangle + \sum_{\beta \in \mathcal{D}_N^3} \langle \Lambda_+^N J_\beta \psi, (\mathcal{H}^{BR} - \mu)[J_\beta, \Lambda_+^N]\psi \rangle. \end{aligned} \tag{I.54}$$

First we use Cauchy-Schwarz inequality, then since $\mathcal{H}_\beta^{BR}$ is form bounded in the $H^{1/2}$, see the proof of [32, Lemma 6] norm we get

$$\begin{aligned} &\left|\langle (\mathcal{H}^{BR} - \mu)[J_\beta, \Lambda_+^N]\psi, J_\beta \Lambda_+^N \psi \rangle + \langle \Lambda_+^N J_\beta \psi, (\mathcal{H}^{BR} - \mu)[J_\beta, \Lambda_+^N]\psi \rangle\right| \\ &\leq C\|[J_\beta, \Lambda_+^N]\psi\|_{H^{1/2}} \|J_\beta \psi\|_2 + C\|\Lambda_+^N J_\beta \psi\|_2 \|[J_\beta, \Lambda_+^N]\psi\|_{H^{1/2}}. \end{aligned} \tag{I.55}$$

Note that $\Lambda_+^N \psi = \psi$ since $\psi \in \Lambda_+^N H^{1/2}(\mathbb{R}^{3N}, \mathbb{C}^{4N})$ and using (I.52) we get

$$\begin{aligned} &C\|[J_\beta, \Lambda_+^N]\psi\|_{H^{1/2}} \|J_\beta \psi\|_{H^{1/2}} + C\|\Lambda_+^N J_\beta \psi\|_{H^{1/2}} \|[J_\beta, \Lambda_+^N]\psi\|_{H^{1/2}} \\ &\leq C\Big((|D|^{-1} + |D|^{-2}) \sum_{i=1}^N \|\mathbf{1}_\Omega^{(i)} J_\beta \psi\| + e^{-a|D|}\|J_\beta \psi\|^2\Big). \end{aligned} \tag{I.56}$$

□

As in the main part, we need to partition the state space with respect to three cluster decompositions. Let $\{J_\beta\}_{\beta \in \mathcal{D}_N^3}$ be a partition of unity as defined in Chapter 11.1. In abuse of notation we will not define the unitary shifts of the system. It's correct implementation can be seen in the main part.

Remark I.6. *In abuse of notation we omit the shifting of the system towards the origin as we do in the Herbst case. Notions such as exponential decay, rotational symmetry and questions of support have to be undestood with respect to the relative electron positions. That is, if we say that a ground state $\phi \in \Lambda_+^N H^{1/2}(\mathbb{R}^{3N}, \mathbb{C}^{4N})$ of H_β^{BR} is exponentially localized, we mean that ϕ is exponentially localized around X_1 in coordinates $j \in \mathcal{C}_1$ and around X_2 in directions $i \in \mathcal{C}_2$. This is particularly relevant for the intercluster expansion, where we require the relative positions to be small with respect to $|D|$.*

I.3. **Estimate from below.** We will proceed similar to the Herbst case. Let us define the energy functionals $L[J_\beta\psi]$ where we compare the mean value of $(H^{BR}-\mu)$ for the localized state with the localization error similar to (12.4).

More explicitly, for $\psi \in H^{1/2}(\mathbb{R}^{3N}, \mathbb{C}^{4N})$

$$\langle \psi, (H^{BR} - \mu)\psi\rangle = \sum_{\beta \in \mathcal{D}_N^3} L[J_\beta\psi], \tag{I.57}$$

where the energy functional is given by

$$\begin{aligned} L[J_\beta\psi] &= \langle J_\beta\psi, \Lambda_+^N(\mathcal{H}^{BR} - \mu)\Lambda_+^N J_\beta\psi\rangle \\ &\quad + \big(\langle \psi, (H^{BR} - \mu)\psi - \langle J_\beta\psi, \Lambda_+^N(\mathcal{H}^{BR} - \mu)\Lambda_+^N J_\beta\psi\rangle\big). \end{aligned} \tag{I.58}$$

The second line is the full localization error for the operator H^{BR}. Once again, we will distinguish between the three cases.
Case 1: Decompositions $\beta \in \mathcal{D}_N^3$ such that $\mathcal{C}_0(\beta) \neq \emptyset$
Case 2: Ionic decompositions $\beta \in \mathcal{D}_N^3$ with $\mathcal{C}_0(\beta) = \emptyset$ and $\sharp\mathcal{C}_1(\beta) \neq Z_1$
Case 3: Atomic decompositions $\beta \in \mathcal{D}^{at}$

For the first case, where $\mathcal{C}_0(\beta) \neq \emptyset$, there is at least one electron that is far from both the nuclei. This is the setting that is studied in the proof of the HVZ theorem. Thus the energy functional $L[J_\beta\psi]$ must be non-negative for large $|D|$. The HVZ theorem in the Brown-Ravenhall case, [30, Theorem 6] thus yields

$$L[J_\beta\psi] \geq 0 \quad \text{for all } \beta \text{ with } \mathcal{C}_0(\beta) \neq \emptyset. \tag{I.59}$$

For an ionic decomposition, that is $\mathcal{C}_0(\beta) = \emptyset$ and $\sharp\mathcal{C}_1(\beta) \neq Z_1$, we will use Condition 1) in Theorem I.1 which ensures that we have $d = \mu_\beta - \mu > 0$, independently of $|D|$ and ψ. This yields

$$\begin{aligned} L[J_\beta\psi] &= \langle J_\beta\psi, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N J_\beta\psi\rangle + \langle J_\beta\psi, \Lambda_+^N I_\beta \Lambda_+^N J_\beta\psi\rangle + \mathcal{LE}[J_\beta\psi] \\ &\geq d\|J_\beta\psi\|^2 + \langle J_\beta\psi, \Lambda_+^N I_\beta \Lambda_+^N J_\beta\psi\rangle + \mathcal{LE}[J_\beta\psi]. \end{aligned} \tag{I.60}$$

Similar to the Herbst case, the term $\mathcal{LE}[J_\beta\psi]$ is the portion of the localization error attributed to $J_\beta\psi$. Notice that for the two last terms in the second line of (I.60) there exist $\epsilon_1(|D|)$, $\epsilon_2(|D|) > 0$ with $\epsilon_{1,2}(|D|) \xrightarrow{|D|\to\infty} 0$ such that

$$\langle J_\beta\psi, \Lambda_+^N I_\beta \Lambda_+^N J_\beta\psi\rangle \geq -\epsilon_1(|D|)\|J_\beta\psi\|^2 \tag{I.61}$$

and

$$\mathcal{LE}[J_\beta\psi] \geq -\epsilon_2(|D|)\|J_\beta\|^2. \tag{I.62}$$

The latter follows from Lemma I.5, this crude estimate is sufficient in this case. Thus, since d is independent of $|D|$, going back to (I.60) we get that for large $|D|$

$$L[J_\beta\psi] > 0 \text{ for all } \beta \in \mathcal{D}_N^3 \text{ with } \mathcal{C}_0(\beta) = \emptyset \text{ and } \sharp\mathcal{C}_1(\beta) \neq Z_1. \tag{I.63}$$

In the third case, the estimates have to be handled with more care. We will fix an atomic decomposition $\beta \in \mathcal{D}^{at}$ and similar to the Herbst case, will project the state

$J_\beta\psi$ onto the ground state space of H_β^{BR}, we thus get a complex valued γ and a normalized ground state ϕ, together with a remainder $g \perp \phi$ such that

$$J_\beta\psi = \gamma\phi + g. \tag{I.64}$$

Recall that we aim to bound the energy functional

$$L[J_\beta\psi] = \langle J_\beta\psi, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N J_\beta\psi\rangle + \langle J_\beta\psi, \Lambda_+^N I_\beta \Lambda_+^N J_\beta\psi\rangle + \mathcal{LE}[J_\beta\psi]. \tag{I.65}$$

We will use the decomposition of (I.64) for the first term of (I.64) to get

$$\begin{aligned}\langle J_\beta\psi, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N J_\beta\psi\rangle &= \langle \gamma\phi, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N \gamma\phi\rangle \\ &\quad + 2\,\mathrm{Re}\langle g, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N \gamma\phi\rangle \\ &\quad + \langle g, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N g\rangle.\end{aligned} \tag{I.66}$$

The first two terms on the right hand side of (I.66) are zero as ϕ is a ground state and thus

$$\langle J_\beta\psi, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N J_\beta\psi\rangle = \langle g, \Lambda_+^N(\mathcal{H}_\beta^{BR} - \mu)\Lambda_+^N g\rangle. \tag{I.67}$$

The second term of the energy functional in (I.60) is the intercluster interaction term. We have

$$\langle J_\beta\psi, \Lambda_+^N I_\beta \Lambda_+^N J_\beta\psi\rangle = \langle \gamma\phi, \Lambda_+^N I_\beta \Lambda_+^N \gamma\phi\rangle + 2\,\mathrm{Re}\,\gamma\langle g, \Lambda_+^N I_\beta \Lambda_+^N \phi\rangle + \langle g, \Lambda_+^N I_\beta \Lambda_+^N g\rangle. \tag{I.68}$$

Remark I.7. *Note that the potential I_β is not locally supported. This is necessary in order to ensure that the multipole expansion converges and that we can apply Newton's theorem. We can write $J_\beta = \mathbf{1}_\beta J_\beta$, where $\mathbf{1}_\beta$ is the indicator function of* $\mathrm{supp}(J_\beta)$. *If we then commute J_β towards the states ψ we can keep the indicator functions besides the potential. By this small sleight we can identify I_β with the local potential $\mathbf{1}_\beta I_\beta$.*

The ground state space is contained in the positive spectral subspace onto which Λ_+^N projects. Thus $\Lambda_+^N\phi = \phi$. By Condition 2) and Remark I.7 we can apply Newton's theorem and get

$$\langle \gamma\phi, \Lambda_+^N I_\beta \Lambda_+^N \gamma\phi\rangle = 0. \tag{I.69}$$

We will use the multipole expansion for the intercluster potential I_β. We will use exponential decay of the ground state ϕ, see [31], to bound the remainder in the intercluster potential expansion, compare Lemma D.2 and Corollary D.4. For the functions f_ℓ defined in (D.8) we get that for any $k \geq 3$

$$2\,\mathrm{Re}\,\gamma\langle g, \Lambda_+^N I_\beta \Lambda_+^N \phi\rangle = \sum_{\ell=2}^{k} \frac{2\,\mathrm{Re}\,\gamma\langle g, \Lambda_+^N f_\ell \phi\rangle}{|D|^{\ell+1}} + C_k |D|^{-(k+2)} \|\gamma\phi\| \|g\|. \tag{I.70}$$

Note that for $\ell \geq 3$ we can use Young's product inequality to get

$$\frac{2\,\mathrm{Re}\,\gamma\langle g, \Lambda_+^N f_\ell \phi\rangle}{|D|^{\ell+1}} \geq -\frac{\|g\|^2}{|D|^{2\ell-5}} - \frac{|\gamma|^2 \|\Lambda_+^N f_\ell \phi\|^2}{|D|^7}. \tag{I.71}$$

Since ϕ decays exponentially, for any $\ell \in \mathbb{N}$ there is a constant C_ℓ such that $\|\Lambda_+^N f_\ell \phi\| \leq C_\ell \|\phi\|$. Overall, expanding the intercluster potential up to the term $k = 5$ yields

$$2\,\mathrm{Re}\,\gamma\langle g, \Lambda_+^N I_\beta \Lambda_+^N \phi\rangle \geq 2|D|^{-3}\,\mathrm{Re}\,\gamma\langle g, \Lambda_+^N f_2 \phi\rangle - \frac{C\|g\|^2}{|D|} - \frac{C\|g\|\|\gamma\phi\|}{|D|^7} - \frac{C\|\gamma\phi\|^2}{|D|^7}. \tag{I.72}$$

By Young's inequality the term $-C|D|^{-7}\|g\|\|\gamma\phi\|$ is dominated by the two adjacent terms for large $|D|$ and we get

$$2\operatorname{Re}\gamma\langle g,\Lambda_+^N I_\beta\Lambda_+^N\phi\rangle \ \geq 2|D|^{-3}\operatorname{Re}\gamma\langle g,\Lambda_+^N f_2\phi\rangle - \frac{C\|g\|^2}{|D|} - \frac{C\|\gamma\phi\|^2}{|D|^7}. \tag{I.73}$$

For the first summand on the right hand side of (I.73) we would like to apply the operators $(H_\beta^{BR}-\mu)^{\pm\frac{1}{2}}$ to g and $\Lambda_+^N f_2\phi$ respectively. Notice that $(H_\beta^{BR}-\mu)$ is a positive operator on the orthogonal complement of the ground state space to which both $f_2\phi$ and g belong. For g this is clear and for $f_2\phi$ this follows from the rotational symmtery of $f_2\phi$, which is discussed at length in the Herbst case. We thus have

$$2|D|^{-3}\operatorname{Re}\gamma\langle g,\Lambda_+^N f_2\phi\rangle = 2|D|^{-3}\operatorname{Re}\gamma\langle (H_\beta^{BR}-\mu)^{\frac{1}{2}}g,(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\rangle. \tag{I.74}$$

Together with a fraction (of order $b^{-2}<1$) of the term

$$\langle g,(H_\beta^{BR}-\mu)g\rangle = \|(H_\beta^{BR}-\mu)^{\frac{1}{2}}g\|^2$$

we obtain

$$\begin{aligned}
&\frac{1}{b^2}\|(H_\beta^{BR}-\mu)^{\frac{1}{2}}g\|^2 + 2|D|^{-3}\operatorname{Re}\gamma\langle (H_\beta^{BR}-\mu)^{\frac{1}{2}}g,(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\rangle\\
&= \frac{-b^2}{|D|^6}\|(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2 + \|\frac{1}{b}(H_\beta^{BR}-\mu)^{\frac{1}{2}}g + \frac{b}{|D|^3}(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2\\
&\quad\geq -\frac{b^2}{|D|^6}\|(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2.
\end{aligned} \tag{I.75}$$

Remark I.8. *Note that this estimate holds for any value of b, thus one might be tempted to conclude that choosing very large b will minimize the energy. In this case however the non-negative term will be very positive if we choose a big value for b. In fact, in the estimate from above we will see that the optimal choice is $b=1+\epsilon$. We will choose $b^2=2$, this will prevent us from obtaining the optimal coefficient for the vdW interaction, but it is sufficient to prove the result.*

To finish with the bound of the intercluster potential, we note that the first term in the intercluster expansion is of order $|D|^{-3}$. Furthermore, since $J_\beta\psi$ is locally supported and ϕ decays exponentially, the equation $g=J_\beta\psi-J_\beta\psi$ yields exponential decay of g outside the support of J_β. Which allows us to estimate the expected intercluster energy of the state g as

$$\langle\Lambda_+^N g, I_\beta\Lambda_+^N g\rangle \geq -\frac{C}{|D|^3}\|\Lambda_+^N g\|^2. \tag{I.76}$$

This estimate contains only the remainder estimate, since the terms $\ell=0$ and $\ell=1$ in the expansion are zero. The inequality uses that the second term in the second line of (I.75) is non-negative. Coming back to (I.66), gathering bounds in (I.67),(I.69),(I.70),(I.76) and (I.75) we get

$$\begin{aligned}
L[J_\beta\psi] \geq &-\frac{2|\gamma|^2}{|D|^6}\|(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2 - \frac{C\|\gamma\phi\|^2}{|D|^7} + \frac{1}{2}\|(H_\beta^{BR}-\mu)^{\frac{1}{2}}g\|^2\\
&- \frac{C\|g\|^2}{|D|} - \frac{C\|\Lambda_+^N g\|^2}{|D|^3} - \mathcal{LE}[J_\beta\psi].
\end{aligned} \tag{I.77}$$

By Lemma I.5, the localization error is bounded by

$$-\mathcal{LE}[J_\beta\psi] \geq -C\big((|D|^{-1}+|D|^{-2})\sum_{i=1}^N \|\mathbf{1}_\Omega J_\beta\psi\|\|J_\beta\psi\| + e^{-a|D|}\|J_\beta\psi\|^2\big). \tag{I.78}$$

Using the orthogonal decomposition $J_\beta\psi = \gamma\phi + g$ and the fact that the terms $\|\mathbf{1}_\Omega^{(i)}\phi\|$ are exponentially small in $|D|$ we obtain

$$-\mathcal{LE}[J_\beta\psi] \geq -C\big(|D|^{-1} + e^{-a|D|}\big)\|g\|^2 - Ce^{-c|D|}\|\gamma\phi\|^2). \tag{I.79}$$

We will gather the terms that depend on the norm of g and contribute to the energy functional formally as $L[g]$ where

$$L[g] := \frac{1}{2}\|(H_\beta^{BR} - \mu)^{\frac{1}{2}} g\|^2 - C\left(\frac{\|g\|^2}{|D|} + \frac{\|\Lambda_+^N g\|^2}{|D|^3} + e^{-a|D|}\|g\|^2\right). \tag{I.80}$$

The lower bound of the operator $(H_\beta^{BR} - \mu)^{\frac{1}{2}}$ is independent of $|D|$. For large $|D|$ the first (positive) factor on the right hand side will dominate the latter negative contributions to the energy. Overall for large $|D|$ the energy functional $L[g]$ is thus non-negative

$$L[g] \geq 0. \tag{I.81}$$

This yields that for sufficiently large $|D|$ we get

$$L[J_\beta\psi] \geq -\frac{2|\gamma|^2}{|D|^6}\|(H_\beta^{BR} - \mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2 - \frac{C\|\gamma\phi\|^2}{|D|^7} - Ce^{-c|D|}\|\gamma\phi\|^2. \tag{I.82}$$

Summation over all decompositions $\beta \in \mathcal{D}^{at}$ yields the bound from below.

I.4. **Estimate from above.** For the construction of a trial function we pick one atomic decomposition $\beta \in \mathcal{D}^{at}$ and a normalized ground state ϕ of the Hamiltonian H_β^{BR}. We will set

$$\psi_0 := \phi - |D|^{-3}(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi \tag{I.83}$$

and calculate

$$\begin{aligned}
&\langle\psi_0, (H^{BR} - \mu)\psi_0\rangle \\
&= \langle\phi, (H_\beta^{BR} - \mu)\phi\rangle + \langle\phi, I_\beta^{BR}\phi\rangle - 2\,\mathrm{Re}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, (H_\beta^{BR} - \mu)\phi\rangle \\
&\quad - \frac{2}{|D|^3}\,\mathrm{Re}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, I_\beta^{BR}\phi\rangle + \frac{1}{|D|^6}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, \Lambda_+^N f_2\phi\rangle \\
&\quad + \frac{1}{|D|^6}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, I_\beta^{BR}\phi\rangle.
\end{aligned} \tag{I.84}$$

By the ground state properties of ϕ, the terms in the first line on the right hand side of (I.84) are zero, that is

$$\langle\phi, (H_\beta^{BR} - \mu)\phi\rangle = \langle\phi, I_\beta^{BR}\phi\rangle = 2\,\mathrm{Re}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, (H_\beta^{BR} - \mu)\phi\rangle = 0 \tag{I.85}$$

We will once again apply a multipole expansion for the intercluster potential. For the last term in (I.84), since the expansion starts with a term of order $|D|^{-3}$ we get

$$\frac{1}{|D|^6}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, I_\beta^{BR}\phi\rangle \leq \frac{C\|(H_\beta^{BR} - \mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2}{|D|^9}. \tag{I.86}$$

The other intercluster interaction term yields

$$\begin{aligned}
&-\frac{2}{|D|^3}\,\mathrm{Re}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, I_\beta^{BR}\phi\rangle \\
&\qquad \leq -\frac{2}{|D|^6}\,\mathrm{Re}\langle(H_\beta^{BR} - \mu)^{-1}\Lambda_+^N f_2\phi, \Lambda_+^N f_2\phi\rangle + \frac{C\|\phi\|^2}{|D|^7}.
\end{aligned} \tag{I.87}$$

Let us go back to (I.84), note that the term of order $|D|^{-9}$ in (I.85) is dominated for large $|D|$, with (I.87) we get

$$\begin{aligned}\langle\psi_0,(H^{BR}-\mu)\psi_0\rangle \leq -\frac{2}{|D|^6}\,\mathrm{Re}\langle(H_\beta^{BR}-\mu)^{-1}\Lambda_+^N f_2\phi,\Lambda_+^N f_2\phi\rangle + \frac{C\|\phi\|^2}{|D|^7} \\ + \frac{1}{|D|^6}\langle(H_\beta^{BR}-\mu)^{-1}\Lambda_+^N f_2\phi,\Lambda_+^N f_2\phi\rangle.\end{aligned} \tag{I.88}$$

This yields

$$\langle\psi_0,(H^{BR}-\mu)\psi_0\rangle \leq -\frac{1}{|D|^6}\|(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2 + \frac{C\|\phi\|^2}{|D|^7}. \tag{I.89}$$

Note that $\|(H_\beta^{BR}-\mu)^{-\frac{1}{2}}\Lambda_+^N f_2\phi\|^2$ is finite since the resolvent is bounded and $\|\psi_0\| = 1+\frac{C}{|D|^3}$. Dividing by $\|\psi_0\|^2$ in (I.89) yields the bound from above which finishes the proof.

References

[1] Milton Abramowitz and Irene A Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. Number 55. Courier Corporation, 1964.

[2] Shmuel Agmon. *Lectures on exponential decay of solutions of second-order elliptic equations: bounds on eigenfunctions of N-body Schrödinger operators*, volume 29 of *Mathematical Notes*. Princeton University Press, Princeton, NJ; University of Tokyo Press, Tokyo, 1982.

[3] Ioannis Anapolitanos and Israel Michael Sigal. Long-range behavior of the van der Waals force. *Comm. Pure Appl. Math.*, 70(9):1633–1671, 2017.

[4] George E. Andrews, Richard Askey, and Ranjan Roy. *Special Functions*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 1999.

[5] A. A. Balinsky and W. D. Evans. *Spectral analysis of relativistic operators*. Imperial College Press, London, 2011.

[6] Jean-Marie Barbaroux, Thomas Chen, Vitali Vougalter, and Semjon Vugalter. On the ground state energy of the translation invariant Pauli-Fierz model. *Proc. Amer. Math. Soc.*, 136(3):1057–1064, 2008.

[7] Jean-Marie Barbaroux, Thomas Chen, Vitali Vougalter, and Semjon Vugalter. Quantitative estimates on the binding energy for hydrogen in non-relativistic QED. *Ann. Henri Poincaré*, 11(8):1487–1544, 2010.

[8] Jean-Marie Barbaroux and Semjon Vugalter. Quantitative estimates on the binding energy for hydrogen in non-relativistic QED, II: The spin case. *Rev. Math. Phys.*, 26(8):1450016, 57, 2014.

[9] Jean-Marie Barbaroux and Semjon A. Vugalter. On the ground state energy of the translation invariant Pauli-Fierz model. II. *Doc. Math.*, 17:401–415, 2012.

[10] Rafael Benguria and Elliott H. Lieb. *Proof of the Stability of Highly Negative Ions in the Absence of the Pauli Principle*, pages 82–85. Springer Berlin Heidelberg, Berlin, Heidelberg, 1991.

[11] Rafael Benguria, Heinz Siedentop, and Edgardo Stockmeyer. Dissociation of Homonuclear Relativistic Molecular Ions. *Annales Henri Poincaré*, 2(1):27–40, 2001.

[12] René Carmona, Wen Chen Masters, and Barry Simon. Relativistic schrödinger operators: asymptotic behavior of the eigenfunctions. *Journal of functional analysis*, 91(1):117–142, 1990.

[13] H.L. Cycon, R. G. Froese, W. Kirsch, and B. Simon. *Schrödinger Operators: With Applications to Quantum Mechanics and Global Geometry*. Theoretical and Mathematical Physics. Springer-Verlag Berlin Heidelberg, 1987.

[14] Charles L. Fefferman and Rafael de la Llave. Relativistic stability of matter (I). *Revista Matemática Iberoamericana*, 2(1-2):119–213, 1986.

[15] Charles L. Fefferman and Luis A. Seco. Asymptotic neutrality of large ions. *Comm. Math. Phys.*, 128(1):109–130, 1990.

[16] Marcel Griesemer, Elliott H. Lieb, and Michael Loss. *Ground states in non-relativistic quantum electrodynamics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[17] Stephen J Gustafson and Israel Michael Sigal. *Mathematical concepts of quantum mechanics*. Springer Science & Business Media, 2011.

[18] Morton Hamermesh. *Group theory and its application to physical problems*. Addison-Wesley Series in Physics. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1962.

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

[20] Walter Hunziker. On the spectra of Schrödinger multiparticle Hamiltonians. *Helv. Phys. Acta*, 39:451–462, 1966.

[21] Enno Lenzmann and Mathieu Lewin. Minimizers for the hartree-fock-bogoliubov theory of neutron stars and white dwarfs. *Duke Math. J.*, 152(2):257–315, 04 2010.

[22] R. T. Lewis, H. Siedentop, and S. Vugalter. The essential spectrum of relativistic multiparticle operators. *Annales de l'institut Henri Poincaré (A) Physique théorique*, 67(1):1–28, 1997.

[23] David R Lide. *CRC handbook of chemistry and physics*, volume 85. CRC press, 2004.

[24] E. H. Lieb and M. Loss. *Analysis*. Crm Proceedings & Lecture Notes. American Mathematical Society, 2001.

[25] Elliott H. Lieb. The stability of matter: from atoms to stars. *Bull. Amer. Math. Soc. (N.S.)*, 22(1):1–49, 1990.

[26] Elliott H. Lieb, Israel M. Sigal, Barry Simon, and Walter Thirring. Asymptotic neutrality of large-Z ions. *Phys. Rev. Lett.*, 52(12):994–996, 1984.

[27] Elliott H. Lieb and Walter E. Thirring. Universal nature of van der Waals forces for Coulomb systems. *Phys. Rev. A*, 34:40–46, Jul 1986.

[28] Elliott H. Lieb and Horng-Tzer Yau. The stability and instability of relativistic matter. *Comm. Math. Phys.*, 118(2):177–213, 1988.

[29] John D. Morgan III and Barry Simon. Behavior of molecular potential energy curves for large nuclear separations. *International journal of quantum chemistry*, 17(6):1143–1166, 1980.

[30] Sergey Morozov. Essential spectrum of multiparticle Brown-Ravenhall operators in external field. *Doc. Math.*, 13:51–79, 2008.

[31] Sergey Morozov. Exponential decay of eigenfunctions of Brown-Ravenhall operators. *J. Phys. A*, 42(47):475206, 16, 2009.

[32] Sergey Morozov and Semjon Vugalter. Stability of atoms in the Brown-Ravenhall model. *Ann. Henri Poincaré*, 7(4):661–687, 2006.

[33] P. T. Nam. On the number of electrons that a nucleus can bind. In *XVIIth International Congress on Mathematical Physics*, pages 504–511. World Sci. Publ., Hackensack, NJ, 2014.

[34] Phan Thành Nam. New bounds on the maximum ionization of atoms. *Comm. Math. Phys.*, 312(2):427–445, 2012.

[35] F. Nardini. Exponential decay for the eigenfunctions of the two body relativistic hamiltonian. *Journal d'Analyse Mathématique*, 47(1):87–109, Dec 1986.

[36] F. Nardini. On the asymptotic behaviour of the eigenfunctions of the relativistic N-body Schrödinger operator. *Boll. Un. Mat. Ital. A (7)*, 2:365–369, 1988.

[37] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[38] Michael Reed and Barry Simon. *Methods of modern mathematical physics. IV. Analysis of operators.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1978.

[39] Michael Reed and Barry Simon. *Methods of modern mathematical physics. III.* Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1979. Scattering theory.

[40] Michael Reed and Barry Simon. *Methods of modern mathematical physics. I.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, second edition, 1980. Functional analysis.

[41] L. A. Seco, I. M. Sigal, and J. P. Solovej. Bound on the ionization energy of large atoms. *Comm. Math. Phys.*, 131(2):307–315, 1990.

[42] Elias M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[43] S. A. Vugal' ter and G. M. Zhislin. On the asymptotics of the discrete spectrum of given symmetry of multiparticle Hamiltonians. *Trudy Moskov. Mat. Obshch.*, 54:186–212, 279, 1992.

[44] S. A. Vugal' ter and G. M. Zhislin. On the discrete spectrum of the negative hydrogen ion with a homogeneous magnetic field. *Lett. Math. Phys.*, 27(3):163–177, 1993.

[45] S. A. Vugal' ter and G. M. Zhislin. Spectral asymptotics of N-particle Schrödinger operators with a homogeneous magnetic field on subspaces with fixed SO(2) symmetry. *Algebra i Analiz*, 5(2):108–125, 1993.

[46] Hermann Weyl. *The theory of groups and quantum mechanics.* Dover Publications, Inc., New York, 1950. Translated from the second (revised) German edition by H. P. Robertson, Reprint of the 1931 English translation.

[47] Zong-Chao Yan, James F. Babb, A. Dalgarno, and G. W. F. Drake. Variational calculations of dispersion coefficients for interactions among h, he, and li atoms. *Phys. Rev. A*, 54:2824–2833, Oct 1996.

[48] G. M. Zhislin. A study of the spectrum of the Schrödinger operator for a system of several particles'. *Tr. Mosk. Mat. Obs.*, 9:81–120, 1960.

[49] G. M. Zhislin. Spectra of hamiltonians of molecule pseudorelativistic electrons in spaces of functions with permutational and point symmetry. *Functional Analysis and Its Applications*, 40(2):134–138, 2006.

[50] G. M. Zislin and A. G. Sigalov. The spectrum of the energy operator for atoms with fixed nuclei on subspaces corresponding to irreducible representations of the group of permutations. *Izv. Akad. Nauk SSSR Ser. Mat.*, 29:835–860, 1965.

Michael C. Hartig, Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France.

Email address: michael.hartig@univ-tln.fr

Michael HARTIG

Centre de Physique Théorique, Université de Toulon

Loi de van der Waals-London pour les systèmes d'atomes et de molécules relativistes

On considère un système constitué de plusieurs atomes où les noyaux sont supposés fixes et ponctuels. Les particules interagissent via le potentiel de Coulomb et les électrons ont une énergie cinétique pseudo-relativiste donnée par $(p^2+m^2)^{1/2}$-m. On démontre la loi de van der Waals-London qui dit que l'énergie d'interaction entre atomes neutres décroit comme $|D|^{-6}$ où $|D|$ est la distance entre atomes. Nous calculons rigoureusement tous les termes de l'énergie de liaison jusqu'à l'ordre $|D|^{-9}$ avec un terme d'erreur en $O(|D|^{-10})$. Comme étape intermédiaire, nous établissons la décroissance exponentielle des fonctions propres de l'opérateur de Schrödinger à plusieurs particules avec les symétries imposées par le principe de Pauli, et des estimations sur le terme d'erreur de localisation. Nous montrons de plus la loi de van der Waals-London pour l'opérateur de Dirac projeté connu sous le nom d'opérateur de Brown et Ravenhall. Dans ce dernier cas on obtient un terme d'erreur en $O(|D|^{-7})$.

Mot clés : loi de van der Waals-London, énergie cinétique pseudo-relativiste, erreur de localisation.

Van der Waals-London interaction for atoms with pseudo-relativistic kinetic energy

We consider a multiatomic system where the nuclei are assumed to be point charges at fixed positions. Particles interact via Coulomb potential and electrons have relativistic kinetic energy given by $(p^2+m^2)^{1/2}$-m. We prove the van der Waals-London law, which states that the interaction energy between neutral atoms decays as the sixth power of the distance $|D|$ between the atoms. We rigorously compute all terms in the binding energy up to order $|D|^{-9}$ with error term of order $O(|D|^{-10})$. As intermediate steps we prove exponential decay of eigenfunctions of multiparticle Schrödinger operators with permutation symmetry imposed by the Pauli principle and new estimates of the localization error. In addition we prove the van der Waals-London law for the projected Dirac operator, known as the Brown-Ravenhall operator. In this case we do not calculate the coefficients explicitly and we obtain an error term of order $O(|D|^{-7})$.

Keywords : Van der Waals-London forces, pseudo-relativistic kinetic energy operator, localization error.