



HAL
open science

Mathematical modelling and numerical simulation of the electronic structure of crystals in presence of local defects

Amélie Deleurence

► **To cite this version:**

Amélie Deleurence. Mathematical modelling and numerical simulation of the electronic structure of crystals in presence of local defects. Mathematics [math]. Ecole des Ponts ParisTech, 2008. English. NNT: . pastel-00004782

HAL Id: pastel-00004782

<https://pastel.hal.science/pastel-00004782>

Submitted on 17 Feb 2009

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.

THÈSE

présentée pour l'obtention du titre de

Docteur de l'Ecole Nationale des Ponts et Chaussées

Spécialité : Mathématique, Informatique

par

Amélie DELEURENCE

Sujet : *Modélisation mathématique et simulation numérique
de la structure électronique de cristaux en présence
de défauts ponctuels*

Soutenue le 5 décembre 2008 devant le jury composé de :

Président :	Christian Brouder
Rapporteurs :	François Castella Gianluca Panati Eric Séré
Examineurs :	Xavier Blanc Thierry Deutsch
Directeur de thèse :	Eric Cancès
Co-directeur de thèse :	Mathieu Lewin
Invité :	Romain Vidal



Ecole Nationale des Ponts et Chaussées

Région Ile de France

*Es verdad ; pues reprimamos
esta fiera condición,
esta furia, esta ambición,
por sí alguna vez soñamos ;
y sí haremos, pues estamos
en mundo tan singular,
que el vivir sólo es soñar ;
y la experiencia me enseña
que el hombre que vive, sueña
lo que es, hasta despertar.
Sueña el rey que es rey, y vive
con este engaño mandando,
disponiendo y gobernando ;
y este aplauso, que recibe
prestado, en el viento escribe,
y en cenizas le convierte
la muerte, ¡desdicha fuerte !
¿Que hay quien intente reinar,
viendo que ha de despertar
en el sueño de la muerte !
Sueña el rico en su riqueza,
que más cuidados le ofrece ;
sueña el pobre que padece
su miseria y su pobreza ;
sueña el que a medrar empieza,
sueña el que afana y pretende,
sueña el que agravia y ofende,
y en el mundo, en conclusión,
todos sueñan lo que son,
aunque ninguno lo entiende.
Yo sueño que estoy aquí
de estas prisiones cargado,
y soñé que en otro estado
más lisonjero me vi.
¿Qué es la vida ? Un frenesí.
¿Qué es la vida ? Una ilusión,
una sombra, una ficción,
y el mayor bien es pequeño ;
que toda la vida es sueño,
y los sueños, sueños son.*

Calderón de la Barca

*À mes parents,
À la mémoire de mes grands-parents.*

Remerciements

Un travail de thèse ne peut se mener seul, je tiens donc à remercier toutes les personnes sans qui je n'aurais pu prendre plaisir à l'effectuer. Ainsi ma gratitude s'adressera en premier à mon directeur de thèse, Eric Cancès, que j'ai d'abord apprécié durant ma scolarité pour ses grandes qualités de pédagogie comme professeur à l'Ecole Nationale des Ponts et Chaussées, puis qui m'a proposé ce travail de thèse et qui n'a ménagé ni son temps ni ses conseils tout au long de son élaboration. Son enthousiasme et son dynamisme furent un soutien précieux. Je remercie également Frédéric Legoll et Mathieu Lewin qui ont participé à mon encadrement, pour la gentillesse dont ils ont fait preuve à mon égard, j'ai par ailleurs beaucoup appris à leur contact.

Que les rapporteurs de ce mémoire, François Castella, Gianluca Panati et Eric Séré, trouvent ici l'expression de ma reconnaissance pour l'attention qu'ils ont bien voulu porter à ce manuscrit et pour les remarques judicieuses qu'ils ont formulées à son sujet. Je tiens aussi à exprimer ma gratitude à Xavier Blanc, Christian Brouder et Thierry Deutsch pour avoir accepté de faire partie du jury.

Ce travail de thèse fut effectué au sein du laboratoire de mathématiques appliquées de l'Ecole des Ponts, le Cermics. Je tiens à remercier son directeur, Serge Piperno, pour l'accueil qu'il m'a réservé, ainsi que pour l'instauration du rôle de représentant des doctorants, que j'eus la chance d'être la première à assumer ! Au sein du laboratoire mes remerciements vont d'abord à l'équipe de Simulation Moléculaire et Multi-Echelles, et principalement à Claude Le Bris, brillant guide de cette équipe, conjointement à Eric Cancès. Merci aussi à Andreea Grigoriu, Anthony Scemama, Antoine Gloria, Antonin Orriols, Arnaud Anantharaman, Gabriel Stoltz, Gabriel Turinici, Guy Bencteux, Ismaila Dabo, Kimiya Minoukadeh, Mathias Rousset, Raphaël Roux, Sébastien Boyaval, Tony Lelièvre et Yanli Li pour leur sympathie et l'ambiance conviviale qui règne dans l'équipe. Ma gratitude s'étend à tous les chercheurs, thésards et post-docs que j'ai pu connaître durant ces trois années, en particulier merci à Annette Stephansen et Pierre Sochala pour leur amitié et les pauses salutaires entre deux démonstrations de mathématiques, ainsi qu'à Pablo Tassi pour son café le matin, note chaleureuse avant une journée de labeur scientifique ... Je n'oublierai pas les secrétaires, Catherine Baccaert, Sylvie Berte, Khadija Elouali et Martine Ouhanna qui nous facilitent grandement toutes les tâches administratives, et ceux dont l'aide se révèle précieuse dès lors qu'un problème informatique se présente : Jean-Philippe Chancelier au Cermics ainsi que Jacques Daniel et Roland

Jarry à la DIT.

Au sein de l'ENPC et l'Université Paris Est, je remercie le Directeur du département des Formations Doctorales, Bernard Lapeyre, pour son écoute et ses conseils avisés au sujet de la thèse et l'après-thèse. Merci aussi à Alice Tran pour son aide - toujours avec un grand sourire ! - dans le cadre des formalités d'inscription et de fin de thèse.

Que la région Ile de France soit aussi remerciée, au travers de son représentant invité à mon jury de thèse, Romain Vidal, de la confiance qui m'a été témoignée lors de la sélection de ma candidature à la bourse Ile de France. Cette thèse constitue le fructueux aboutissement de ce parrainage. J'en profiterai pour remercier à nouveau le Cermics, ainsi que l'INRIA, qui prennent en charge les dépenses pour les conférences auxquelles nous participons.

Par ailleurs, j'adresse mes remerciements à ceux qui m'ont honorée de leur confiance en me donnant la chance de pouvoir enseigner au cours de mes années de thèse : Alexandre Ern et Michel de Lara à l'ENPC, Sophie Gaultier-Gaillard et Thierry Lafay à la Sorbonne et Alain Prignet à l'ESIEE. Qu'ils trouvent ici l'expression de ma reconnaissance.

Enfin je voudrais mentionner l'affection et le soutien de ma famille et mes amis durant toutes ces années d'étude, je ne les remercierai jamais assez et c'est à eux que je dédie en premier lieu ce travail.

Modélisation mathématique et simulation numérique de la structure électronique de cristaux en présence de défauts ponctuels

Nous présentons des résultats mathématiques obtenus pour un nouveau modèle de champ moyen dédié à la description d'électrons quantiques interagissant dans des cristaux comportant des défauts locaux. Ce modèle est dérivé du modèle dit de supercellule par un procédé de limite thermodynamique. Nous travaillons avec un modèle de Hartree-Fock réduit, obtenu à partir du modèle de Kohn-Sham étendu en négligeant le terme d'échange-corrélation. Les modèles utilisés et les résultats obtenus sont présentés au chapitre 2 puis démontrés au chapitre 4. Les chapitres 3 et 5 sont consacrés à la simulation numérique de notre modèle. Notre approche consiste à mettre en œuvre une approximation variationnelle dans une base précalculée de fonctions de Wannier du cristal parfait de référence. Nous présentons quelques résultats numériques obtenus sur un modèle unidimensionnel avec un potentiel d'interaction de Yukawa.

Mathematical modelling and numerical simulation of the electronic structure of crystals in presence of local defects

We present mathematical results obtained for a new mean-field model dedicated to the description of interacting electrons in crystals with local defects. Our model is derived from the so-called supercell model by a thermodynamic limit procedure. We work with a reduced Hartree-Fock model, obtained from the extended Kohn-Sham model by neglecting the exchange-correlation term. The models used and the obtained results are presented in chapter 2 and then proved in chapter 4. Chapters 3 and 5 deal with the numerical simulation of our model. Our approach consists of a variational approximation in a precomputed basis set of Wannier functions of the reference perfect crystal. Some numerical results obtained on a one-dimensional model with Yukawa interaction potential are presented.

Sommaire

1	Introduction générale	1
1.1	Présentation des résultats obtenus durant la thèse	1
1.2	Publications et communications orales	3
2	Introduction à la modélisation des cristaux en chimie quantique	5
2.1	Modélisation des systèmes à nombre fini d'électrons	5
2.1.1	Mise en place du modèle	5
2.1.2	Modèle de Hartree-Fock	8
2.1.3	Théorie de la fonctionnelle de la densité	9
2.2	Modélisation des cristaux parfaits	15
2.2.1	Réseaux et cellules unités	15
2.2.2	Rappels sur la théorie de Bloch	16
2.2.3	Modèle de Hartree-Fock réduit pour le cristal parfait	21
2.3	Modélisation des cristaux présentant des défauts ponctuels	26
2.3.1	Présentation du modèle de défaut	27
2.3.2	Existence de minimiseurs	29
2.4	Supercellule et limite thermodynamique	31
2.4.1	Description du modèle de supercellule	32
2.4.2	Limite thermodynamique pour le cristal parfait	35
2.4.3	Limite thermodynamique pour le cas avec défaut	36
3	Introduction à la simulation numérique des cristaux	39
3.1	Résolution numérique du problème de cellule	40
3.1.1	Calcul de la densité électronique fondamentale associée à un potentiel de champ moyen donné	42
3.1.2	Calcul du potentiel de champ moyen associé à une densité donnée	45
3.1.3	Optimal Damping Algorithm (ODA)	47
3.2	Résolution du problème de supercellule	49
3.2.1	Calcul de la densité électronique fondamentale associée à un potentiel de champ moyen donné	50
3.2.2	Calcul du potentiel de Yukawa associé à une densité donnée	51
3.2.3	Algorithme ODA pour le problème de supercellule	53
3.3	Résolution numérique du modèle de défaut	53

3.3.1	Fonctions de Wannier	53
3.3.2	Construction de MLWF en dimension 1	55
3.3.3	Approximation variationnelle du modèle de défaut	57
3.3.4	Quelques tests numériques	59
4	Local defects in periodic crystals	61
4.1	The reduced Hartree-Fock model for N electrons	63
4.2	The reduced Hartree-Fock model for a perfect crystal	64
4.3	The rHF model for a crystal with a defect	70
4.3.1	Definition of the reduced Hartree-Fock energy of a defect	71
4.3.2	Existence of minimizers with a chemical potential	82
4.3.3	Existence of minimizers under a charge constraint	94
4.4	Thermodynamic limit of the supercell model	109
4.4.1	Thermodynamic limit without defect	110
4.4.2	Thermodynamic limit with defect	119
4.5	Some useful lemmas	136
4.5.1	Commutator estimates	136
4.5.2	Density of finite rank operators in \mathcal{K}	138
5	Numerical simulation of crystals models	141
5.1	Derivation of the model	142
5.2	Proper definition of the variational set	145
5.3	Interpretation in terms of Bogoliubov states	147
5.4	Variational approximation	148
5.5	Numerical results	150
6	Annexes : Notions d'analyse fonctionnelle	153
6.1	Notions d'analyse fonctionnelle	153
	Bibliographie	165

Chapitre 1

Introduction générale

1.1 Présentation des résultats obtenus durant la thèse

Au cours de ce travail de thèse, nous avons abordé deux thèmes différents de la simulation moléculaire et multi-échelle. Tout d'abord, nous avons réalisé une étude théorique et numérique de modèles de structures électroniques en phase cristalline, d'abord dans le cadre de cristaux parfaits, puis de cristaux présentant des défauts ponctuels. La détermination de la structure électronique est en effet une étape nécessaire pour évaluer de façon fiable les forces interatomiques au sein du matériau. Les modèles sous-jacents sont des systèmes d'équations aux dérivées partielles non linéaires, - modèles de Kohn-Sham issus de la théorie de la fonctionnelle de la densité. Le modèle mathématique que nous proposons ici pour la description des cristaux présentant des défauts est nouveau. Il a fait l'objet d'une étude numérique unidimensionnelle et son implémentation dans le code Quantum-Espresso est en cours [78].

La seconde partie du travail de thèse consistait en une étude de calibration des modèles de Monte Carlo cinétique à l'aide de simulations de dynamique moléculaire. La connaissance des potentiels interatomiques permet de mener à bien des simulations de dynamique moléculaire sur des temps très courts - quelques nanosecondes. Or les processus mis en jeu dans le phénomène d'irradiation (par exemple sur les cuves des centrales nucléaires) correspondent pour certains à des événements rares (sauts de lacunes, formation d'amas de défauts cristallins, etc) qui se déroulent sur des échelles de temps beaucoup plus longues, de l'ordre de la seconde, et dont les conséquences macroscopiques (durcissement, fragilisation du matériau) ne deviennent visibles qu'à l'échelle de plusieurs années. Les méthodes de Monte Carlo cinétique permettent de réaliser l'intégration de ces échelles temporelles mais nécessitent, pour être efficaces, d'être paramétrées correctement par des simulations de dynamique moléculaire. L'objet du travail consistait à analyser de façon systématique pour un processus de diffusion de surface les différentes méthodes de calibration utilisées. Nous avons donc pour cela amélioré un code en C++ préexistant au CER-MICS. Les résultats de ces travaux sont en cours d'exploitation et ne seront donc pas présentés dans cette thèse.

Nous détaillons à présent la composition de ce manuscrit.

Dans le chapitre 2, nous présentons les différents types de modèles utilisés en chimie quantique, d'abord pour des systèmes comportant un nombre fini d'électrons, puis pour des cristaux parfaits. Nous introduisons ensuite les résultats théoriques obtenus durant la thèse. Ces résultats permettent en particulier de proposer un nouveau modèle de champ moyen décrivant la structure électronique d'un cristal en présence de défauts locaux. Ce modèle est fondé rigoureusement dans le cadre Hartree-Fock réduit par des arguments de limites thermodynamiques, et peut être étendu au cadre Kohn-Sham LDA et GGA. Pour construire ce modèle, nous reprenons dans un premier temps le procédé de limite thermodynamique pour un cristal parfait étudié par I. Catto, C. Le Bris et P.-L. Lions dans [11, 12], mais en adoptant une approche différente : dans notre étude, les électrons sont confinés à l'intérieur d'un cube (supercellule) dont on fait tendre la taille vers l'infini, et on impose des conditions aux bords périodiques, tandis que dans [11, 12], les électrons sont libres d'occuper l'espace entier. Nous prouvons en particulier un résultat d'unicité de l'état fondamental du cristal parfait dans le cadre Hartree-Fock réduit, résultat essentiel pour notre étude et qui n'était pas prouvé dans [11, 12]. Dans un second temps, en nous basant sur des idées de [13] déjà mises en œuvre par Ch. Hainzl, M. Lewin, E. Séré et J.P. Solovej [22–25] pour l'électrodynamique quantique, nous dérivons un modèle permettant de calculer l'énergie et la densité électronique de l'état fondamental en présence d'un défaut, modèle que nous justifions ensuite par des arguments de limite thermodynamique. Ce chapitre consiste principalement en des annonces de résultats qui seront démontrés ensuite au chapitre 4. Nous y montrons aussi quelques résultats préliminaires qui seront utilisés par la suite.

Le chapitre 3 est dédié à la mise en application numérique des résultats théoriques que nous avons obtenus. Cette mise en œuvre s'est effectuée au cours de la thèse dans un cadre unidimensionnel, mais les résultats théoriques seront utilisables dans un cadre plus général. Dans ce chapitre nous explicitons les approximations effectuées et l'algorithme d'optimisation utilisé (*Optimal Damping Algorithm (ODA)*), d'abord pour calculer l'état fondamental du cristal parfait, puis pour résoudre le problème de supercellule (avec ou sans défaut). Enfin, nous introduisons les bases de fonctions de Wannier qui nous ont paru les plus appropriées pour simuler numériquement notre modèle de défaut.

Les deux derniers chapitres sont plus approfondis et peuvent être lus indépendamment. Le chapitre 4 reprend, en le complétant par des démonstrations plus détaillées, un article que nous avons publié dans *Communications in Mathematical Physics*, présentant les fondements théoriques de notre modèle de défaut. Le chapitre 5, concernant la simulation numérique de ce modèle dans un cadre simplifié unidimensionnel, a été publié en l'état dans *Journal of Physics : Condensed Matter*.

Enfin, pour aider le lecteur, nous mettons en Annexe quelques définitions et résultats d'analyse fonctionnelle dont nous faisons un usage régulier dans ce manuscrit.

1.2 Publications et communications orales

Articles publiés

- E. Cancès, A. Deleurence & M. Lewin, *Non-perturbative embedding of local defects in crystalline materials*. J. Phys. : Condens. Matter **20** (2008), 294213.
- E. Cancès, A. Deleurence & M. Lewin, *A new approach to the modelling of local defects in crystals : the reduced Hartree-Fock case*. Comm. Math. Phys. **281** (2008), no. 1, 129–177.
- A. Deleurence, *Modélisation de cristaux périodiques avec ou sans défauts*, Femmes et Maths, n°8, ISSN 1271-3546, décembre 2006.

Communications orales

- *Modélisation de cristaux périodiques avec ou sans défaut*, CANUM, (Saint Jean de Monts, 26-30 Mai 2008)
- *Modelling of local defects in crystals*, Classical and Quantum Approaches in Molecular Modeling, (IMA, Minneapolis, 23 Juillet - 3 Août 2007)
- *Modélisation de cristaux périodiques avec ou sans défaut*, Forum des Jeunes Mathématiciennes, Mathématiques et Interactions, (Paris, 6-7 octobre 2006).

Chapitre 2

Modélisation des cristaux en chimie quantique

2.1 Modélisation des systèmes à nombre fini d'électrons

2.1.1 Mise en place du modèle

Le but principal des modèles *ab initio* en chimie quantique est de déterminer in fine l'état électronique fondamental, c'est-à-dire l'état électronique de plus basse énergie, d'un système moléculaire. Celui-ci est indispensable pour pouvoir ensuite étudier les propriétés physico-chimiques du système considéré.

Afin de simplifier l'écriture des équations, on travaille habituellement en unités atomiques, obtenues en imposant

$$m_e = 1, \quad e = 1, \quad \hbar = 1, \quad \frac{1}{4\pi\epsilon_0} = 1,$$

où m_e désigne la masse de l'électron, e la charge électrique élémentaire, \hbar la constante de Planck réduite et ϵ_0 la permittivité diélectrique du vide.

On a alors la correspondance suivante avec les unités usuelles

Unité de masse	9.11×10^{-31} kg
Unité de longueur (a_0 , rayon de Bohr)	5.29×10^{-11} m
Unité de temps	2.42×10^{-17} s
Unité d'énergie (le hartree, noté Ha)	4.36×10^{-18} J soit 627 kcal/mol.

Dans la majorité des cas, les modèles de chimie quantique se placent dans le cadre de l'*approximation de Born-Oppenheimer* : en simplifiant, cela revient à considérer les noyaux des atomes comme des particules classiques et les électrons comme des particules quantiques. Cette approximation est justifiée par le fait que les noyaux

sont beaucoup plus lourds que les électrons. Dans leur article de 1927 [5], Born et Oppenheimer ont résolu l'équation de Schrödinger stationnaire, que nous allons donner par la suite, par une méthode de perturbation en cherchant une solution sous la forme d'une série de puissances d'un paramètre ϵ , $\epsilon = \left(\frac{m_e}{m_n}\right)^{\frac{1}{4}}$, $m_e = 1$ étant la masse d'un électron et $m_n \gg 1$ la masse d'un noyau. Des travaux ultérieurs [29] ont montré que la série ainsi obtenue était asymptotique à la solution exacte de l'équation de Schrödinger à tous les ordres en ϵ .

Nous nous plaçons donc dans le cadre de cette approximation dans tout ce qui va suivre et considérons un système composé de

- K noyaux assimilés à des charges ponctuelles, situés aux positions R_1, \dots, R_K dans \mathbb{R}^3 , et de charges électriques z_1, \dots, z_K ;
- N électrons décrits, puisque ce sont des particules quantiques, par une fonction d'onde notée ψ_e .

Afin de simplifier le formalisme, nous allons "omettre" les variables de spin, les preuves mathématiques restant valables en prenant des modèles avec spin.

Sous l'approximation de Born-Oppenheimer, les particules classiques que sont les noyaux évoluent dans un potentiel effectif W et leur configuration d'équilibre correspond au minimum d'énergie potentielle

$$\inf \{W(R_1, \dots, R_K), (R_1, \dots, R_K) \in \mathbb{R}^{3K}\},$$

avec

$$\begin{aligned} W(R_1, \dots, R_K) &= U(R_1, \dots, R_K) + \sum_{1 \leq k < l \leq K} \frac{z_k z_l}{|R_k - R_l|}, \\ U(R_1, \dots, R_K) &= \inf \{ \langle \psi_e, H_e^{\{R_k\}} \psi_e \rangle, \psi_e \in \mathcal{H}_e, \|\psi_e\|_{L^2} = 1 \}, \\ H_e^{\{R_k\}} &= - \sum_{i=1}^N \frac{1}{2} \Delta_{x_i} - \sum_{i=1}^N \sum_{k=1}^K \frac{z_k}{|x_i - R_k|} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|}, \\ \mathcal{H}_e &= \bigwedge_{i=1}^N H^1(\mathbb{R}^3). \end{aligned}$$

Le premier terme du potentiel W représente le potentiel effectif perçu par les noyaux à cause de la présence d'un nuage d'électrons et le deuxième terme correspond à la répulsion internucléaire (forces de Coulomb).

Le potentiel U est obtenu en cherchant l'état fondamental (i.e. l'état stationnaire de plus basse énergie) de l'Hamiltonien électronique $H_e^{\{R_k\}}$ sur l'espace des fonctions d'onde admissibles, antisymétriques d'après le principe de Pauli et normalisées, puisque le module au carré de la fonction d'onde s'interprète comme une densité de probabilité. Notons que la difficulté réside dans le fait que U est lui-même le résultat d'un problème de minimisation sous contrainte ($\|\psi_e\|_{L^2} = 1$) sur un espace de dimension infinie (\mathcal{H}_e). Ce problème de minimisation, couramment appelé problème

électronique, est celui qui nous intéressera par la suite. Nous supposons donc dorénavant que les positions $\{R_k\}$ des noyaux sont fixées et posons $H_e := H_e^{\{R_k\}}$ pour simplifier les notations.

Détaillons à présent ce que représentent les différents termes de l'Hamiltonien H_e

- le premier terme représente l'énergie cinétique des électrons, Δ_{x_i} étant le laplacien par rapport à la variable $x_i \in \mathbb{R}^3$,
- le deuxième terme correspond à l'interaction coulombienne entre noyaux et électrons. Nous définissons au passage afin de simplifier les notations pour la suite,

$$V(x) = - \sum_{k=1}^K \frac{z_k}{|x - R_k|}$$

- enfin le troisième terme décrit la répulsion interélectronique coulombienne.

Précisons que la notation $\bigwedge_{i=1}^N H^1(\mathbb{R}^3)$ désigne l'espace vectoriel des fonctions de $\bigotimes_{i=1}^N H^1(\mathbb{R}^3)$ antisymétriques lorsque l'on échange les coordonnées de deux électrons.

Nous allons renommer l'Hamiltonien H_V afin d'insister sur le rôle que va jouer le potentiel V auquel sont soumis les électrons

$$H_V = H_1 + \sum_{i=1}^N V(x_i),$$

$$H_1 = - \sum_{i=1}^N \frac{1}{2} \Delta_{x_i} + \sum_{1 \leq i < j \leq N} \frac{1}{|x_i - x_j|},$$

Comme nous l'avons dit plus haut, un minimiseur du problème électronique est un vecteur propre de l'Hamiltonien H_V , associé à la plus petite valeur propre E de H_V (qui n'est autre que la valeur du minimum du problème électronique, puisque nous sommes dans un cadre linéaire). Ainsi, tout minimiseur du problème électronique vérifie l'équation de Schrödinger stationnaire

$$H_V \psi = E \psi.$$

On ne dispose d'une solution analytique à ce problème que pour le cas où $N = 1$ et on ne peut procéder à une résolution numérique directe que pour des systèmes à 1 ou 2 électrons. Pour des systèmes plus complexes, on utilise des méthodes de type Hartree-Fock ou fonctionnelle de la densité que nous allons présenter dans les deux prochaines sections.

Enfin à toute fonction d'onde $\psi_e \in \mathcal{H}_e$, nous pouvons associer pour tout p compris entre 1 et N un opérateur densité d'ordre p , noté $D_{\psi_e, p}$, de $\bigotimes_{i=1}^p L^2(\mathbb{R}^3)$ dans lui-même, défini par son noyau

$$\gamma_{\psi_e, p}(x_1, \dots, x_p, x'_1, \dots, x'_p) = \frac{N!}{p! (N-p)!} \int_{\mathbb{R}^{3(N-p)}} \psi_e(x_1, \dots, x_p, x_{p+1}, \dots, x_N) \psi_e(x'_1, \dots, x'_p, x_{p+1}, \dots, x_N) dx_{p+1} \cdots dx_N,$$

lui-même appelé matrice densité d'ordre p . Nous ne nous intéresserons par la suite qu'aux matrices densité d'ordre 1. Enfin on définit la densité électronique ρ_{ψ_e} associée à la fonction d'onde ψ_e par

$$\rho_{\psi_e}(x) = \gamma_{\psi_e,1}(x, x) = N \int_{\mathbb{R}^{3(N-1)}} |\psi_e(x, x_2, \dots, x_N)|^2 dx_2 \cdots dx_N. \quad (2.1)$$

2.1.2 Modèle de Hartree-Fock

Le modèle de Hartree-Fock résulte de l'approximation variationnelle du problème électronique, obtenue en restreignant les fonctions d'onde admissibles à celles qui s'écrivent sous la forme d'un déterminant de Slater de N fonctions d'onde monoélectroniques orthonormées φ_i appelées orbitales moléculaires :

$$\psi_e(x_1, \dots, x_N) = \frac{1}{\sqrt{N!}} \det(\varphi_i(x_j)) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(x_1) & \cdots & \varphi_1(x_N) \\ \vdots & & \vdots \\ \varphi_N(x_1) & \cdots & \varphi_N(x_N) \end{vmatrix}.$$

Introduisons l'espace des configurations de N orbitales moléculaires

$$\mathcal{W}_N = \left\{ \Phi = \{\varphi_i\}_{1 \leq i \leq N}, \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}, 1 \leq i, j \leq N \right\},$$

et l'ensemble des déterminants de Slater

$$\mathcal{S}_N = \left\{ \psi_e \in \mathcal{H}_e / \exists \Phi = \{\varphi_i\}_{1 \leq i \leq N} \in \mathcal{W}_N, \psi_e = \frac{1}{\sqrt{N!}} \det(\varphi_i(x_j)) \right\}.$$

Le problème de Hartree-Fock s'écrit

$$\inf \{ \langle \psi_e, H_V \psi_e \rangle, \psi_e \in \mathcal{S}_N \}. \quad (2.2)$$

Soit $\Phi = \{\varphi_i\}_{1 \leq i \leq N} \in \mathcal{W}_N$ et $\psi_e \in \mathcal{S}_N$ le déterminant de Slater associé à Φ . Soit $\gamma_\Phi(x, x') = \gamma_{\psi_e,1}(x, x')$, $D_\Phi = D_{\psi_e,1}$ et $\rho_\Phi(x) = \rho_{\psi_e}(x)$. On a alors

$$\begin{aligned} \gamma_\Phi(x, x') &= \sum_{i=1}^N \varphi_i(x) \varphi_i(x'), \\ D_\Phi &= \sum_{i=1}^N (\varphi_i, \cdot)_{L^2} \varphi_i, \\ \rho_\Phi(x) &= \sum_{i=1}^N |\varphi_i(x)|^2. \end{aligned}$$

Nous pouvons donc à présent définir la fonctionnelle d'énergie du modèle de Hartree-Fock, $E^{HF}(\Phi) := \langle \psi_e, H_V \psi_e \rangle$ qui s'exprime de la façon suivante en fonction des φ_i

$$\begin{aligned} E^{HF}(\Phi) &= \sum_{i=1}^N \frac{1}{2} \int_{\mathbb{R}^3} |\nabla \varphi_i|^2 + \int_{\mathbb{R}^3} \rho_\Phi V \\ &\quad + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\Phi(x) \rho_\Phi(x')}{|x - x'|} dx dx' \\ &\quad - \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{|\gamma_\Phi(x, x')|^2}{|x - x'|} dx dx'. \end{aligned}$$

Physiquement, le premier terme de l'énergie de Hartree-Fock représente l'énergie cinétique de la fonction d'onde, et le deuxième l'interaction électrostatique entre noyaux et électrons; le troisième terme correspond à l'énergie coulombienne classique de la distribution de charge ρ_Φ ; le dernier est appelé le terme d'échange. Il est d'origine purement quantique puisqu'il résulte de l'antisymétrie de la fonction d'onde. Le problème de Hartree-Fock (2.2) s'écrit donc aussi

$$\inf \{ E^{HF}(\Phi), \Phi \in \mathcal{W}_N \}. \quad (2.3)$$

Notons dès à présent que la difficulté à résoudre ce nouveau problème provient du fait que la fonctionnelle d'énergie n'est pas quadratique.

Par ailleurs, puisque nous minimisons la fonctionnelle d'énergie $\langle \psi_e, H_V \psi_e \rangle$ sur un ensemble plus restreint que l'ensemble de départ, l'énergie de Hartree-Fock sera supérieure à l'énergie électronique du système. La différence de ces deux énergies est dénommée énergie de corrélation.

Enfin pour les modèles de type Hartree-Fock, l'existence du minimiseur pour les potentiels de type Coulombien est démontrée dans le cas où $Z = \sum_{i=1}^K z_k \geq N$ [15, 42, 49]. En revanche son unicité reste un problème ouvert et difficile.

2.1.3 Théorie de la fonctionnelle de la densité

Les méthodes provenant de la théorie de la fonctionnelle de la densité (DFT, pour *Density Functional Theory*) déjà très utilisées en physique du solide, prennent aussi depuis quelques années une part essentielle dans les simulations de chimie moléculaire. Il s'agit de rechercher l'énergie et la densité électronique du fondamental du problème électronique, en résolvant un problème de minimisation de la forme

$$\inf \left\{ F(\rho) + \int_{\mathbb{R}^3} \rho V, \rho \in L^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \rho = N \right\},$$

où F est une fonctionnelle de la densité électronique ρ .

La justification théorique de ces modèles a d'abord été faite par Hohenberg et Kohn [27] puis une autre approche fut proposée par Levy et ensuite reprise et prolongée par Lieb [33].

Le problème des méthodes de type DFT est qu'on ne connaît pas d'expression exacte de la fonctionnelle $F(\rho)$, qui pourrait se prêter à des simulations numériques. On en construit donc des approximations en prenant pour référence des systèmes pour lesquels on sait construire une fonctionnelle de la densité :

- pour les modèles de type Thomas-Fermi, le système de référence est un gaz homogène d'électrons,
- quant aux modèles de type Kohn-Sham ils utilisent un système de N électrons *sans interaction*.

2.1.3.1 Modèles de type Thomas-Fermi

Ces modèles sont obtenus par transposition au cas moléculaire du comportement d'un gaz homogène d'électrons et certains ont été introduits bien avant la justification théorique de la DFT par Hohenberg et Kohn. Citons les trois principaux :

- le modèle de Thomas-Fermi (TF)

$$F(\rho) = C_{TF} \int_{\mathbb{R}^3} \rho^{\frac{5}{3}} + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x) \rho(y)}{|x-y|} dx dy,$$

- le modèle de Thomas-Fermi-von Weizsäcker (TFW)

$$F(\rho) = C_W \int_{\mathbb{R}^3} |\nabla \sqrt{\rho}|^2 + C_{TF} \int_{\mathbb{R}^3} \rho^{\frac{5}{3}} + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x) \rho(y)}{|x-y|} dx dy,$$

- et le modèle de Thomas-Fermi-Dirac-von Weizsäcker (TFDW)

$$F(\rho) = C_W \int_{\mathbb{R}^3} |\nabla \sqrt{\rho}|^2 + C_{TF} \int_{\mathbb{R}^3} \rho^{\frac{5}{3}} - C_D \int_{\mathbb{R}^3} \rho^{\frac{4}{3}} + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x) \rho(y)}{|x-y|} dx dy.$$

Ces modèles, bien que considérés comme rudimentaires et ayant cessé d'être utilisés dans les simulations, restent intéressants d'un point de vue théorique car ils présentent des difficultés mathématiques semblables aux modèles de type Hartree-Fock et Kohn-Sham : minimisation sous contrainte, perte de compacité à l'infini, non-convexité (pour TFDW), présence de potentiels coulombiens, etc. Ils ont été étudiés entre autres par Lieb [32] et Simon [37], Benguria, Brezis et Lieb [2], Catto, Lions [8] et Le Bris [9, 10].

2.1.3.2 Modèles de type Kohn-Sham et extended Kohn-Sham

Pour présenter ces modèles nous allons avoir besoin de retourner aux fondements de la théorie de la fonctionnelle de la densité.

Nous revenons au problème de minimisation

$$E(V) := \inf \left\{ \langle \psi_e, H_V \psi_e \rangle, \psi_e \in \bigwedge_{i=1}^N H^1(\mathbb{R}^3), \|\psi_e\|_{L^2} = 1 \right\}$$

et notons $\mathcal{F}_N = \left\{ \psi_e \in \bigwedge_{i=1}^N H^1(\mathbb{R}^3), \|\psi_e\|_{L^2} = 1 \right\}$. On peut montrer que si $\psi_e \in \mathcal{F}_N$, alors la densité ρ_{ψ_e} définie par (2.1) appartient à l'ensemble convexe

$$\mathcal{I}_N = \left\{ \rho \geq 0, \sqrt{\rho} \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \rho = N \right\}.$$

Réciproquement, si $\rho \in \mathcal{I}_N$, alors il existe $\psi_e \in \mathcal{F}_N$ tel que $\rho_{\psi_e} = \rho$. On dit que les densités de \mathcal{I}_N sont *N-représentables* par des fonctions d'onde d'énergie finie. Une idée naturelle pour construire une fonctionnelle de la densité (dont l'existence découle du théorème de Hohenberg et Kohn [27]) est alors de poser

$$F_{LL}(\rho) = \inf \{ \langle \psi_e, H_1 \psi_e \rangle, \psi_e \in \mathcal{F}_N, \psi_e \text{ a pour densité } \rho \}.$$

et de remarquer que

$$\begin{aligned} E(V) &= \inf \{ \langle \psi_e, H_V \psi_e \rangle, \psi_e \in \mathcal{F}_N \}, \\ &= \inf \left\{ \inf \{ \langle \psi_e, H_1 \psi_e \rangle, \psi_e \in \mathcal{F}_N, \psi_e \text{ a pour densité } \rho \} + \int_{\mathbb{R}^3} \rho V, \rho \in \mathcal{I}_N \right\}, \\ &= \inf \left\{ F_{LL}(\rho) + \int_{\mathbb{R}^3} \rho V, \rho \in \mathcal{I}_N \right\}. \end{aligned} \quad (2.4)$$

La fonctionnelle F_{LL} s'appelle *fonctionnelle de la densité de Levy-Lieb*.

On peut aussi définir une autre fonctionnelle bénéficiant de meilleures propriétés mathématiques [33] mais construite à partir d'états mélangés plutôt que purs. L'opérateur de densité à N corps associé à l'état pur correspondant à la fonction d'onde $\psi_e \in \mathcal{F}_N$ est défini par

$$\Gamma_N = |\psi_e\rangle\langle\psi_e|.$$

Les états mélangés sont définis par des opérateurs densité à N corps d'états mixtes, qui sont des combinaisons convexes d'opérateurs de densité à N corps d'états purs

$$\Gamma = \sum_{i=1}^{+\infty} p_i |\psi_e^i\rangle\langle\psi_e^i|, \quad 0 \leq p_i \leq 1, \quad \psi_e^i \in \mathcal{F}_N.$$

Soit \mathcal{D}_N l'ensemble de ces opérateurs. C'est un ensemble convexe - il s'agit en fait de l'enveloppe convexe de l'ensemble des opérateurs densité des états purs. La densité $\rho(x)$ et la matrice densité d'ordre 1 $\gamma(x, x')$ associées à un tel opérateur s'écrivent

$$\rho(x) = \sum_{i=1}^{+\infty} p_i \rho^i(x),$$

et

$$\gamma(x, x') = \sum_{i=1}^{+\infty} p_i \gamma^i(x, x'),$$

où ρ^i et γ^i sont la densité et la matrice densité associées à la fonction d'onde ψ_e^i . Par linéarité, on peut écrire

$$\mathrm{Tr}(\Gamma) = 1, \quad \mathrm{Tr}(H_1\Gamma) = \sum_{i=1}^{+\infty} p_i \langle \psi_e^i, H_1 \psi_e^i \rangle, \quad \mathrm{Tr}(H_V\Gamma) = \sum_{i=1}^{+\infty} p_i \langle \psi_e^i, H_1 \psi_e^i \rangle + \int_{\mathbb{R}^3} \rho V.$$

Il est alors facile de voir que

$$E(V) = \inf \{ \mathrm{Tr}(H_V\Gamma), \Gamma \in \mathcal{D}_N \},$$

et que

$$\{ \rho, \exists \Gamma \in \mathcal{D}_N \text{ tel que } \Gamma \text{ ait pour densité } \rho \} = \mathcal{I}_N.$$

Par conséquent,

$$E(V) = \inf \left\{ F_L(\rho) + \int_{\mathbb{R}^3} \rho V, \rho \in \mathcal{I}_N \right\} \quad (2.5)$$

où $F_L(\rho)$ est la *fonctionnelle de la densité de Lieb*, définie par

$$F_L(\rho) = \inf \{ \mathrm{Tr}(H_1\Gamma), \Gamma \in \mathcal{D}_N, \Gamma \text{ a pour densité } \rho \}.$$

Jusqu'à présent aucune approximation n'a été effectuée mais aucune des deux fonctionnelles F_{LL} et F_L , ni la fonctionnelle de Hohenberg-Kohn [27] dont elles découlent, ne peuvent être évaluées facilement pour un système de N électrons. Des approximations doivent donc être faites pour que la théorie de la fonctionnelle de la densité ait un intérêt pratique. Actuellement les approximations faites reposent sur des calculs pour des systèmes de référence relativement proches du système "idéal" réel. La contribution la plus connue est celle de Kohn et Sham [30] qui ont donné l'idée de prendre un système de référence où les N électrons n'interagissent pas entre eux. La fonctionnelle de densité est alors obtenue en remplaçant l'Hamiltonien H_1 du système réel par

$$H_0 = - \sum_{i=1}^N \frac{1}{2} \Delta_{x_i}.$$

La fonctionnelle analogue à la fonctionnelle de densité de Levy-Lieb est alors la fonctionnelle d'énergie cinétique de Kohn-Sham

$$T_{KS}(\rho) = \inf \{ \langle \psi_e, H_0 \psi_e \rangle, \psi_e \in \mathcal{F}_N, \psi_e \text{ a pour densité } \rho \}. \quad (2.6)$$

Ceci peut se simplifier en

$$T_{KS}(\rho) = \inf \left\{ \frac{1}{2} \sum_{i=1}^N \int_{\mathbb{R}^3} |\nabla \varphi_i|^2, \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}, \sum_{i=1}^N |\varphi_i|^2 = \rho \right\}, \quad (2.7)$$

pourvu que (2.6) admette pour minimiseur un déterminant de Slater. Le problème est que ce n'est pas toujours le cas (il y a des contre-exemples). Cependant, la

formule (2.7) a un sens pour tout $\rho \in \mathcal{I}_N$ et c'est cette formule qui est utilisée en général dans les simulations numériques.

Ce problème n'a plus lieu pour la fonctionnelle de la densité de Lieb associée à l'Hamiltonien H_0 ne prenant pas en compte l'interaction des électrons. Définissons la *fonctionnelle d'énergie cinétique de Janak*

$$T_J(\rho) = \inf \{ \text{Tr}(H_0\Gamma), \Gamma \in \mathcal{D}_N, \Gamma \text{ a pour densité } \rho \}.$$

Remarquons d'abord que

- pour tout opérateur densité d'états mixtes Γ on a

$$\text{Tr}(H_0\Gamma) = \int_{\mathbb{R}^3} \left(\frac{1}{2} \Delta_x \gamma(x, x') \right) \Big|_{x=x'} dx$$

où γ est la matrice densité d'ordre 1 associée à Γ ;

- l'ensemble des matrices densité d'ordre 1 associées à des opérateurs densité d'états mixtes est connu : il s'agit de

$$\left\{ \gamma(x, x') = \sum_{i=1}^{+\infty} n_i \varphi_i(x) \varphi_i(x'), \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}, 0 \leq n_i \leq 1, \sum_{i=1}^{\infty} n_i = N \right\}.$$

Remarque 1. *Ce résultat est connu sous le nom de théorème de représentabilité des matrices densité de premier ordre d'états mixtes [17, 33], et il n'en existe pas de similaire pour les opérateurs densité des états purs.*

Ceci nous amène à redéfinir la fonctionnelle d'énergie cinétique de Janak, qui a un sens pour tout $\rho \in \mathcal{I}_N$

$$T_J(\rho) = \inf \left\{ \frac{1}{2} \sum_{i=1}^N n_i \int_{\mathbb{R}^3} |\nabla \varphi_i|^2, \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}, \right. \\ \left. 0 \leq n_i \leq 1, \sum_{i=1}^{+\infty} n_i = N, \sum_{i=1}^{\infty} n_i |\varphi_i|^2 = \rho \right\}.$$

Les fonctionnelles T_{KS} et T_J associées à l'Hamiltonien H_0 fournissent des approximations assez bonnes de l'énergie cinétique du système réel. Par ailleurs, l'énergie de Coulomb

$$J(\rho) = \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x)\rho(y)}{|x-y|} dx dy$$

représentant l'énergie électrostatique d'une distribution de charge classique de densité ρ est une estimation raisonnable de l'énergie d'interaction entre les N électrons d'un système de densité ρ . Les erreurs dues à chacun de ces deux termes sont regroupées dans l'*énergie d'échange-corrélation* définie comme la différence

$$E_{xc}(\rho) = F_{LL}(\rho) - T_{KS}(\rho) - J(\rho),$$

ou

$$E_{xc}(\rho) = F_L(\rho) - T_J(\rho) - J(\rho),$$

selon que l'on a choisi de suivre l'approche Levy-Lieb ou l'approche Lieb. La construction de fonctionnelles d'échange-corrélation approchées rendant les simulations numériques possibles est basée sur l'analyse du gaz uniforme d'électron (fonctionnelles LDA) ou sur une analyse du trou d'échange-corrélation (fonctionnelles GGA). Une partie de l'échange Hartree-Fock peut aussi être inclus et on parle alors de fonctionnelles hybrides.

Lorsque la fonctionnelle d'énergie cinétique de Kohn-Sham est choisie, le problème de minimisation abstrait (2.4) peut être remplacé par le *modèle de Kohn-Sham*

$$E^{KS}(V) = \inf \left\{ \frac{1}{2} \sum_{i=1}^N \int_{\mathbb{R}^3} |\nabla \varphi_i|^2 + \int_{\mathbb{R}^3} \rho V + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x) \rho(y)}{|x-y|} dx dy + E_{xc}(\rho), \right. \\ \left. \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij} \right\},$$

avec ici $\rho = \sum_{i=1}^N |\varphi_i|^2$.

Si l'on utilise la fonctionnelle d'énergie cinétique de Janak pour reformuler le problème (2.5), on obtient le modèle "*extended Kohn-Sham*"

$$E^{EKS}(V) = \inf \left\{ \frac{1}{2} \sum_{i=1}^{+\infty} n_i \int_{\mathbb{R}^3} |\nabla \varphi_i|^2 + \int_{\mathbb{R}^3} \rho V + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x) \rho(y)}{|x-y|} dx dy + E_{xc}(\rho), \right. \\ \left. \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}, 0 \leq n_i \leq 1, \sum_{i=1}^{+\infty} n_i = N, \sum_{i=1}^{+\infty} n_i \|\nabla \varphi_i\|_{L^2}^2 < \infty \right\},$$

où cette fois

$$\rho = \sum_{i=1}^{+\infty} n_i |\varphi_i|^2. \quad (2.8)$$

C'est du modèle extended Kohn-Sham que découle le modèle *reduced Hartree-Fock* (rHF) puisque celui-ci consiste à considérer une fonctionnelle d'énergie de type extended Kohn-Sham dans laquelle le terme d'échange-corrélation est nul. Ce modèle s'écrit donc

$$E^{rHF}(V) = \inf \left\{ \frac{1}{2} \sum_{i=1}^{+\infty} n_i \int_{\mathbb{R}^3} |\nabla \varphi_i|^2 + \int_{\mathbb{R}^3} \rho V + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho(x) \rho(y)}{|x-y|} dx dy, \right. \\ \left. \varphi_i \in H^1(\mathbb{R}^3), \int_{\mathbb{R}^3} \varphi_i \varphi_j = \delta_{ij}, 0 \leq n_i \leq 1, \sum_{i=1}^{+\infty} n_i = N, \sum_{i=1}^{+\infty} n_i \|\nabla \varphi_i\|_{L^2}^2 < \infty \right\},$$

où ρ est donné par (2.8). On peut simplifier l'écriture du modèle ci-dessus en le reformulant à l'aide des matrices densité d'ordre 1. On obtient ainsi

$$E^{rHF}(V) = \inf \left\{ \mathcal{E}^{rHF}(\gamma), \gamma \in \mathcal{P}_N \right\} \quad (2.9)$$

où

$$\mathcal{E}^{rHF}(\gamma) = \int_{\mathbb{R}^3} \text{Tr}(-\frac{1}{2}\Delta\gamma) + \int_{\mathbb{R}^3} \rho_\gamma V + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_\gamma(x) \rho_\gamma(y)}{|x-y|} dx dy,$$

avec $\rho_\gamma(x) = \gamma(x, x)$, et où

$$\mathcal{P}_N = \left\{ \gamma \in \mathcal{S}(L^2(\mathbb{R}^3)), 0 \leq \gamma \leq 1, \text{Tr}(\gamma) = N, \text{Tr}(-\Delta\gamma) < \infty \right\}.$$

C'est le modèle (2.9) qui servira de base à tout ce qui va suivre.

2.2 Modélisation des cristaux parfaits

Nous commençons ici par présenter les outils et concepts généraux utilisés dans le cadre de la modélisation des cristaux parfaits, avant de présenter les résultats relatifs aux modèles de cristaux présentant des défauts, qui seront exposés et démontrés au Chapitre 4.

2.2.1 Réseaux et cellules unités

Un cristal parfait est un système moléculaire de taille infinie, constitué d'un ensemble de noyaux dont la charge est représentée par une mesure de Radon positive périodique, et d'une infinité d'électrons donnant lieu à une densité électronique périodique (de même périodicité que la distribution de charge des noyaux). En supposant que les noyaux sont fixés (approximation de Born-Oppenheimer), nous pouvons décrire leurs positions par un motif m répété sur les noeuds d'un réseau \mathcal{R} :

$$\mathcal{R} = \left\{ ua + vb + wc, \quad (u, v, w) \in \mathbb{Z}^3 \right\}.$$

Les vecteurs a, b, c étant supposés linéairement indépendants dans \mathbb{R}^3 , ils forment une base de \mathcal{R} . Pour revenir à l'approximation classique des noyaux, le motif m représente K particules ponctuelles de charge positive qui correspondent mathématiquement parlant à un ensemble de masses de Dirac positives comme nous l'avons vu auparavant :

$$m = \sum_{k=1}^K z_k \delta_{R_k},$$

où les R_k représentent les positions des noyaux et peuvent être choisis dans l'ensemble

$$\Gamma = \left\{ xa + yb + zc, \quad (x, y, z) \in \left[-\frac{1}{2}, \frac{1}{2} \right]^3 \right\},$$

appelé cellule unité. Plusieurs choix de cellules unité sont bien sûr possibles. En physique, on utilise souvent la cellule de Wigner-Seitz, ou zone de Dirichlet, définie comme étant la cellule de Voronoï associée au point $\{0\}$ de la distribution de points \mathcal{R} . La cellule de Wigner-Seitz a la propriété de partager les symétries du réseau. Notons que lorsque les vecteurs a , b et c sont deux à deux orthogonaux, Γ coïncide avec la cellule de Wigner-Seitz du réseau.

Notre objectif est de décrire la structure électronique du cristal, la distribution périodique des noyaux étant donnée. Pour cela, nous aurons à manipuler de nombreuses fonctions \mathcal{R} -périodiques, pouvant être traitées grâce à l'analyse de Fourier. C'est pourquoi nous introduisons à présent la notion de *réseau réciproque*. Pour une base donnée (a, b, c) de \mathcal{R} , on définit :

$$\mathcal{R}^* = \{ua' + vb' + wc', \quad (u, v, w) \in \mathbb{Z}^3\},$$

où la base (a', b', c') vérifie

$$\begin{aligned} a \cdot a' &= b \cdot b' = c \cdot c' = 2\pi, \\ b \cdot a' &= c \cdot a' = a \cdot b' = c \cdot b' = a \cdot c' = b \cdot c' = 0. \end{aligned}$$

Cette définition de \mathcal{R}^* ne dépend pas de la base choisie. La cellule de Wigner-Seitz du réseau réciproque est appelée première *zone de Brillouin* du cristal. Nous la notons Γ^* .

Nous supposons dans toute la suite que $\mathcal{R} = \mathbb{Z}^3$ (l'extension à un réseau quelconque ne pose pas de difficulté). On aura donc $\Gamma = [-\frac{1}{2}, \frac{1}{2})^3$, $\mathcal{R}^* = 2\pi\mathbb{Z}^3$ et $\Gamma^* = [-\pi, \pi)^3$.

2.2.2 Rappels sur la théorie de Bloch

Les ondes de Bloch ont d'abord été introduites par Floquet [19] dans le cas unidimensionnel et par Bloch [4] dans le cas général.

Une onde de Bloch, ou fonction quasi-périodique, est une fonction appartenant à l'espace

$$L_\xi^2(\Gamma) = \{\varphi \in L_{\text{loc}}^2(\mathbb{R}^3) \mid \varphi(x - k) = e^{-ik \cdot \xi} \varphi(x), \quad \forall k \in \mathbb{Z}^3, \text{ pour presque tout } x \in \Gamma\},$$

où $L_{\text{loc}}^2(\mathbb{R}^3)$ désigne l'espace des fonctions mesurables de \mathbb{R}^3 à valeurs dans \mathbb{C} qui sont de carré intégrable sur tout compact inclus dans \mathbb{R}^3 . De manière équivalente,

$$L_\xi^2(\Gamma) = \{\varphi \in L_{\text{loc}}^2(\mathbb{R}^3) \mid e^{-i\xi \cdot \cdot} \varphi \in L_{\text{per}}^2(\Gamma)\},$$

où

$$L_{\text{per}}^2(\Gamma) = \{\varphi \in L_{\text{loc}}^2(\mathbb{R}^3) \mid \varphi \text{ est } \mathbb{Z}^3 - \text{périodique}\}.$$

Ainsi, si l'on dote $L_\xi^2(\Gamma)$ du produit scalaire

$$(\varphi, \psi)_{L_\xi^2(\Gamma)} = \int_\Gamma \overline{\varphi} \psi,$$

on obtient un espace de Hilbert isomorphe à $L^2_{\text{per}}(\Gamma)$.

Considérons maintenant l'espace

$$\frac{1}{(2\pi)^3} \int_{\Gamma^*} L^2_{\xi}(\Gamma) d\xi = \left\{ (\varphi_{\xi})_{\xi \in \Gamma^*} \mid \varphi_{\xi} \in L^2_{\xi}(\Gamma), \xi \mapsto \|\varphi_{\xi}\|_{L^2_{\xi}(\Gamma)} \text{ mesurable,} \right. \\ \left. \int_{\Gamma^*} \|\varphi_{\xi}\|_{L^2_{\xi}(\Gamma)}^2 < \infty \right\}$$

qui, muni du produit scalaire défini par

$$(\varphi, \psi) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (\varphi_{\xi}, \psi_{\xi})_{L^2_{\xi}(\Gamma)} d\xi,$$

a une structure d'espace de Hilbert.

Le théorème de décomposition de Bloch assure que l'application linéaire définie sur la classe de Schwartz $\mathcal{S}(\mathbb{R}^3)$ par

$$\forall \varphi \in \mathcal{S}(\mathbb{R}^3), \quad U\varphi = ((U\varphi)_{\xi})_{\xi \in \Gamma^*} \quad \text{avec} \quad (U\varphi)_{\xi}(x) = \sum_{k \in \mathbb{Z}^3} e^{ik \cdot \xi} \varphi(x - k),$$

se prolonge de manière unique en une isométrie U de $L^2(\mathbb{R}^3)$ dans $\frac{1}{(2\pi)^3} \int_{\Gamma^*} L^2_{\xi}(\Gamma) d\xi$.

L'inverse de U , notée U^* , est définie pour tout $\varphi = (\varphi_{\xi})_{\xi \in \Gamma^*}$ par

$$(U^*\varphi)(x - k) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} e^{-ik \cdot \xi} \varphi_{\xi}(x) d\xi,$$

pour presque tout $x \in \Gamma$ et pour tout $k \in \mathbb{Z}^3$. L'isométrie U permet donc d'écrire que pour tout $\varphi \in L^2(\mathbb{R}^3)$,

$$\varphi(x) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \varphi_{\xi}(x) d\xi$$

où

$$\varphi_{\xi}(x) = \sum_{k \in \mathbb{Z}^3} e^{ik \cdot \xi} \varphi(x - k) \in L^2_{\xi}(\Gamma)$$

et

$$\|\varphi\|_{L^2(\mathbb{R}^3)}^2 = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \|\varphi_{\xi}\|_{L^2_{\xi}(\Gamma)}^2 d\xi.$$

Soit maintenant un opérateur auto-adjoint $A \in \mathcal{B}(L^2(\mathbb{R}^3))$ qui commute avec les translations du réseau périodique, c'est-à-dire tel que

$$\forall k \in \mathbb{Z}^3, \quad \tau_k A = A \tau_k.$$

A la décomposition ci-dessus des fonctions de $L^2(\mathbb{R}^3)$ en ondes de Bloch correspond la décomposition suivante de l'opérateur A : il existe une unique fonction $\xi \mapsto A_{\xi}$

dans $L^\infty(\Gamma^*, \mathcal{B}(L_\xi^2(\Gamma)))$ telle que pour toute fonction $\varphi \in L^2(\mathbb{R}^3)$ et presque tout $\xi \in \Gamma^*$

$$(UA\varphi)_\xi = A_\xi(U\varphi)_\xi, \quad (2.10)$$

soit de façon équivalente

$$(A\varphi)(x) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (A\varphi)_\xi(x) d\xi = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (A_\xi\varphi_\xi)(x) d\xi.$$

On écrit de manière symbolique

$$A = \frac{1}{(2\pi)^3} \int_{\Gamma^*} A_\xi d\xi.$$

L'égalité (2.10) implique en particulier que

$$\sup_{\xi \in \Gamma^*} \|A_\xi\|_{\mathcal{B}(L_\xi^2(\Gamma))} = \|A\|_{\mathcal{B}(L^2(\mathbb{R}^3))}.$$

Si de plus A est positif sur $L^2(\mathbb{R}^3)$, chacun des A_ξ est positif. On peut alors définir dans $\mathbb{R}_+ \cup \{+\infty\}$

$$\mathrm{Tr}_\Gamma(A) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \mathrm{Tr}_{L_\xi^2(\Gamma)}(A_\xi) d\xi.$$

Dans la modélisation de la structure électronique des cristaux parfaits, on est amené à manipuler des opérateurs (ou matrices) densité $\gamma \in L^2(\mathbb{R}^3)$ vérifiant les propriétés suivantes :

- γ commute avec les translations du réseau

$$\forall k \in \mathbb{Z}^3, \quad \tau_k \gamma = \gamma \tau_k,$$

- $0 \leq \gamma \leq 1$ au sens des opérateurs auto-adjoints sur $L^2(\mathbb{R}^3)$, 1 représentant l'opérateur identité sur $L^2(\mathbb{R}^3)$,
- le nombre d'électrons par unité de volume est fini,
- l'énergie par unité de volume de γ est finie.

Il résulte de ce qui précède qu'un tel opérateur γ peut se décomposer sous la forme

$$\gamma = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \gamma_\xi d\xi,$$

avec $0 \leq \gamma_\xi \leq 1$ pour presque tout $\xi \in \Gamma^*$, 1 désignant cette fois-ci l'opérateur identité sur $L_\xi^2(\Gamma)$. Le nombre d'électrons par cellule unité est égal à $\mathrm{Tr}_\Gamma(\gamma)$. Notons que

$$\mathrm{Tr}_\Gamma(\gamma) = \int_\Gamma \rho_\gamma$$

où

$$\rho_\gamma(x) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \gamma_\xi(x, x) d\xi$$

est la densité électronique (c'est une fonction positive et \mathbb{Z}^3 -périodique). La condition sur la finitude de l'énergie par cellule unité s'exprime par

$$\mathrm{Tr}_\Gamma(-\Delta\gamma) := \frac{1}{(2\pi)^3} \int_{\Gamma^*} \mathrm{Tr}_{L_\xi^2(\Gamma)}(-\Delta_\xi\gamma_\xi) d\xi < \infty,$$

où Δ_ξ est l'opérateur de Laplace, considéré comme un opérateur $L_\xi^2(\Gamma)$. Dans l'expression ci-dessus, on a utilisé la convention usuelle

$$\mathrm{Tr}_\Gamma(-\Delta\gamma) := \mathrm{Tr}_\Gamma((-\Delta)^{-1/2}\gamma(-\Delta)^{-1/2})$$

et

$$\mathrm{Tr}_{L_\xi^2(\Gamma)}(-\Delta_\xi\gamma_\xi) := \mathrm{Tr}_{L_\xi^2(\Gamma)}((-\Delta_\xi)^{1/2}\gamma_\xi(-\Delta_\xi)^{1/2}).$$

Pour toute matrice densité d'énergie finie, on a en fait

$$\sqrt{\rho_\gamma} \in H_{\mathrm{per}}^1(\Gamma) = \{u \in L_{\mathrm{per}}^2(\Gamma) \mid \nabla u \in (L_{\mathrm{per}}^2(\Gamma))^3\}.$$

Ceci est une conséquence de l'inégalité

$$\int_\Gamma |\nabla \sqrt{\gamma}|^2 \leq \mathrm{Tr}_\Gamma(-\Delta\gamma).$$

Dans certains problèmes variationnels que nous considérerons, le nombre d'électrons Z par cellule unité sera une inconnue du problème. On travaillera alors dans l'ensemble

$$\mathcal{P}_{\mathrm{per}} = \left\{ \gamma \in \mathcal{B}(L^2(\mathbb{R}^3)) \mid 0 \leq \gamma \leq 1, \forall k \in \mathbb{Z}^3, \tau_k \gamma = \gamma \tau_k, \int_{\Gamma^*} \mathrm{Tr}_{L_\xi^2(\Gamma)}((1 - \Delta_\xi)^{1/2}\gamma_\xi(1 - \Delta_\xi)^{1/2}) d\xi < \infty \right\}.$$

Dans d'autres cas, il sera fixé à l'avance, et l'ensemble variationnel sera

$$\mathcal{P}_{\mathrm{per}}^Z = \left\{ \gamma \in \mathcal{P}_{\mathrm{per}} \mid \mathrm{Tr}_\Gamma(\gamma) = \int_\Gamma \rho_\gamma(x) dx = Z \right\}.$$

Dans les modèles de type Kohn-Sham, l'Hamiltonien électronique est un opérateur sur $L^2(\mathbb{R}^3)$ de la forme

$$H = -\frac{1}{2}\Delta + V_{\mathrm{per}} \tag{2.11}$$

où $V_{\mathrm{per}} \in L_{\mathrm{per}}^2(\Gamma)$ est une fonction (un potentiel effectif) \mathbb{Z}^3 -périodique. Un tel opérateur est appelé un opérateur de Schrödinger périodique. L'opérateur H est non-borné, mais on peut néanmoins le décomposer en ondes de Bloch :

$$H = \frac{1}{(2\pi)^3} \int_{\Gamma^*} H_\xi d\xi,$$

où H_ξ est un opérateur non-borné sur $L_\xi^2(\Gamma)$. L'expression de H_ξ est très simple. On a en effet

$$H_\xi = -\frac{1}{2}\Delta_\xi + V_{\text{per}}. \quad (2.12)$$

Comme Γ est borné, l'opérateur H_ξ est d'inverse compact. Il existe donc pour tout $\xi \in \Gamma^*$, une suite $(\epsilon_{n,\xi})_{n \geq 1}$ et une base hilbertienne $(\psi_{n,\xi})_{n \geq 1}$ de $L_\xi^2(\Gamma)$ telle que pour tout $n \in \mathbb{N} \setminus \{0\}$ et tout $\xi \in \Gamma^*$,

$$H_\xi \psi_{n,\xi} = \epsilon_{n,\xi} \psi_{n,\xi}.$$

Posons $\varphi_{n,\xi}(x) = e^{-i\xi \cdot x} \psi_{n,\xi}(x)$. On vérifie facilement que $(\varphi_{n,\xi})_{n \geq 1}$ est une base hilbertienne de $L_{\text{per}}^2(\Gamma)$ formée de vecteurs propres de l'opérateur auto-adjoint H^ξ sur $L_{\text{per}}^2(\Gamma)$ défini par

$$H^\xi = -\frac{1}{2}\Delta - i\xi \cdot \nabla + \frac{|\xi|^2}{2} + V_{\text{per}}.$$

Plus précisément,

$$H^\xi \varphi_{n,\xi} = \epsilon_{n,\xi} \varphi_{n,\xi}.$$

Comme l'application $\xi \mapsto H^\xi$ est analytique, on peut montrer à partir de l'équation ci-dessus que les fonctions $\xi \mapsto \epsilon_{n,\xi}$ sont lipschitziennes. De plus pour tout $E \in \mathbb{R}$, l'ensemble

$$\{\xi \in \Gamma^* \mid \exists n \in \mathbb{N} \setminus \{0\} \text{ tel que } \epsilon_{n,\xi} = E\}$$

est de mesure nulle.

On peut aussi montrer [42, Thm. XIII.100] que le spectre de H est purement absolument continu et que

$$\sigma(H) = \bigcup_{n=1}^{\infty} b_n \quad \text{où} \quad b_n = \bigcup_{\xi \in \Gamma^*} \{\epsilon_{n,\xi}\} = [\Sigma_n^-, \Sigma_n^+], \quad \Sigma_n^- = \min_{\xi \in \Gamma^*} \epsilon_{n,\xi}, \quad \Sigma_n^+ = \max_{\xi \in \Gamma^*} \epsilon_{n,\xi}$$

L'intervalle b_n est appelé la n -ième bande du spectre de H .

Si H représente l'Hamiltonien de champ moyen relatif à l'état électronique fondamental, la matrice densité de l'état fondamental est

$$\gamma = \chi_{(-\infty, \epsilon_F]}(H)$$

pour un certain $\epsilon_F \in \mathbb{R}$, appelé niveau de Fermi. On peut montrer que γ appartient bien à l'ensemble \mathcal{P}_{per} . Le niveau de Fermi peut être interprété comme un potentiel chimique, autrement dit comme le multiplicateur de Lagrange associé à la contrainte fixant le nombre d'électrons par cellule unité. En utilisant la décomposition de Bloch, on obtient

$$\gamma = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \gamma_\xi d\xi \quad \text{avec} \quad \gamma_\xi = \chi_{(-\infty, \epsilon_F]}(H_\xi) = \sum_{(n,\xi) \mid \epsilon_{n,\xi} \leq \epsilon_F} |\psi_{n,\xi}\rangle \langle \psi_{n,\xi}|.$$

L'opérateur γ possède la densité

$$\rho_\gamma(x) = \gamma(x, x) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{(n,\xi) | \epsilon_{n,\xi} \leq \epsilon_F} |\psi_{n,\xi}(x)|^2 d\xi$$

vérifiant $\rho_\gamma \geq 0$, $\sqrt{\rho_\gamma} \in H_{\text{per}}^1(\Gamma)$ et

$$\int_{\Gamma} \rho_\gamma = Z$$

où Z est le nombre d'électrons par cellule unité (le niveau de Fermi est choisi de telle sorte que la contrainte ci-dessus soit vérifiée). Lorsqu'il y a un gap entre la Z -ième et la $(Z+1)$ -ième bande, c'est-à-dire lorsque $\Sigma_{Z+1}^- > \Sigma_Z^+$, ϵ_F peut prendre n'importe quelle valeur dans l'intervalle $[\Sigma_Z^+, \Sigma_{Z+1}^-]$. Sinon, ϵ_F a une valeur bien définie. Le système se comporte comme un isolant dans le premier cas et comme un conducteur dans le second cas.

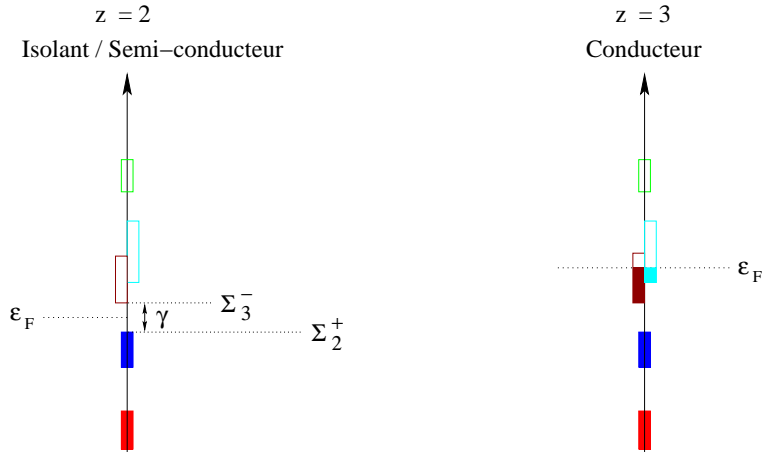


FIG. 2.1 – Spectre de l'opérateur Hamiltonien pour des matériaux isolants/semi-conducteurs et des matériaux conducteurs.

2.2.3 Modèle de Hartree-Fock réduit pour le cristal parfait

Pour un cristal parfait comportant Z électrons par cellule unité, le modèle de Hartree-Fock réduit s'écrit

$$\inf \{ \mathcal{E}_{\text{per}}^0(\gamma), \gamma \in \mathcal{P}_{\text{per}}^Z \} \quad (2.13)$$

où l'énergie de Hartree Fock est donnée par

$$\mathcal{E}_{\text{per}}^0(\gamma) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left(-\frac{1}{2} \Delta_\xi \gamma_\xi \right) + \frac{1}{2} D_{G_1} (\rho_\gamma - \mu_{\text{per}}, \rho_\gamma - \mu_{\text{per}}).$$

La forme bilinéaire D_{G_1} est définie pour toutes fonctions f et g \mathbb{Z}^3 -périodiques suffisamment régulières par

$$D_{G_1}(f, g) := \int_{\Gamma} \int_{\Gamma} G_1(x - y) f(x) g(y) dx dy.$$

où G_1 est la fonction définie de manière unique par

$$\begin{cases} -\Delta G_1 = 4\pi \left(\sum_{k \in \mathbb{Z}^3} \delta_k - 1 \right) \\ G_1 \text{ } \mathbb{Z}^3\text{-périodique, } \min_{\mathbb{R}^3} G_1 = 0. \end{cases}$$

La fonction G_1 est en fait le noyau de Green \mathbb{Z}^3 -périodique de l'interaction de Poisson [37]. Comme G_1 est \mathbb{Z}^3 -périodique, il peut aussi s'écrire sous la forme d'une série de Fourier

$$G_1(x) = c + \sum_{k \in 2\pi\mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} e^{ik \cdot x} \quad (2.14)$$

où $c = \int_{\Gamma} G_1 > 0$. Cela sera démontré dans un cas plus général à la section 2.4. La mesure μ_{per} représente la densité périodique des noyaux définie comme suit

$$\mu_{\text{per}}(x) = \sum_{R \in \mathbb{Z}^3} Z m(x - R). \quad (2.15)$$

Nous supposons pour simplifier que m est une fonction positive de $C_c^\infty(\mathbb{R}^3)$ de support inclus dans Γ et que $\int_{\mathbb{R}^3} m(x) dx = 1$. Un calcul simple montre alors que

$$D_{G_1}(\rho_\gamma - \mu_{\text{per}}, \rho_\gamma - \mu_{\text{per}}) = \sum_{k \in 2\pi\mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} |\hat{\rho}_\gamma(k) - \hat{\mu}_{\text{per}}(k)|^2,$$

où l'on a tenu compte du fait que $\int_{\Gamma} (\rho_\gamma - \mu_{\text{per}}) = 0$ pour tout $\gamma \in \mathcal{P}_{\text{per}}^Z$. Notons que chacun des deux termes de l'énergie $\mathcal{E}_{\text{per}}^0(\gamma)$ est fini pour tout $\gamma \in \mathcal{P}_{\text{per}}$. C'est clair pour le premier, et c'est vrai aussi pour le second puisque $\sqrt{\rho_\gamma} \in H_{\text{per}}^1(\Gamma)$ pour tout $\gamma \in \mathcal{P}_{\text{per}}$.

Une solution de (2.13) est par définition un état électronique fondamental (*electronic ground state*) du cristal parfait, que nous noterons dans toute la suite γ_{per}^0 . La densité électronique correspondante sera notée $\rho_{\gamma_{\text{per}}^0}$ ($\rho_{\gamma_{\text{per}}^0}(x) = \gamma_{\text{per}}^0(x, x)$).

I. Catto, C. Le Bris et P.-L. Lions ont démontré dans [12] que le problème (2.13) admettait un minimiseur γ_{per}^0 et que tous les minimiseurs de (2.13) partageaient la même densité $\rho_{\gamma_{\text{per}}^0}$. Nous donnons ici les idées principales de la preuve. Soit $(\gamma_j)_{j \in \mathbb{N}}$ une suite minimisante pour le problème (2.13). Chacun des opérateurs γ_j peut être décomposé en ondes de Bloch

$$\gamma_j = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (\gamma_j)_\xi d\xi.$$

De plus, comme

$$\mathrm{Tr}_\Gamma(\gamma_j) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \mathrm{Tr}_{L_\xi^2(\Gamma)}((\gamma_j)_\xi) d\xi < \infty,$$

$(\gamma_j)_\xi$ est un opérateur à trace, donc compact, pour tout j et presque tout ξ . En conséquence,

$$(\gamma_j)_\xi(x, y) = \sum_{p \geq 1} \lambda_p^{(j)}(\xi) u_p^{(j)}(\xi, x) \overline{u_p^{(j)}(\xi, y)}, \quad (2.16)$$

où $0 \leq \lambda_p^{(j)}(\xi) \leq 1$, et où $(u_p^{(j)}(\xi, \cdot))_{p \geq 1}$ est une base hilbertienne de $L_\xi^2(\Gamma)$. En exploitant à nouveau le fait que $(\gamma_j)_\xi$ est à trace pour tout j et presque tout ξ , on peut aussi définir la densité

$$(\rho_j)_\xi(x) = \sum_{p \geq 1} \lambda_p^{(j)}(\xi) |u_p^{(j)}(\xi, x)|^2.$$

La fonction $(\rho_j)_\xi$ est \mathbb{Z}^3 -périodique, positive, et si ρ_j désigne la densité de l'opérateur γ_j , on a

$$\rho_j(x) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (\rho_j)_\xi(x) d\xi.$$

Le but de la première étape de la démonstration est de borner les fonctions $(\rho_j)_\xi$ et ρ_j et les opérateurs γ_j et $(\gamma_j)_\xi$ afin de pouvoir extraire de ces suites des sous-suites convergentes permettant, dans les deux autres étapes, de passer successivement à la limite dans la contrainte de charge et dans l'expression de l'énergie.

Pour cela, on remarque d'abord que $0 \leq (\gamma_j)_\xi \leq 1$ pour tout γ_j dans $\mathcal{P}_{\mathrm{per}}^Z$. On obtient ensuite les bornes souhaitées, en utilisant d'une part la contrainte de charge et d'autre part le fait que la suite $(\mathcal{E}^0(\gamma_j))_{j \in \mathbb{N}}$ est bornée. On montre ainsi successivement les propriétés suivantes.

1. La suite $(\gamma_j)_\xi(x, y)$ est bornée dans $L^2(\Gamma^* \times \Gamma \times \Gamma)$, et donc la suite $\gamma_j(x, y)$ est bornée dans $L^2(\Gamma \times \mathbb{R}^3)$, et même dans $L_{\mathrm{unif}}^2(\mathbb{R}^3; L^2(\mathbb{R}^3))$ grâce à l'invariance par translation.
2. En utilisant la décomposition (2.16), on obtient que la suite $\sqrt{(\rho_j)_\xi(x)}$ est bornée dans $H^1(\Gamma)$ et donc dans $L^p(\Gamma)$ pour $1 \leq p \leq 6$ (injections de Sobolev).
3. En corollaire du résultat précédent, la suite $\gamma_j(x, y)$ est bornée dans $L^p(\Gamma \times \Gamma)$, pour $1 \leq p \leq 6$ et $\sqrt{(\rho_j)_\xi(x)}$ est bornée dans $L^2(\Gamma^*; H^1(\Gamma))$.
4. Enfin, en utilisant l'inégalité de Lieb-Thirring, on obtient que la suite $(\rho_j)_\xi(x)$ est bornée dans $L^{\frac{5}{3}}(\Gamma^* \times \Gamma)$.

Après avoir montré que les différentes suites sont bornées, il est possible de supposer (quitte à extraire des sous-suites) que ces suites convergent. On peut ainsi supposer que

1. la suite $(\gamma_j)_\xi(x, y)$ converge faiblement dans $L^2(\Gamma^* \times \Gamma \times \Gamma)$ vers une fonction $(\gamma_\infty)_\xi(x, y)$. On peut dès lors définir un opérateur auto-adjoint γ_∞ sur $L^2(\mathbb{R}^3)$, par

$$\gamma_\infty = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (\gamma_\infty)_\xi d\xi,$$

où $(\gamma_\infty)_\xi$ est un opérateur de Hilbert-Schmidt sur $L^2_\xi(\Gamma)$, dont le noyau est $(\gamma_\infty)_\xi(x, y)$. On a $0 \leq (\gamma_\infty)_\xi \leq 1$ et donc $0 \leq \gamma_\infty \leq 1$.

2. la suite $\sqrt{\rho_j(x)}$ converge faiblement dans $H^1_{\text{per}}(\Gamma)$, fortement dans $L^p(\Gamma)$ pour $1 \leq p < 6$ et presque partout sur \mathbb{R}^3 vers une fonction $\sqrt{\bar{\rho}_\infty(x)} \in H^1_{\text{per}}(\Gamma)$.
3. la suite de fonctions positives $(\rho_j)_\xi(x)$ converge faiblement dans $L^{\frac{5}{3}}(\Gamma^* \times \Gamma)$ vers une fonction positive que nous notons $(\tilde{\rho}_\infty)_\xi(x)$.

On vérifie ensuite facilement que l'on a

$$\frac{1}{(2\pi)^3} \int_{\Gamma^*} (\tilde{\rho}_\infty)_\xi(x) d\xi = \bar{\rho}_\infty(x).$$

On écrira donc désormais $(\bar{\rho}_\infty)_\xi(x)$ au lieu de $(\tilde{\rho}_\infty)_\xi(x)$.

Comme nous l'avons dit plus haut, le but de la deuxième étape de la preuve est de montrer que γ_∞ vérifie la contrainte de charge

$$\frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}(\gamma_\infty)_\xi d\xi = Z,$$

ce qui va être fait en utilisant le fait que l'énergie cinétique est bornée. En effet, on ne peut ici utiliser directement la convergence faible de $(\gamma_j)_\xi$ car l'opérateur identité n'est pas un opérateur de Hilbert-Schmidt. Comme $(\gamma_j)_\xi$ et $-\Delta_\xi(\gamma_j)_\xi$ sont des opérateurs à trace sur $L^2_\xi(\Gamma)$, il en est de même pour $(1 - \Delta)_\xi^{\frac{1}{2}}(\gamma_j)_\xi(1 - \Delta)_\xi^{\frac{1}{2}}$ et de plus la suite $\int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}[(1 - \Delta)_\xi^{\frac{1}{2}}(\gamma_j)_\xi(1 - \Delta)_\xi^{\frac{1}{2}}] d\xi$ est bornée. On peut donc supposer, à extraction près d'une sous-suite, que $(1 - \Delta)_\xi^{\frac{1}{2}}(\gamma_j)_\xi(1 - \Delta)_\xi^{\frac{1}{2}}$ converge de la même manière que γ_ξ et sa limite sera nécessairement $(1 - \Delta)_\xi^{\frac{1}{2}}(\gamma_\infty)_\xi(1 - \Delta)_\xi^{\frac{1}{2}}$. Cette convergence signifie que pour presque tout $\xi \in \Gamma^*$ et pour tout opérateur T auto-adjoint sur $L^2(\mathbb{R}^3)$ se décomposant sous la forme $T = \frac{1}{(2\pi)^3} \int_{\Gamma^*} T_\xi d\xi$ où les T_ξ sont des opérateurs de Hilbert-Schmidt sur $L^2_\xi(\Gamma)$ dont les noyaux $T_\xi(x, y)$ appartiennent à $L^2(\Gamma^*; L^2(\Gamma \times \Gamma))$, on a

$$\begin{aligned} \lim_{j \rightarrow \infty} \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}[(1 - \Delta)_\xi^{\frac{1}{2}}(\gamma_j)_\xi(1 - \Delta)_\xi^{\frac{1}{2}} \cdot T_\xi^*] d\xi \\ = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}[(1 - \Delta)_\xi^{\frac{1}{2}}(\gamma_\infty)_\xi(1 - \Delta)_\xi^{\frac{1}{2}} \cdot T_\xi^*] d\xi. \end{aligned}$$

Alors, en prenant comme fonction test les opérateurs $T_\xi = (1 - \Delta)_\xi^{-1}$, on obtient que

$$\lim_{j \rightarrow \infty} \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}(\gamma_j)_\xi d\xi = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}(\gamma_\infty)_\xi d\xi,$$

et comme les opérateurs γ_j vérifient la contrainte de charge pour tout j , il en est de même pour γ_∞ .

Enfin, la troisième étape de la preuve concerne le passage à la limite dans l'énergie. En utilisant l'étape précédente et le lemme de Fatou, on obtient que

$$\liminf_{j \rightarrow \infty} \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left[-\frac{1}{2} \Delta_\xi (\gamma_j)_\xi \right] d\xi \geq \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left[-\frac{1}{2} \Delta_\xi (\gamma_\infty)_\xi \right] d\xi. \quad (2.17)$$

Par ailleurs, en utilisant la convergence forte de $\rho_j(x)$ dans $L_{\text{per}}^p(\Gamma)$ pour $1 \leq p < 3$, on obtient

$$\lim_{j \rightarrow \infty} \frac{1}{2} D_{G_1}(\rho_j - \mu_{\text{per}}, \rho_j - \mu_{\text{per}}) = \frac{1}{2} D_{G_1}(\bar{\rho}_\infty - \mu_{\text{per}}, \bar{\rho}_\infty - \mu_{\text{per}}). \quad (2.18)$$

Grâce à (2.17) et (2.18), on aura l'existence d'un minimiseur pour le modèle réduit de Hartree-Fock périodique, à condition de montrer que pour presque tout $\xi \in \Gamma^*$ et $x \in \Gamma$,

$$(\bar{\rho}_\infty)_\xi(x) = (\gamma_\infty)_\xi(x, x). \quad (2.19)$$

Ceci se montre en utilisant la convergence faible des opérateurs $(1 - \Delta)_\xi^{\frac{1}{2}} (\gamma_j)_\xi (1 - \Delta)_\xi^{\frac{1}{2}}$ et en utilisant comme opérateurs-tests les opérateurs de Hilbert-Schmidt

$$T_\xi = (1 - \Delta)_\xi^{-\frac{1}{2}} \theta_\xi(\cdot) (1 - \Delta)_\xi^{-\frac{1}{2}}$$

où $\theta_\xi(x)$ est une fonction arbitraire de $L^\infty(\Gamma^* \times \Gamma)$. Il vient alors

$$\begin{aligned} \lim_{j \rightarrow \infty} \int \int_{\Gamma^* \times \Gamma} (\rho_j)_\xi(x) \theta_\xi(x) dx d\xi &= \lim_{j \rightarrow \infty} \int \int_{\Gamma^* \times \Gamma} (\gamma_j)_\xi(x, x) \theta_\xi(x) dx d\xi \\ &= \int \int_{\Gamma^* \times \Gamma} (\gamma_\infty)_\xi(x, x) \theta_\xi(x) dx d\xi \end{aligned}$$

et, grâce à la convergence de $(\rho_j)_\xi$ dans $L^{\frac{5}{3}}(\Gamma^* \times \Gamma)$, on a aussi

$$\lim_{j \rightarrow \infty} \int \int_{\Gamma^* \times \Gamma} (\rho_j)_\xi(x) \theta_\xi(x) dx d\xi = \int \int_{\Gamma^* \times \Gamma} (\bar{\rho}_\infty)_\xi(x) \theta_\xi(x) dx d\xi,$$

d'où (2.19).

Nous prouvons à présent l'unicité de la densité périodique ρ_{per}^0 . Pour cela, supposons qu'il existe deux minimiseurs de l'énergie sur $\mathcal{P}_{\text{per}}^Z$, γ_1 et γ_2 , de densités associées ρ_1 et ρ_2 . Soit $\gamma = \frac{1}{2}\gamma_1 + \frac{1}{2}\gamma_2$, alors

$$\begin{aligned} \mathcal{E}_{\text{per}}^0(\gamma) &= \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left(-\frac{1}{2} \Delta_\xi \gamma_\xi \right) + \frac{1}{2} D_{G_1}(\rho_\gamma - \mu_{\text{per}}, \rho_\gamma - \mu_{\text{per}}), \\ &= \frac{1}{2} \mathcal{E}_{\text{per}}^0(\gamma_1) + \frac{1}{2} \mathcal{E}_{\text{per}}^0(\gamma_2) - \frac{1}{8} D_{G_1}(\rho_1 - \rho_2, \rho_1 - \rho_2), \\ &= I_{\text{per}}^0 - \frac{1}{8} D_{G_1}(\rho_1 - \rho_2, \rho_1 - \rho_2), \end{aligned}$$

où I_{per}^0 est le minimum de l'énergie périodique pour $\gamma \in \mathcal{P}_{\text{per}}^Z$. Par conséquent, $D_{G_1}(\rho_1 - \rho_2, \rho_1 - \rho_2) = 0$. Comme G_1, ρ_1, ρ_2 sont tous \mathbb{Z}^3 -périodiques, on peut récrire l'expression de D_{G_1} avec les coefficients de Fourier de G_1 et $\rho_1 - \rho_2$:

$$D_{G_1}(\rho_1 - \rho_2, \rho_1 - \rho_2) = \sum_{k \in \mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} |\hat{\rho}_1(k) - \hat{\rho}_2(k)|^2.$$

Ainsi, $D_{G_1}(\rho_1 - \rho_2, \rho_1 - \rho_2) = 0$ uniquement si la différence $\rho_1 - \rho_2$ est constante. Or $\int_{\Gamma} \rho_1 - \rho_2 = Z - Z = 0$ et donc $\rho_1 = \rho_2$.

Nous allons en outre montrer dans le prochain chapitre que le minimiseur de (2.13) est unique, que c'est un projecteur orthogonal ($\gamma_{\text{per}}^0 = \gamma_{\text{per}}^0 = \gamma_{\text{per}}^{0*}$) et qu'il vérifie l'équation auto-cohérente

$$\gamma_{\text{per}}^0 = \chi_{(-\infty, \epsilon_F]}(H_{\text{per}}^0) \quad (2.20)$$

où $\epsilon_F \in \mathbb{R}$ (niveau de Fermi), où

$$H_{\text{per}}^0 = -\frac{1}{2}\Delta + V_{\text{per}}$$

et où le potentiel de champ moyen

$$V_{\text{per}}(x) = \int_{\Gamma} G_1(x - y) \left(\rho_{\gamma_{\text{per}}^0}(y) - \mu_{\text{per}}(y) \right) dy$$

est solution du problème

$$\begin{cases} -\Delta V_{\text{per}}(x) = 4\pi \left(\rho_{\gamma_{\text{per}}^0}(x) - \mu_{\text{per}}(x) \right), & \text{(Equation de Poisson)} \\ \forall z \in \mathcal{R}, V_{\text{per}}(x + z) = V_{\text{per}}(x) & \text{(Conditions aux bords périodiques)}. \end{cases}$$

Notons que la solution du système du problème elliptique ci-dessus n'est pas unique, mais que toutes ses solutions sont de la forme $V_{\text{per}} + C$, où C est une constante. Si l'on remplace V_{per} par $V_{\text{per}} + C$, il suffit de remplacer ϵ_F par $\epsilon_F + C$ dans (2.20). Comme $V_{\text{per}} \in L^2_{\text{per}}(\Gamma)$, l'Hamiltonien H_{per}^0 est un opérateur de Schrödinger périodique et son spectre possède la structure décrite à la section précédente.

2.3 Modélisation des cristaux présentant des défauts ponctuels

Notre approche consiste à considérer le défaut comme une “pseudo-molécule” plongée dans le cristal parfait hôte. Cette approche reprend dans un cadre différent l'approche suivie dans [22–25] pour étudier des modèles d'électrodynamique quantique (QED). Dans tout ce qui suit, nous supposons qu'il existe un gap $\Sigma_{Z+1}^- - \Sigma_Z^+ > 0$ (autrement dit que le cristal hôte est un isolant ou un semi-conducteur).

2.3.1 Présentation du modèle de défaut

Si μ désigne la distribution de charge nucléaire du cristal avec défaut, la charge nucléaire de la pseudo-molécule sera la distribution ν telle que

$$\mu(x) = \mu_{\text{per}}(x) + \nu(x).$$

Ainsi par exemple, si on remplace le noyau du cristal parfait situé à l'origine par un noyau de charge Z' (impureté), il vient

$$\nu(x) = (Z' - Z)m(x).$$

Dans ce qui suit, nous supposons que ν est une fonction donnée de $L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$.

De même, on écrira la matrice densité du cristal avec défaut sous la forme

$$\gamma = \gamma_{\text{per}}^0 + Q,$$

Q désignant la matrice densité du défaut. Notons que Q vérifie les contraintes

$$Q^* = Q \quad \text{et} \quad -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0. \quad (2.21)$$

Nous définissons maintenant l'énergie électronique du défaut pour une matrice densité test γ par

$$\mathcal{E}_{\mu_{\text{per}} + \nu}^{\text{rHF}}(\gamma) - \mathcal{E}_{\mu_{\text{per}} + \nu}^{\text{rHF}}(\gamma_{\text{per}}^0).$$

Bien évidemment, cette expression n'a qu'un sens formel puisque les deux quantités sont infinies (μ_{per} est périodique et γ et γ_{per}^0 sont de rang infini). En revanche, en récrivant cette différence en fonction de la variable $Q = \gamma - \gamma_{\text{per}}^0$, on obtient

$$\begin{aligned} \mathcal{E}_{\mu_{\text{per}} + \nu}^{\text{rHF}}(\gamma) - \mathcal{E}_{\mu_{\text{per}} + \nu}^{\text{rHF}}(\gamma_{\text{per}}^0) \text{ " = " } & \text{Tr}(H_{\text{per}}^0 Q) \\ & - \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\nu(x)\rho_Q(y)}{|x-y|} dx dy + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_Q(x)\rho_Q(y)}{|x-y|} dx dy, \end{aligned}$$

et nous verrons qu'il est possible de donner un sens mathématique précis au membre de droite de l'égalité ci-dessus.

En fait, l'ensemble des matrices densité Q sur lequel minimiser le membre de droite de l'égalité ci-dessus pour obtenir l'énergie électronique du défaut va nous être fourni par des arguments de limite thermodynamique. Pour décrire cet ensemble, nous devons d'abord introduire la notion de trace- γ_{per}^0 : un opérateur Q est à trace- γ_{per}^0 si $Q \in \mathfrak{S}_2$ et $\gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0, (1 - \gamma_{\text{per}}^0)Q(1 - \gamma_{\text{per}}^0) \in \mathfrak{S}_1$. On notera $\mathfrak{S}_1^{\gamma_{\text{per}}^0}$ et pour simplifier \mathfrak{S}_1^0 l'ensemble des opérateurs à trace- γ_{per}^0 . La trace- γ_{per}^0 d'un opérateur de la classe \mathfrak{S}_1^0 est alors définie comme

$$\text{Tr}_0(Q) := \text{Tr}_{\gamma_{\text{per}}^0}(Q) = \text{Tr}(\gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0 + (1 - \gamma_{\text{per}}^0)Q(1 - \gamma_{\text{per}}^0)).$$

Notons que $\text{Tr}_0(Q) = \text{Tr}(Q)$ si Q est un opérateur à trace. Pour simplifier les notations, nous définissons

$$Q^{--} := \gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0 \quad \text{et} \quad Q^{++} := (1 - \gamma_{\text{per}}^0)Q(1 - \gamma_{\text{per}}^0).$$

Nous pouvons à présent introduire l'espace de Banach

$$\mathcal{Q} = \{Q \in \mathfrak{S}_1^0 \mid Q^* = Q, |\nabla|Q \in \mathfrak{S}_2, |\nabla|Q^{++}|\nabla| \in \mathfrak{S}_1, |\nabla|Q^{--}|\nabla| \in \mathfrak{S}_1\},$$

doté de la norme

$$\begin{aligned} \|Q\|_{\mathcal{Q}} := & \|Q\|_{\mathfrak{S}_2} + \|Q^{++}\|_{\mathfrak{S}_1} + \|Q^{--}\|_{\mathfrak{S}_1} \\ & + \||\nabla|Q\|_{\mathfrak{S}_2} + \||\nabla|Q^{++}|\nabla|\|_{\mathfrak{S}_1} + \||\nabla|Q^{--}|\nabla|\|_{\mathfrak{S}_1}. \end{aligned} \quad (2.22)$$

La proposition 1 (Section 4.3.1) établira qu'à tout $Q \in \mathcal{Q}$ on peut associer une densité $\rho_Q \in L^2(\mathbb{R}^3) \cap \mathcal{C}$ définie de manière unique par

$$\forall W \in C_c^\infty(\mathbb{R}^3), \quad \text{Tr}(WQ) = \int_{\mathbb{R}^3} \rho_Q(x)W(x) dx$$

et que l'application $Q \mapsto \rho_Q$ est continue. Ici, la notation \mathcal{C} désigne l'espace de Coulomb

$$\mathcal{C} := \{\rho \in \mathcal{S}'(\mathbb{R}^3) \mid \widehat{\rho} \in L_{\text{loc}}^1(\mathbb{R}^3), D(\rho, \rho) < \infty\},$$

où $D(\cdot, \cdot)$ est la forme quadratique positive définie sur $\{\rho \in \mathcal{S}'(\mathbb{R}^3) \mid \widehat{\rho} \in L_{\text{loc}}^1(\mathbb{R}^3)\}$ par

$$D(f, g) := 4\pi \int_{\mathbb{R}^3} \frac{\overline{\widehat{f}(k)} \widehat{g}(k)}{|k|^2} dk.$$

Notons que $L^{\frac{6}{5}}(\mathbb{R}^3) \subset \mathcal{C}$ et que pour des fonctions f et g appartenant à $L^{\frac{6}{5}}(\mathbb{R}^3)$,

$$D(f, g) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{f(x)g(y)}{|x-y|} dx dy.$$

Nous définissons également pour $Q \in \mathcal{Q}$

$$\text{Tr}_0(H_{\text{per}}^0 Q) := \text{Tr}(|H_{\text{per}}^0 - \kappa|^{1/2}(Q^{++} - Q^{--})|H_{\text{per}}^0 - \kappa|^{1/2}) + \kappa \text{Tr}_0(Q)$$

où κ est un nombre réel arbitraire choisi dans le gap $(\Sigma_Z^+, \Sigma_{Z+1}^-)$ (nous montrerons au chapitre suivant que le membre de droite de l'expression ci-dessus a bien un sens et ne dépend pas du choix de κ). Nous sommes maintenant en mesure de donner une définition mathématique précise de la fonctionnelle d'énergie de notre modèle de défaut. Celle-ci s'écrit

$$\mathcal{E}^\nu(Q) := \text{Tr}_0(H_{\text{per}}^0 Q) - D(\rho_Q, \nu) + \frac{1}{2}D(\rho_Q, \rho_Q),$$

où $\kappa \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$ (rappelons que $\mathcal{E}^\nu(Q)$ est indépendant du choix de κ). La fonctionnelle $\mathcal{E}^\nu(Q)$ est définie pour tout $Q \in \mathcal{Q}$. Enfin pour tenir compte des contraintes physiques (2.21), nous introduisons l'ensemble de minimisation

$$\mathcal{K} := \{Q \in \mathcal{Q} \mid -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0\}.$$

On peut vérifier que \mathcal{K} est en fait l'enveloppe convexe fermée des états $Q \in \mathcal{Q}$ qui s'écrivent sous la forme particulière $Q = \gamma - \gamma_{\text{per}}^0$, γ étant un projecteur orthogonal sur $L^2(\mathbb{R}^3)$.

2.3.2 Existence de minimiseurs

Nous nous sommes intéressés à deux situations physiques. Dans le premier cas, on impose le potentiel chimique, la charge correspondant à l'état minimisant l'énergie étant alors une inconnue du problème. Dans le deuxième cas, on impose une contrainte de charge, ce qui donne lieu à un problème mathématique plus difficile.

Problème de minimisation avec un potentiel chimique

Considérons le problème de minimisation suivant

$$\boxed{E_{\epsilon_F}^\nu := \inf\{\mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}_0(Q), Q \in \mathcal{K}\} > -\infty} \quad (2.23)$$

pour ϵ_F dans le gap entre les bandes Z et $Z+1$. Dans ce modèle, la charge du défaut est contrôlée par le potentiel chimique, ou niveau de Fermi, ϵ_F .

Nous allons prouver à la Section 4.3.2 que ce problème admet un minimiseur Q^{ν, ϵ_F} . Le problème (2.23) peut *a priori* admettre plusieurs minimiseurs mais tous les minimiseurs partagent la même densité $\rho_{Q^{\nu, \epsilon_F}}$. Par ailleurs, comme nous l'avons annoncé précédemment, tout minimiseur vérifie l'équation auto-cohérente

$$Q^{\nu, \epsilon_F} = \chi_{(-\infty, \epsilon_F)}(H_{Q^{\nu, \epsilon_F}}) - \gamma_{\text{per}}^0 + \delta, \quad (2.24)$$

où

$$H_{Q^{\nu, \epsilon_F}} = H_{\text{per}}^0 + (\rho_{Q^{\nu, \epsilon_F}} - \nu) * |\cdot|^{-1}$$

et où δ est un opérateur auto-adjoint de rang fini vérifiant $0 \leq \delta \leq 1$ et $\text{Ran}(\delta) \subseteq \ker(H_{Q^{\nu, \epsilon_F}} - \epsilon_F)$.

Nous montrerons que l'opérateur $(\rho_{Q^{\nu, \epsilon_F}} - \nu) * |\cdot|^{-1}$ est une perturbation compacte de l'opérateur de Schrödinger périodique H_{per}^0 , et qu'en conséquence les opérateurs $H_{Q^{\nu, \epsilon_F}}$ et H_{per}^0 ont même spectre essentiel. En revanche, et contrairement au spectre de H_{per}^0 , le spectre de $H_{Q^{\nu, \epsilon_F}}$ peut contenir des valeurs propres isolées de multiplicité finie en dessous de la bande de plus basse énergie et dans les gaps. Les valeurs propres de $H_{Q^{\nu, \epsilon_F}}$ inférieures à ϵ_F correspondent à des électrons localisés piégés par le défaut.

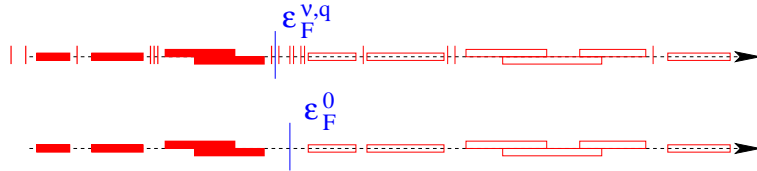


FIG. 2.2 – Spectre de $H_{Q^{\nu, \epsilon_F}}$ (en haut) et de H_{per}^0 (en bas)

Notons $\mathfrak{H}_- = \text{Ran}(\gamma_{\text{per}}^0)$ et $\mathfrak{H}_+ = \text{Ran}(1 - \gamma_{\text{per}}^0)$ les espaces respectivement occupés et virtuels du cristal parfait de base. Supposons maintenant que ϵ_F ne soit pas une valeur propre de $H_{Q^{\nu, \epsilon_F}}$. Dans ce cas $\delta = 0$ et $Q = \gamma - \gamma_{\text{per}}^0$ est la différence de

deux projecteurs orthogonaux. Alors (cf. [23]) il existe une base de Hilbert $(\varphi_i^-)_{i>N_-}$ de \mathfrak{H}_- et une base de Hilbert $(\varphi_i^+)_{i>N_+}$ de \mathfrak{H}_+ telles que dans la base de Hilbert $((\varphi_i^-), (\varphi_i^+))$ de $L^2(\mathbb{R}^3) = \mathfrak{H}_- \hat{\otimes} \mathfrak{H}_+$ on ait

$$Q^{\nu, \epsilon_F} = \left(\begin{array}{c|c|c|c} -I_{N_-} & 0 & 0 & 0 \\ \hline 0 & \text{diag}(-a_1, -a_2, \dots) & 0 & \text{diag}(b_1, b_2, \dots) \\ \hline 0 & 0 & I_{N_+} & 0 \\ \hline 0 & \text{diag}(b_1, b_2, \dots) & 0 & \text{diag}(a_1, a_2, \dots) \end{array} \right)$$

avec

$$0 \leq a_i < 1, \quad \sum_{i=1}^{+\infty} a_i < +\infty, \quad b_i = \sqrt{a_i(1-a_i)}$$

et

$$\text{Tr}_0(Q^{\nu, \epsilon_F}) = N_+ - N_-.$$

Dans notre modèle $\text{Tr}_0(Q)$ sera interprété comme la charge du défaut (le lien précis entre $\text{Tr}_0(Q)$ et les observables physiques restant à éclaircir).

Notons que la matrice Q^{ν, ϵ_F} peut être rendue diagonale en prenant la base de Hilbert

$$((\varphi_i^-)_{-N_- < i \leq 0}, (\psi_i^-)_{i \geq 1}, (\varphi_i^+)_{-N_+ < i \leq 0}, (\psi_i^+)_{i \geq 1})$$

où $\psi_i^\pm = \pm \left(\frac{1 \mp \sqrt{a_i}}{2} \right)^{1/2} \varphi_i^- + \left(\frac{1 \pm \sqrt{a_i}}{2} \right)^{1/2} \varphi_i^+$. Dans cette base

$$Q^{\nu, \epsilon_F} = \left(\begin{array}{c|c|c|c} -I_{N_-} & 0 & 0 & 0 \\ \hline 0 & \text{diag}(-\sqrt{a_1}, -\sqrt{a_2}, \dots) & 0 & 0 \\ \hline 0 & 0 & I_{N_+} & 0 \\ \hline 0 & 0 & 0 & \text{diag}(\sqrt{a_1}, \sqrt{a_2}, \dots) \end{array} \right).$$

On voit que Q^{ν, ϵ_F} sera à trace si et seulement si

$$\sum_{i=1}^{+\infty} \sqrt{a_i} < +\infty,$$

ce qui n'est pas évident puisqu'*a priori* on a seulement $\sum_{i=1}^{+\infty} a_i < +\infty$.

Problème de minimisation avec une contrainte de charge

Dans le cas précédent on ignore *a priori* la charge du défaut, tandis qu'ici nous allons minimiser l'énergie en imposant une contrainte de charge. Dans notre modèle, la charge du défaut est *par définition* $q := \text{Tr}_0(Q)$ (notons que la relation précise entre $\text{Tr}_0(Q)$ et les observables physiques reste à éclaircir). Nous obtenons ainsi le problème suivant

$$\boxed{E^\nu(q) := \inf \{ \mathcal{E}^\nu(Q), Q \in \mathcal{K}, \text{Tr}_0(Q) = q \},} \quad (2.25)$$

dans lequel la charge $q \in \mathbb{R}$ est donnée.

Le premier résultat que nous donnons est l'expression analytique de $E^\nu(q)$, pour le cas sans défaut ($\nu = 0$)

$$E^0(q) = \begin{cases} \Sigma_{Z+1}^- q & \text{pour } q \geq 0 \\ \Sigma_Z^+ q & \text{pour } q \leq 0, \end{cases}$$

où Σ_Z^+ est la borne supérieure de la Z ième bande et Σ_{Z+1}^- la borne inférieure de la $(Z + 1)$ ième bande. Dans ce cas où $\nu = 0$, le problème de minimisation n'a pas de solution, sauf si $q = 0$.

Nous montrons ensuite que l'existence d'un minimiseur $Q^{\nu,q}$ pour le problème (2.25) est équivalente à chacune des deux conditions suivantes :

- toute suite minimisante pour le problème est précompacte dans \mathcal{Q} et converge vers un minimiseur $Q^{\nu,q}$ de l'énergie ;
- des *inégalités de type HVZ* sont vérifiées :

$$\forall q' \in \mathbb{R} \setminus \{0\}, \quad E^\nu(q) < E^\nu(q - q') + E^0(q').$$

Dans le cas où ces conditions sont vérifiées, nous obtenons comme dans le cas précédent que le minimiseur n'est pas nécessairement unique mais que la densité associée $\rho_{Q^{\nu,q}}$ l'est. En outre il existe $\epsilon_F^{\nu,q}$ dans le gap $[\Sigma_Z^+, \Sigma_{Z+1}^-]$ tel que le minimiseur obtenu soit aussi un minimiseur du problème de minimisation (2.23) pour $\epsilon_F = \epsilon_F^{\nu,q}$. Ce minimiseur vérifie l'équation auto-cohérente (2.24) pour un opérateur δ vérifiant $0 \leq \delta \leq 1$ et $\text{Ran}(\delta) \subseteq \ker(H_{Q^{\nu,q}} - \epsilon_F^{\nu,q})$. L'opérateur δ est de rang fini si $\epsilon_F^{\nu,q}$ n'est pas l'une des extrémités du gap. Dans le cas contraire, c'est un opérateur à trace.

Enfin nous établissons que l'ensemble des $q \in \mathbb{R}$ satisfaisant les conditions requises pour l'existence d'un minimiseur est un intervalle fermé non vide de \mathbb{R} et que c'est le plus grand intervalle sur lequel la fonction $q \mapsto E^\nu(q)$ est strictement convexe.

2.4 Supercellule et limite thermodynamique

Nous allons à présent relier les modèles introduits dans les sections précédentes pour les cas avec et sans défaut avec la notion de limite thermodynamique du modèle de supercellule.

Le modèle de supercellule, que nous allons décrire plus en détails dans un instant, consiste à considérer un "grand" domaine de simulation, composé de plusieurs cellules unités et appelé *supercellule*, et à imposer des conditions limites périodiques au bord de la supercellule.

Ce modèle est le plus couramment utilisé dans le calcul de structures électroniques de solides cristallins, mais présente néanmoins plusieurs inconvénients lorsqu'il s'agit de simuler un cristal comportant un défaut. En premier lieu, les conditions au bord périodiques créent des interactions parasites entre le défaut et ses images périodiques. En second lieu, le potentiel électrostatique qui règne au sein de la supercellule doit

par construction respecter les conditions limites périodiques au bord de la supercellule. En conséquence, la supercellule doit être globalement neutre. Lorsqu'on simule un défaut chargé, il faut donc compenser l'excès ou le défaut de charge par une densité de charge uniforme (*jellium background*), ce qui conduit à des approximations mal contrôlées.

Nous commençons par décrire le modèle de supercellule, après quoi nous présenterons les résultats de convergence obtenus pour les cas avec et sans défaut.

2.4.1 Description du modèle de supercellule

Pour $L \in \mathbb{N} \setminus \{0\}$, nous définissons la supercellule

$$\Lambda_L = \left[-L/2, L/2 \right]^3.$$

Le réseau, le réseau réciproque et la zone de Brillouin correspondants sont donc respectivement $\mathcal{R} = LZ^3$, $\mathcal{R}^* = \frac{2\pi}{L}Z^3$ et $[-\frac{\pi}{L}, \frac{\pi}{L}]^3$. Notons que la cellule unité $\Gamma = [-\frac{1}{2}, \frac{1}{2}]^3$ coïncide avec Λ_1 .

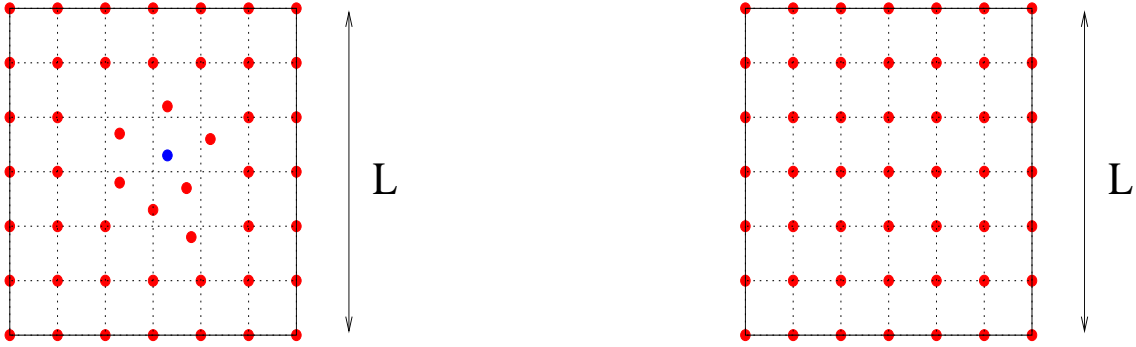


FIG. 2.3 – Supercellule pour le cas avec défaut (à gauche) et sans défaut (à droite).

L'espace de Hilbert sur lequel agissent les opérateurs relatifs à la supercellule est

$$L_{\text{per}}^2(\Lambda_L) = \{ \varphi \in L_{\text{loc}}^2(\mathbb{R}^3) \mid \varphi \text{ } (LZ^3)\text{-périodique} \},$$

et un état électronique admissible est décrit par une matrice densité à un corps γ appartenant à l'ensemble convexe

$$\mathcal{P}_{\text{sc},L} = \left\{ \gamma \in \mathfrak{S}_1(L_{\text{per}}^2(\Lambda_L)) \mid \gamma^* = \gamma, 0 \leq \gamma \leq 1, \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(-\Delta\gamma) < \infty \right\}.$$

Un opérateur $\gamma \in \mathcal{P}_{\text{sc},L}$ est un opérateur auto-adjoint à trace sur $L_{\text{per}}^2(\Lambda_L)$, compris entre 0 et 1 au sens des opérateurs. La condition $\text{Tr}(-\Delta\gamma) < +\infty$ permet de donner un sens à chacun des termes de l'énergie du modèle de Hartree-Fock réduit pour la supercellule (voir les deux sections suivantes). De plus, la densité de charge $\rho_\gamma(x) = \gamma(x, x)$ associée à γ vérifie $\rho_\gamma \geq 0$ et $\sqrt{\rho_\gamma} \in H_{\text{per}}^1(\Lambda_L)$.

Nous introduisons à présent le potentiel de Coulomb $L\mathbb{Z}^3$ -périodique G_L défini comme l'unique solution du système

$$\begin{cases} -\Delta G_L = 4\pi \left(\sum_{k \in \mathbb{Z}^3} \delta_{kL} - \frac{1}{L^3} \right) \\ G_L \text{ } L\mathbb{Z}^3\text{-périodique, } \min_{\mathbb{R}^3} G_L = 0. \end{cases} \quad (2.26)$$

Nous allons vérifier que G_L est relié au noyau de Green G_1 introduit à la Section 2.2.3 par la relation $G_L(x) = L^{-1}G_1(x/L)$ et qu'on a

$$G_L(x) = \frac{c}{L} + \sum_{k \in \mathbb{R}^* \setminus \{0\}} \frac{4\pi}{|k|^2} \frac{1}{L^3} e^{ik \cdot x}, \quad (2.27)$$

la constante c étant celle apparaissant dans (2.14).

Vérifions d'abord que la fonction G_L définie par (2.27) est telle que $G_L(x) = L^{-1}G_1(x/L)$. En partant de (2.14), on peut écrire

$$G_1(x) = c + \sum_{k \in \mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|2\pi k|^2} e^{i2\pi k \cdot x}.$$

D'où

$$G_1\left(\frac{x}{L}\right) = c + \sum_{k \in \mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|2\pi k|^2} e^{i\frac{2\pi k}{L} \cdot x},$$

et par suite

$$\frac{1}{L}G_1\left(\frac{x}{L}\right) = \frac{c}{L} + \sum_{k \in \mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{\left|\frac{2\pi}{L}k\right|^2} \frac{1}{L^3} e^{i\frac{2\pi}{L}k \cdot x} = G_L(x).$$

Vérifions à présent que la fonction G_L définie par (2.27) est solution du système (2.26). Il est évident que G_L est $L\mathbb{Z}^3$ -périodique. Par ailleurs,

$$-\Delta G_L = \frac{4\pi}{L^3} \sum_{k \in \mathbb{Z}^3 \setminus \{0\}} e^{i\frac{2\pi}{L}k \cdot x} \quad \text{dans } \mathcal{D}'(\mathbb{R}^3).$$

On obtient alors (2.26) grâce à l'extension suivante de la formule de Poisson

$$\frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} e^{i\frac{2\pi}{L}k \cdot x} = \sum_{l \in \mathbb{Z}^3} \delta_{lL}, \quad \text{dans } \mathcal{D}'(\mathbb{R}^3). \quad (2.28)$$

Comme la formule ci-dessus joue un rôle important dans la suite, nous en rappelons la preuve. Dans le cas unidimensionnel, cette formule est appelée formule de Poisson et s'exprime ainsi :

$$\frac{1}{L} \sum_{k \in \mathbb{Z}} e^{i\frac{2\pi}{L}k \cdot x} = \sum_{l \in \mathbb{Z}} \delta_{lL}, \quad \text{dans } \mathcal{D}'(\mathbb{R}). \quad (2.29)$$

Pour établir cette formule, nous utilisons la fonction L -périodique f définie par

$$f(x) = x \quad \text{pour } x \in [0, L),$$

pour laquelle

$$f'(x) = 1 - L \sum_{k \in \mathbb{Z}} \delta_{kL} \text{ dans } \mathcal{D}'(\mathbb{R}).$$

En développant f en série de Fourier, on obtient

$$f(x) = \frac{1}{\sqrt{L}} \sum_{l \in \mathbb{Z}} c_l(f) e^{i \frac{2\pi}{L} l x}, \quad (2.30)$$

avec

$$\begin{aligned} c_l(f) &= \frac{1}{\sqrt{L}} \int_0^L f(x) e^{-i \frac{2\pi}{L} l x} dx, \\ &= \frac{1}{\sqrt{L}} \int_0^L x e^{-i \frac{2\pi}{L} l x} dx, \\ &= \frac{iL^{\frac{3}{2}}}{2\pi l}. \end{aligned}$$

La série (2.30) est dérivable dans $\mathcal{D}'(\mathbb{R}^3)$, et

$$\begin{aligned} f'(x) &= \frac{1}{\sqrt{L}} \sum_{l \in \mathbb{Z} \setminus \{0\}} i \frac{2\pi}{L} l c_l(f) e^{i \frac{2\pi}{L} l x}, \\ &= - \sum_{l \in \mathbb{Z} \setminus \{0\}} e^{i \frac{2\pi}{L} l x}. \end{aligned}$$

Par ailleurs, en utilisant la formule des sauts, on obtient

$$f'(x) = 1 - L \sum_{k \in \mathbb{Z}} \delta_{kL},$$

d'où (2.29).

Nous passons à présent au cas tridimensionnel. Nous allons prouver que pour tout $\varphi \in \mathcal{D}(\mathbb{R}^3)$,

$$\left\langle \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} e^{i \frac{2\pi}{L} k}, \varphi \right\rangle = \left\langle \sum_{l \in \mathbb{Z}^3} \delta_{lL}, \varphi \right\rangle. \quad (2.31)$$

Comme l'espace des combinaisons linéaires de fonctions du type $\varphi(x) = \varphi_1(x_1)\varphi_2(x_2)\varphi_3(x_3)$ est dense dans $\mathcal{D}(\mathbb{R}^3)$, il est en réalité suffisant de prouver (2.31) pour une telle fonc-

tion. Il vient

$$\begin{aligned}
\left\langle \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} e^{i\frac{2\pi}{L}k}, \varphi \right\rangle &= \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} \left\langle e^{i\frac{2\pi}{L}k}, \varphi \right\rangle, \\
&= \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} \int_{\Lambda_L} e^{i\frac{2\pi}{L}k \cdot x} \varphi(x) dx, \\
&= \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} \int_{\Lambda_L} \int_{\Lambda_L} \int_{\Lambda_L} e^{i\frac{2\pi}{L}(k_1x_1+k_2x_2+k_3x_3)} \varphi_1(x_1)\varphi_2(x_2)\varphi_3(x_3) dx_1 dx_2 dx_3, \\
&= \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} \left(\int_{\Lambda_L} e^{i\frac{2\pi}{L}k_1x_1} \varphi_1(x_1) dx_1 \right) \left(\int_{\Lambda_L} e^{i\frac{2\pi}{L}k_2x_2} \varphi_2(x_2) dx_2 \right) \left(\int_{\Lambda_L} e^{i\frac{2\pi}{L}k_3x_3} \varphi_3(x_3) dx_3 \right), \\
&= \frac{1}{L^3} \left(\sum_{k_1 \in \mathbb{Z}} \int_{\Lambda_L} e^{i\frac{2\pi}{L}k_1x_1} \varphi_1(x_1) dx_1 \right) \left(\sum_{k_2 \in \mathbb{Z}} \int_{\Lambda_L} e^{i\frac{2\pi}{L}k_2x_2} \varphi_2(x_2) dx_2 \right) \left(\sum_{k_3 \in \mathbb{Z}} \int_{\Lambda_L} e^{i\frac{2\pi}{L}k_3x_3} \varphi_3(x_3) dx_3 \right), \\
&= \frac{1}{L^3} \left(\left\langle L \sum_{k_1 \in \mathbb{Z}} \delta_{k_1L}, \varphi_1 \right\rangle \right) \left(\left\langle L \sum_{k_2 \in \mathbb{Z}} \delta_{k_2L}, \varphi_2 \right\rangle \right) \left(\left\langle L \sum_{k_3 \in \mathbb{Z}} \delta_{k_3L}, \varphi_3 \right\rangle \right), \\
&= \left\langle \sum_{k_1 \in \mathbb{Z}} \delta_{k_1L}, \varphi_1 \right\rangle \left\langle \sum_{k_2 \in \mathbb{Z}} \delta_{k_2L}, \varphi_2 \right\rangle \left\langle \sum_{k_3 \in \mathbb{Z}} \delta_{k_3L}, \varphi_3 \right\rangle, \\
&= \left\langle \sum_{l \in \mathbb{Z}^3} \delta_{lL}, \varphi \right\rangle.
\end{aligned}$$

Cela prouve que

$$-\Delta G_L = 4\pi \left(\frac{1}{L^3} \sum_{k \in \mathbb{Z}^3} e^{i\frac{2\pi}{L}k \cdot x} - \frac{1}{L^3} \right) = 4\pi \left(\sum_{k \in \mathbb{Z}^3} \delta_{kL} - \frac{1}{L^3} \right).$$

À présent, nous pouvons définir pour toute fonction g $L\mathbb{Z}^3$ -périodique et suffisamment régulière

$$(g \star_{\Lambda_L} G_L)(x) := \int_{\Lambda_L} G_L(x-y) g(y) dy,$$

et

$$D_{G_L}(f, g) := \int_{\Lambda_L} \int_{\Lambda_L} G_L(x-y) f(x) g(y) dx dy.$$

2.4.2 Limite thermodynamique pour le cristal parfait

Pour un cristal parfait, la fonctionnelle d'énergie du modèle de supercellule de Hartree-Fock réduit s'écrit

$$\mathcal{E}_{\text{sc},L}^0(\gamma) = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(-\frac{1}{2} \Delta \gamma \right) + \frac{1}{2} D_{G_L}(\rho_\gamma - \mu_{\text{per}}, \rho_\gamma - \mu_{\text{per}})$$

où nous rappelons que $\mu_{\text{per}}(x) = \sum_{R \in \mathbb{Z}^3} Zm(x-R)$ est une fonction \mathbb{Z}^3 - (et donc $L\mathbb{Z}^3$ -) périodique. L'énergie de l'état fondamental est alors donnée par

$$I_{\text{sc},L}^0 = \inf \left\{ \mathcal{E}_{\text{sc},L}^0(\gamma), \gamma \in \mathcal{P}_{\text{sc},L}, \int_{\Lambda_L} \rho_\gamma = \int_{\Lambda_L} \mu_{\text{per}} = ZL^3 \right\}. \quad (2.32)$$

Nous montrerons que (2.32) admet un minimiseur $\gamma_{sc,L}^0$, que tous les minimiseurs de (2.32) ont la même densité $\rho_{\gamma_{sc,L}^0}$, et que cette densité est unique. Par ailleurs, les convergences suivantes ont lieu :

- convergence de l'énergie

$$\lim_{L \rightarrow +\infty} \frac{I_{sc,L}^0}{L^3} = I_{per}^0,$$

- convergence de la densité électronique

$$\sqrt{\rho_{\gamma_{sc,L}^0}} \rightharpoonup \sqrt{\rho_{\gamma_{per}^0}} \quad \text{faiblement dans } H_{loc}^1(\mathbb{R}^3),$$

$$\rho_{\gamma_{sc,L}^0} \rightarrow \rho_{\gamma_{per}^0} \quad \text{fortement dans } L_{loc}^p(\mathbb{R}^3) \quad \text{pour } 1 \leq p < 3 \text{ et p.p..}$$

Soit maintenant

$$H_{sc,L}^0 = -\frac{\Delta}{2} + (\rho_{\gamma_{sc,L}^0} - \mu_{per}) \star_{\Gamma} G_1$$

l'Hamiltonien de champ moyen du modèle de supercellule, vu comme un opérateur agissant sur $L^2(\mathbb{R}^3)$. Nous montrerons que pour tout $L \in \mathbb{N} \setminus \{0\}$, $H_{sc,L}^0 - H_{per}^0$ est un opérateur borné et que

$$\lim_{L \rightarrow \infty} \|H_{sc,L}^0 - H_{per}^0\| = 0.$$

En notant $(\lambda_n^L(\xi))_{n \in \mathbb{N} \setminus \{0\}}$ la suite croissante des valeurs propres de $(H_{sc,L}^0)_{\xi}$ pour $\xi \in \Gamma^*$, il résulte en outre des formules de Courant-Fischer que

$$\lim_{L \rightarrow \infty} \sup_{n \in \mathbb{N} \setminus \{0\}} \sup_{\xi \in \Gamma^*} |\lambda_n^L(\xi) - \lambda_n(\xi)| = 0$$

où $(\lambda_n(\xi))_{n \geq 1}$ sont les valeurs propres de $(H_{per}^0)_{\xi}$ introduites précédemment.

Si l'on suppose enfin qu'il existe un gap entre la Z ième bande et la $(Z+1)$ ième, autrement dit que le cristal est un isolant ou un semi-conducteur, et que l'on fixe un $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, alors pour L suffisamment grand, le minimiseur $\gamma_{sc,L}^0$ de (2.32) est unique, et c'est aussi l'unique minimiseur du modèle grand-canonique suivant :

$$I_{sc,L,\epsilon_F}^0 := \inf \left\{ \mathcal{E}_{sc,L}^0(\gamma) - \epsilon_F \text{Tr}_{L_{per}^2(\Lambda_L)}(\gamma), \quad \gamma \in \mathcal{P}_{sc,L} \right\}.$$

Tout ceci prouve que, pour un cristal parfait, le modèle de supercellule converge bien vers le modèle de Hartree-Fock réduit périodique introduit à la section 2.2.3, ce qui justifie l'utilisation du modèle de supercellule dans les simulations numériques.

2.4.3 Limite thermodynamique pour le cas avec défaut

Après avoir traité le cas sans défaut, il nous est à présent possible de traiter celui avec défaut. L'idée est de calculer l'énergie du défaut en prenant comme référence le cas sans défaut. Nous supposons toujours que le cristal hôte est un isolant ou un

semi-conducteur, autrement dit qu'il existe un gap entre les bandes Z et $Z + 1$ de l'Hamiltonien périodique H_{per}^0 .

Nous rappelons que nous avons pris comme densité de charge du défaut $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) \subset \mathcal{C}$. Dans le modèle de supercellule, il nous faut "périodiser" le défaut. Nous effectuons cette opération en introduisant la fonction

$$\nu_L(x) := \sum_{z \in \mathbb{Z}^3} (\mathbb{1}_{\Lambda_L} \nu)(\cdot - Lz).$$

La fonctionnelle d'énergie pour le modèle de supercellule dans le cas avec défaut est alors définie pour tout $\gamma \in \mathcal{P}_{\text{sc},L}$ par

$$\mathcal{E}_{\text{sc},L}^\nu(\gamma) = \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{1}{2} \Delta \gamma \right) + \frac{1}{2} D_{G_L}(\rho_\gamma - \mu_{\text{per}} - \nu_L, \rho_\gamma - \mu_{\text{per}} - \nu_L).$$

Si l'on choisit de contrôler la charge du défaut à l'aide d'un potentiel chimique (ou niveau de Fermi), $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, le problème de minimisation résultant est le suivant

$$I_{\text{sc},L,\epsilon_F}^\nu = \inf \left\{ \mathcal{E}_{\text{sc},L}^\nu(\gamma) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(\gamma), \gamma \in \mathcal{P}_{\text{sc},L} \right\}. \quad (2.33)$$

Pour passer à la limite thermodynamique, nous allons soustraire respectivement de l'énergie, de la densité électronique et de la matrice densité du minimiseur de (2.33), l'énergie, la densité électronique et la matrice densité du minimiseur de (2.32), et montrer que ces différences admettent une limite lorsque $L \rightarrow +\infty$.

Pour ce qui est de l'énergie, nous obtenons que $I_{\text{sc},L,\epsilon_F}^\nu - I_{\text{sc},L,\epsilon_F}^0$ admet une limite finie lorsque $L \rightarrow +\infty$ et nous sommes capables d'identifier cette limite :

$$\lim_{L \rightarrow \infty} (I_{\text{sc},L,\epsilon_F}^\nu - I_{\text{sc},L,\epsilon_F}^0) = E_{\epsilon_F}^\nu - \int_{\mathbb{R}^3} \nu \left((\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_\Gamma G_1 \right) + \frac{1}{2} D(\nu, \nu)$$

où $E_{\epsilon_F}^\nu$ est défini à la Section 4.3.2 comme l'énergie électronique du défaut lorsque la charge est contrôlée par le potentiel chimique ϵ_F .

De même, la différence des matrices densités des minimiseurs et la différence des densités électroniques associées, admettent des limites faibles : en notant $\gamma_{\text{sc},L}^{\nu,\epsilon_F}$ un minimiseur du problème (2.33), on a, à extraction près,

$$(\gamma_{\text{sc},L}^{\nu,\epsilon_F} - \gamma_{\text{sc},L}^0)(x, y) \rightharpoonup Q^{\nu,\epsilon_F}(x, y)$$

faiblement dans $H_{\text{loc}}^1(\mathbb{R}^3 \times \mathbb{R}^3)$ et fortement dans $L_{\text{loc}}^2(\mathbb{R}^3 \times \mathbb{R}^3)$, où Q^{ν,ϵ_F} est un minimiseur de (2.23). Enfin, pour la densité électronique,

$$\rho_{\gamma_{\text{sc},L}^{\nu,\epsilon_F}} - \rho_{\gamma_{\text{sc},L}^0} \rightharpoonup \rho_{Q^{\nu,\epsilon_F}}$$

faiblement dans $L_{\text{loc}}^2(\mathbb{R}^3)$.

Ces résultats seront prouvés dans la Section 4.4.2, en suivant la démarche de [25]. Comme pour le cas avec défaut traité précédemment nous redéfinissons d'abord un

problème de minimisation dont la variable est la différence Q_L entre une matrice densité test $\gamma \in \mathcal{P}_{\text{sc},L}$ du modèle de supercellule avec défaut et le minimiseur $\gamma_{\text{sc},L}^0$ du modèle de supercellule relatif au cas périodique

$$I_{\text{sc},L,\epsilon_F}^\nu - I_{\text{sc},L,\epsilon_F}^0 = \inf \left\{ E_{\text{sc},L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(Q_L), Q_L \in \mathcal{K}_L \right\} \\ - D_{G_L}(\nu_L, \rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) + \frac{1}{2} D_{G_L}(\nu_L, \nu_L)$$

où

$$E_{\text{sc},L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(Q_L) := -D_{G_L}(\rho_{Q_L}, \nu_L) + \frac{1}{2} D_{G_L}(\rho_{Q_L}, \rho_{Q_L}) \\ + \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(|H_{\text{sc},L}^0 - \epsilon_F|^{1/2} (Q_L^{++L} - Q_L^{--L}) |H_{\text{sc},L}^0 - \epsilon_F|^{1/2} \right)$$

et

$$\mathcal{K}_L = \left\{ Q_L \in \mathfrak{S}_1(L_{\text{per}}^2(\Lambda_L)) \mid Q_L^* = Q_L, |\nabla|Q_L|\nabla| \in \mathfrak{S}_1(L_{\text{per}}^2(\Lambda_L)), \right. \\ \left. -\gamma_{\text{sc},L}^0 \leq Q_L \leq 1 - \gamma_{\text{sc},L}^0 \right\}.$$

Il est assez facile de montrer que $-D_{G_L}(\nu_L, \rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) + \frac{1}{2} D_{G_L}(\nu_L, \nu_L)$ converge vers $-\int_{\mathbb{R}^3} \nu((\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1) + \frac{1}{2} D(\nu, \nu)$ lorsque L tend vers l'infini. La difficulté sera d'établir que

$$\lim_{L \rightarrow \infty} E_{\epsilon_F, L}^\nu = E_{\epsilon_F}^\nu$$

où

$$E_{\epsilon_F, L}^\nu = \inf \left\{ E_{\text{sc},L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(Q_L), Q_L \in \mathcal{K}_L \right\}.$$

Chapitre 3

Simulation numérique des cristaux

Décrire l'état électronique des cristaux comportant des défauts locaux est un problème majeur en physique des états solides, en science des matériaux et en nano-électronique [61, 68, 76].

Ce problème a déjà été abordé sous différentes formes. En effet, comme nous l'avons vu précédemment au Chapitre 2, l'ajout d'un défaut a notamment pour conséquence l'apparition de valeurs propres dans les gaps. Soussi, dans [75], a déjà effectué des calculs de valeurs propres dans des gaps pour la propagation d'ondes dans les cristaux photoniques présentant des perturbations à support compact, mais en utilisant un modèle de supercellule différent du nôtre (et en particulier linéaire). Pisani et ses collaborateurs ont proposé des méthodes de "cluster perturbé" pour traiter des défauts locaux [65, 70], mais celles-ci présentent des problèmes de polarisation du milieu et de transfert de charge entre la zone localisée où se situe le défaut et le reste du cristal. Par ailleurs, on retrouve aussi des problèmes similaires dans les articles concernant la pollution spectrale, mais là encore avec des modèles linéaires (voir notamment [54] et les références incluses).

Pour mettre en place les méthodes numériques associées à notre modèle mathématique pour les défauts ponctuels, nous nous sommes d'abord placés dans un cadre unidimensionnel et nous nous limiterons ici à cette approche. Cependant l'étude du cadre tridimensionnel est actuellement en cours [56]. Notre approche comporte trois étapes principales que nous présenterons dans l'ordre. D'abord, nous résoudrons le problème périodique sur la cellule unité Γ . Nous introduirons ensuite les fonctions de Wannier permettant une localisation de la perturbation de la matrice densité engendrée par le défaut. Enfin nous effectuerons sur cette base d'états localisés une approximation variationnelle qui permettra la résolution numérique du modèle de défaut (2.25). L'approximation variationnelle en question et les résultats numériques obtenus sont détaillés au Chapitre 5.

3.1 Résolution numérique du problème de cellule

Nous avons choisi pour cellule unité $\Gamma = [0, 1)$ et pour zone de Brillouin l'intervalle (ouvert à gauche) $\Gamma^* = (-\pi, \pi]$. Dans ce modèle unidimensionnel, les espaces sur lesquels agissent les divers opérateurs sont

$$L_{\text{per}}^2(\Gamma) = \{u \in L_{\text{loc}}^2(\mathbb{R}), u \mathbb{Z} - \text{périodique}\},$$

et

$$L_{\xi}^2(\Gamma) = \{u \in L_{\text{loc}}^2(\mathbb{R}), u(x)e^{-i\xi x} \mathbb{Z} - \text{périodique}\}.$$

La fonctionnelle d'énergie du modèle 1D considéré dans ce chapitre s'écrit pour une matrice densité admissible

$$\gamma = \frac{1}{2\pi} \int_{\Gamma^*} \gamma_{\xi} d\xi$$

sous la forme

$$E(\gamma) = T(\gamma) + E_{\text{ne}}(\rho_{\gamma}) + E_{\text{ee}}(\rho_{\gamma}),$$

avec

$$\begin{aligned} T(\gamma) &= \frac{A_{\text{ke}}}{2\pi} \int_{\Gamma^*} \text{Tr}_{L_{\xi}^2(\Gamma)} \left(-\frac{1}{2} \Delta \gamma_{\xi} \right) d\xi, \\ E_{\text{ne}}(\rho) &= \int_{\Gamma} V_{\text{nuc}}(x) \rho(x) dx, \\ E_{\text{ee}}(\rho) &= \frac{A_{\text{ee}}}{2} D_{\kappa}(\rho, \rho). \end{aligned}$$

Dans les expressions ci-dessus, V_{nuc} désigne le potentiel de Yukawa \mathbb{Z} -périodique généré par des noyaux de charge Z disposés selon la distribution périodique

$$\rho_{\text{nuc}} = Z \sum_{k \in \mathbb{Z}} \delta_k$$

et D_{κ} correspond à l'interaction de Yukawa

$$D_{\kappa}(\rho, \rho) = \int_{\Gamma} \rho(x) V_{\rho}(x) dx$$

où V_{ρ} est l'unique solution dans $H_{\text{per}}^1(\Gamma)$ de l'équation de Poisson-Boltzmann linéarisée

$$-V_{\rho}'' + \kappa^2 V_{\rho} = 4\pi \rho.$$

Ainsi, avec ces notations, V_{nuc} est égal à $V_{-\rho_{\text{nuc}}}$ et est l'unique solution dans $H_{\text{per}}^1(\Gamma)$ de

$$-V_{\text{nuc}}'' + \kappa^2 V_{\text{nuc}} = -4\pi \rho_{\text{nuc}}. \quad (3.1)$$

Notons que le potentiel V_{nuc} a une expression analytique simple : V_{nuc} est la fonction \mathbb{Z} -périodique telle que

$$\forall x \in [0, 1], \quad V_{\text{nuc}}(x) = 4\pi Z \left(\frac{e^{-\kappa x}}{2\kappa(e^{-\kappa} - 1)} - \frac{e^{\kappa x}}{2\kappa(e^{\kappa} - 1)} \right). \quad (3.2)$$

Pour vérifier cela, notons V la fonction \mathbb{Z} -périodique telle que

$$\forall x \in [0, 1], \quad V(x) = 4\pi Z \left(\frac{e^{-\kappa x}}{2\kappa(e^{-\kappa} - 1)} - \frac{e^{\kappa x}}{2\kappa(e^{\kappa} - 1)} \right)$$

et calculons V' et V'' en utilisant la formule des sauts :

$$\begin{aligned} V'(x) &= W(x) + \sum_{k \in \mathbb{Z}} (V(k+0) - V(k-0)) \delta_k, \\ &= W(x) + (V(0) - V(1)) \sum_{k \in \mathbb{Z}} \delta_k, \\ &= W(x) \end{aligned}$$

où W est l'unique fonction \mathbb{Z} -périodique telle que

$$\forall x \in [0, 1], \quad W(x) = 4\pi Z \left(\frac{-e^{-\kappa x}}{2(e^{-\kappa} - 1)} - \frac{e^{\kappa x}}{2(e^{\kappa} - 1)} \right),$$

et

$$\begin{aligned} V''(x) &= \kappa^2 V(x) + (V'(0) - V'(1)) \sum_{k \in \mathbb{Z}} \delta_k, \\ &= \kappa^2 V(x) + 4\pi Z \sum_{k \in \mathbb{Z}} \delta_k. \end{aligned}$$

Ceci montre que V est une solution de (3.1) dans $H_{\text{per}}^1(\Gamma)$. L'unicité de la solution de (3.1) dans $H_{\text{per}}^1(\Gamma)$ découle du lemme de Lax-Milgram.

Les constantes A_{ke} et A_{ee} présentes dans la définition de l'énergie $E(\gamma)$ servent uniquement à équilibrer les différents termes de la fonctionnelle d'énergie afin qu'ils soient du même ordre de grandeur.

Pour les besoins du numérique nous allons discrétiser Γ et Γ^* avec des grilles uniformes. Ainsi, sur Γ^* , on prend un pas égal à $\frac{2\pi}{N_k}$ avec $N_k = 2\tilde{N}_k \in 2\mathbb{N}^*$ et on définit la suite de points $(\xi(k))_{1 \leq k \leq N_k}$ par

$$\xi(k) = 2\pi \frac{k - \tilde{N}_k}{N_k} = \pi \frac{k - \tilde{N}_k}{\tilde{N}_k} = -\pi + \frac{k\pi}{\tilde{N}_k}.$$

On approche ensuite l'espace $L_{\text{per}}^2(\Gamma)$ introduit précédemment par l'espace X_{N_j} de dimension $N_x = 2N_j + 1$ défini par

$$X_{N_j} = \text{Span} \{e_j(x)\}_{j \in [1, 2N_j + 1]}$$

où $e_j(x) = e^{2i\pi(j - N_j - 1)x}$. Enfin, la grille d'espace $(x(l))_{1 \leq l \leq N_x}$ est définie par

$$x(l) = \frac{l - 1}{N_x}.$$

Pour simplifier l'implémentation, nous n'utilisons qu'une grille de taille N_x (et qu'un espace variationnel X_{N_j} associé), mais la convergence de la méthode aurait été *a priori* meilleure si on avait utilisé une grille de taille $2N_j + 1$ pour discrétiser les fonctions de Bloch et une grille de taille $4N_x + 1$ pour discrétiser la densité.

Maintenant que nous avons introduit ces définitions et notations nous allons pouvoir présenter la méthode de calcul de la densité électronique fondamentale associée à un potentiel de champ moyen \mathbb{Z} -périodique V_{mf} donné.

3.1.1 Calcul de la densité électronique fondamentale associée à un potentiel de champ moyen donné

Soit V_{mf} un potentiel de champ moyen \mathbb{Z} -périodique. Nous fixons le nombre d'électrons par cellule à N_e et considérons l'opérateur Hamiltonien périodique agissant sur $L^2(\mathbb{R})$ défini par

$$H = -\frac{A_{\text{ke}}}{2} \frac{d^2}{dx^2} + V_{\text{mf}}.$$

La densité de l'état fondamental et l'énergie cinétique de cet état sont définies respectivement par

$$\rho^{\text{GS}}(x) = \gamma^0(x, x)$$

et

$$T^{\text{GS}} = \frac{A_{\text{ke}}}{2\pi} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)} \left(-\frac{1}{2} \Delta(\gamma^0)_\xi \right) d\xi$$

où

$$\gamma^0 = \chi_{(-\infty, \epsilon_F]}(H),$$

le niveau de Fermi ϵ_F étant choisi de telle manière que $\text{Tr}_{L^2_{\text{per}}(\Gamma)}(\gamma^0) = N_e$.

Pour construire des approximations $\rho_{\text{app}}^{\text{GS}}$ et $T_{\text{app}}^{\text{GS}}$ et ρ^{GS} et T^{GS} , on considère les N_k opérateurs

$$H_{\xi(k)} = -\frac{A_{\text{ke}}}{2} \frac{d^2}{dx^2} - iA_{\text{ke}}\xi(k) \frac{d}{dx} + \frac{A_{\text{ke}}}{2} |\xi(k)|^2 + V_{\text{mf}}$$

agissant sur $L^2_{\text{per}}(\Gamma)$, et on note $\epsilon_{1, \xi(k)}^{\text{app}} \leq \dots \leq \epsilon_{N_x, \xi(k)}^{\text{app}}$ et

$$u_{n, \xi(k)}^{\text{app}}(x) = \sum_{j=1}^{N_x} C_j^{n, k} e_j(x)$$

les valeurs propres et les vecteurs propres de $H_{\xi(k)}$ obtenus par approximation variationnelle du problème aux valeurs propres

$$H_{\xi(k)} u_{n, \xi(k)} = \epsilon_{n, \xi(k)} u_{n, \xi(k)}$$

dans l'espace X_{N_j} . On obtient ainsi les approximations suivantes de $\gamma^0(x, x')$, ρ^{GS} et T^{GS} :

$$\begin{aligned} \gamma_{\text{app}}^0(x, x') &= \frac{1}{N_k} \sum_{k=1}^{N_k} \sum_{n=1}^{N_e} u_{n, \xi(k)}^{\text{app}}(x) \overline{u_{n, \xi(k)}^{\text{app}}(x')} e^{i\xi(k)x} e^{-i\xi(k)x'}, \\ &= \frac{1}{N_k} \sum_{k=1}^{N_k} \sum_{n=1}^{N_e} \sum_{j=1}^{N_x} \sum_{j'=1}^{N_x} C_j^{n, k} \overline{C_{j'}^{n, k}} e_j(x) \overline{e_{j'}(x')} e^{i\xi(k)x} e^{-i\xi(k)x'}, \\ &= \frac{1}{N_k} \sum_{k=1}^{N_k} \sum_{n=1}^{N_e} \sum_{j=1}^{N_x} \sum_{j'=1}^{N_x} C_j^{n, k} \overline{C_{j'}^{n, k}} e^{2i\pi(j - N_j - 1 + (k - \tilde{N}_k)/N_k)x} e^{-2i\pi(j' - N_j - 1 + (k - \tilde{N}_k)/N_k)x'}, \end{aligned}$$

$$\rho_{\text{app}}^{\text{GS}}(x) = \frac{1}{N_k} \sum_{k=1}^{N_k} \sum_{n=1}^{N_e} |u_{n,\xi(k)}^{\text{app}}(x)|^2 = \frac{1}{N_k} \sum_{k=1}^{N_k} \sum_{n=1}^{N_e} \sum_{j=1}^{N_x} \sum_{j'=1}^{N_x} C_j^{n,k} \overline{C_{j'}^{n,k}} e^{2i\pi(j-j')x}$$

$$T_{\text{GS}} = \frac{1}{N_k} \sum_{k=1}^{N_k} \sum_{n=1}^{N_e} \sum_{j=1}^{N_x} 2\pi^2 A_{\text{ke}} \left(j - N_j - 1 + \frac{k - \tilde{N}_k}{N_k} \right)^2 |C_j^{n,k}|^2.$$

Soit $\tilde{H}_k \in \mathbb{C}^{N_x \times N_x}$ la matrice hermitienne de terme général

$$\begin{aligned} (\tilde{H}_k)_{j_1, j_2} &= \langle e_{j_1} | H_{\xi(k)} | e_{j_2} \rangle, \\ &= 2\pi^2 A_{\text{ke}} \left(\frac{j_1 + j_2}{2} - N_j - 1 + \frac{k - \tilde{N}_k}{N_k} \right)^2 \delta_{j_1, j_2} + \int_{\Gamma} V_{\text{mf}}(x) e^{2i\pi(j_2 - j_1)x} dx. \end{aligned}$$

Les $(\epsilon_{n,\xi(k)}^{\text{app}}, C^{n,k})_{1 \leq n \leq N_x}$ sont obtenus en résolvant le problème aux valeurs propres

$$\tilde{H}_k C^{n,k} = \epsilon_{n,\xi(k)}^{\text{app}} C^{n,k}, \quad [C^{n,k}] * C^{n',k} = \delta_{n,n'}, \quad \epsilon_{1,\xi(k)}^{\text{app}} \leq \epsilon_{2,\xi(k)}^{\text{app}} \leq \dots \leq \epsilon_{N_x,\xi(k)}^{\text{app}}.$$

Dans notre code, le potentiel est donné sur la grille $(x(l))_{1 \leq l \leq N_x}$ et on peut évaluer les termes d'énergie potentielle par transformée de Fourier rapide (*Fast Fourier Transform, FFT*). En effet,

- Pour $1 \leq j_1 \leq j_2 \leq N_x$, on a $(j_2 - j_1 + 1) \in \llbracket 1, N_x \rrbracket$ et par conséquent, en utilisant le fait que le potentiel V_{mf} est à valeurs réelles puis en discrétisant,

$$\begin{aligned} \int_{\Gamma} V_{\text{mf}}(x) e^{2i\pi(j_2 - j_1)x} dx &= \overline{\int_{\Gamma} V_{\text{mf}}(x) e^{-2i\pi(j_2 - j_1)x} dx}, \\ &\simeq \overline{\frac{1}{N_x} \sum_{l=1}^{N_x} V_{\text{mf}} \left(\frac{l-1}{N_x} \right) e^{-2i\pi(j_2 - j_1)(l-1)/N_x}}, \\ &= \overline{\frac{1}{N_x} \sum_{l=1}^{N_x} \mathcal{V}_{\text{mf}}(l) e^{-2i\pi(j_2 - j_1)(l-1)/N_x}}, \\ &= \frac{1}{N_x} \overline{\mathcal{V}_{\text{mf}}^{\text{FFT}}(j_2 - j_1 + 1)}, \end{aligned}$$

en notant $\mathcal{V}_{\text{mf}}(l)$ les valeurs de V_{mf} aux points $(x(l))_{1 \leq l \leq N_x}$ de la grille d'espace.

- Ensuite, pour $1 \leq j_2 < j_1 \leq N_x$,

$$\begin{aligned} \int_{\Gamma} V_{\text{mf}}(x) e^{2i\pi(j_2 - j_1)x} dx &= \overline{\int_{\Gamma} V_{\text{mf}}(x) e^{2i\pi(j_1 - j_2)x} dx}, \\ &= \frac{1}{N_x} \mathcal{V}_{\text{mf}}^{\text{FFT}}(j_1 - j_2 + 1). \end{aligned}$$

Nous pouvons donc à présent donner la formule explicite des \tilde{H}_{j_1, j_2}^k

$$\tilde{H}_{j_1, j_2}^k = \begin{cases} \frac{1}{N_x} \overline{\mathcal{V}_{\text{mf}}^{\text{FFT}}(j_2 - j_1 + 1)} & \text{si } 1 \leq j_1 < j_2 \leq N_x \\ 2\pi^2 A_{\text{ke}} \left(j_1 - N_j - 1 + \frac{k - \tilde{N}_k}{N_k} \right)^2 + \frac{1}{N_x} \mathcal{V}_{\text{mf}}^{\text{FFT}}(1) & \text{si } 1 \leq j_1 = j_2 \leq N_x \\ \frac{1}{N_x} \mathcal{V}_{\text{mf}}^{\text{FFT}}(j_1 - j_2 + 1) & \text{si } 1 \leq j_2 < j_1 \leq N_x. \end{cases}$$

On remarque ensuite que

$$u_{n,-\xi(k)} = \overline{u_{n,\xi(k)}}.$$

On a donc

$$\begin{aligned} \overline{u_{n,\xi(k)}}(x) &= \overline{\sum_{j=1}^{N_x} C_j^{n,k} e_j(x)}, \\ &= \sum_{j=1}^{N_x} \overline{C_j^{n,k}} e^{2i\pi(1+N_j-j)x}, \\ &= \sum_{j=1}^{N_x} \overline{C_j^{n,k}} e^{2i\pi(N_x-j+1-N_j-1)x}, \\ &= \sum_{j=1}^{N_x} \overline{C_j^{n,k}} e_{N_x-j+1}(x), \\ &= \sum_{l=1}^{N_x} \overline{C_{N_x-l+1}^{n,k}} e_l(x), \\ &= \sum_{l=1}^{N_x} C_l^{n,N_k-k} e_l(x) = u_{n,-\xi(k)}(x), \end{aligned}$$

d'où

$$C_j^{n,N_k-k} = \overline{C_j^{n,k}}.$$

En tenant compte de ceci, nous pouvons récrire $T_{\text{app}}^{\text{GS}}$ et $\rho_{\text{app}}^{\text{GS}}$ comme suit

$$\begin{aligned} T_{\text{app}}^{\text{GS}} &= \frac{1}{N_k} \sum_{n=1}^{N_e} \sum_{j=1}^{N_x} 2\pi^2 A_{\text{ke}} \left((j - N_j - 1)^2 |C_j^{n,\tilde{N}_k}|^2 + 2 \sum_{k=\tilde{N}_k+1}^{N_k-1} \left(j - N_j - 1 + \frac{k - \tilde{N}_k}{N_k} \right)^2 |C_j^{n,k}|^2 \right. \\ &\quad \left. + \left(j - N_j - 1 + \frac{1}{2} \right)^2 |C_j^{n,N_k}|^2 \right), \\ \rho_{\text{app}}^{\text{GS}} &= \frac{1}{N_k} \sum_{n=1}^{N_e} \left(|u_{n,\xi(\tilde{N}_k)}|^2 + 2 \sum_{k=\tilde{N}_k+1}^{N_k-1} |u_{n,\xi(k)}|^2 + |u_{n,\xi(N_k)}|^2 \right). \end{aligned}$$

Les valeurs $u_{n,\xi(k)}$ aux points de la grille d'espace peuvent à leur tour se récrire de manière simplifiée

$$\begin{aligned} u_{n,\xi(k)}(x(l)) &= \sum_{j=1}^{N_x} C_j^{n,k} e_j(x(l)), \\ &= \sum_{j=1}^{N_x} C_j^{n,k} e^{2i\pi(j-1-N_j)(l-1)/N_x}, \\ &= w_{n,k}(l) e^{-2i\pi N_j(l-1)/N_x}, \end{aligned}$$

où l'on a posé

$$w_{n,k}(l) = \sum_{j=1}^{N_x} C_j^{n,k} e^{2i\pi(j-1)(l-1)/N_x} = N_x [C^{n,k}]^{\text{IFFT}}(l),$$

IFFT désignant la transformée de Fourier rapide inverse.

On en déduit l'expression finale de $\rho_{\text{app}}^{\text{GS}}$ aux points $(x(l))_{1 \leq l \leq N_x}$:

$$\rho_{\text{app}}^{\text{GS}}(x(l)) = \rho(l) = \frac{N_x^2}{N_k} \sum_{n=1}^{N_e} \left(|[C^{n, \tilde{N}_k}]^{\text{IFFT}}(l)|^2 + 2 \sum_{k=\tilde{N}_k+1}^{N_k-1} |[C^{n, k}]^{\text{IFFT}}(l)|^2 + |[C^{n, N_k}]^{\text{IFFT}}(l)|^2 \right).$$

Pour résumer, à partir des valeurs du potentiel de champ moyen en les points de la grille $(x(l))_{1 \leq l \leq N_x}$, on obtient une approximation du spectre de bandes $(\epsilon_{n, \xi(k)}^{\text{app}})_{1 \leq n \leq N_x, 1 \leq k \leq N_k}$, de l'énergie cinétique de l'état fondamental et de la densité de l'état fondamental aux points de la grille, $(\rho(l))_{1 \leq l \leq N_x}$.

Notons qu'on peut tester numériquement le caractère isolant du cristal parfait, qui est une hypothèse fondamentale de notre modèle de défaut, en vérifiant que

$$\Sigma_- = \max_{1 \leq k \leq N_k} \epsilon_{N_e, \xi(k)}^{\text{app}} < \Sigma_+ = \min_{1 \leq k \leq N_k} \epsilon_{N_e+1, \xi(k)}^{\text{app}}.$$

3.1.2 Calcul du potentiel de champ moyen associé à une densité donnée

Nous avons vu à la section précédente comment calculer une approximation numérique de la densité de l'état fondamental associé à un potentiel de champ moyen donné. Nous examinons maintenant le problème inverse, qui consiste à calculer numériquement une approximation du potentiel de champ moyen associé à une densité ρ \mathbb{Z} -périodique donnée. Le potentiel de champ moyen associé à la densité ρ est

$$V_{\text{mf}} = V_{\text{nuc}} + V_\rho$$

où V_{nuc} est donné par (3.2) et où V_ρ est l'unique solution dans $H_{\text{per}}^1(\Gamma)$ de l'équation

$$-\mathcal{V}'' + \kappa^2 \mathcal{V} = 4\pi\rho.$$

Dans notre code numérique, la densité ρ appartient à X_{N_j} :

$$\rho(x) = \sum_{j=1}^{N_x} \hat{\rho}_j e_j(x).$$

Nous obtenons donc

$$\mathcal{V}(x) = \sum_{j=1}^{N_x} \hat{V}_j e_j(x),$$

où \hat{V}_j est donné par

$$\hat{V}_j = 4\pi \frac{\hat{\rho}_j}{\kappa^2 + 4\pi^2(j - N_j - 1)^2}.$$

En pratique, nous n'avons besoin que des valeurs de la transformée de Fourier discrète de V (qui serviront à construire la matrice de l'Hamiltonien de champ moyen comme expliqué à la section précédente). Nous allons donc calculer à présent V^{FFT} à partir des valeurs de ρ en les points de la grille $(x(l))_{1 \leq l \leq N_x}$. Soit $N_j + 1 \leq j \leq N_x$. On a

$$\begin{aligned}\hat{\rho}_j &= \int_{\Gamma} \rho(x) \overline{e_j(x)} dx, \\ &= \frac{1}{N_x} \sum_{l=1}^{N_x} \rho_{\text{GS}} \left(\frac{l-1}{N_x} \right) e^{-2i\pi(j-N_j-1)(l-1)/N_x}, \\ &= \frac{1}{N_x} \sum_{l=1}^{N_x} \rho(l) e^{-2i\pi(j-N_j-1)(l-1)/N_x}\end{aligned}$$

d'où

$$\forall N_j + 1 \leq j \leq N_x, \quad \hat{\rho}_j = \frac{1}{N_x} \rho^{\text{FFT}}(j - N_j).$$

De la même manière, on peut écrire

$$\forall N_j + 1 \leq j \leq N_x, \quad \hat{V}_j = \frac{1}{N_x} V^{\text{FFT}}(j - N_j)$$

En reprenant l'expression de \hat{V}_j en fonction de $\hat{\rho}_j$, on obtient

$$\forall N_j + 1 \leq j \leq N_x, \quad V^{\text{FFT}}(j - N_j) = 4\pi \frac{\rho^{\text{FFT}}(j - N_j)}{\kappa^2 + 4\pi^2(j - N_j - 1)^2}$$

et donc

$$\forall 1 \leq m \leq N_j + 1, \quad V^{\text{FFT}}(m) = 4\pi \frac{\rho^{\text{FFT}}(m)}{\kappa^2 + 4\pi^2(m - 1)^2}. \quad (3.3)$$

Nous utilisons maintenant le fait que le potentiel est réel pour calculer les autres termes de la FFT : pour $1 \leq m \leq N_x - 1$,

$$\begin{aligned}V^{\text{FFT}}(N_x - m + 1) &= \sum_{l=1}^{N_x} V(l) e^{-2i\pi(N_x-m)(l-1)/N_x}, \\ &= \sum_{l=1}^{N_x} V(l) e^{2i\pi m(l-1)/N_x}, \\ &= \overline{\sum_{l=1}^{N_x} V(l) e^{-2i\pi m(l-1)/N_x}}, \\ &= \overline{V^{\text{FFT}}(m + 1)}.\end{aligned}$$

Finalement,

$$\forall N_j + 2 \leq m \leq N_x, \quad V^{\text{FFT}}(m) = \overline{V^{\text{FFT}}(N_x - m + 2)}. \quad (3.4)$$

Notons que $D_\kappa(\rho, \rho)$ s'exprime en fonction des valeurs de la FFT de ρ . En utilisant toujours le fait que ρ et V_ρ sont à valeurs réelles, on a $\hat{\rho}_{N_x-j+1} = \overline{\hat{\rho}_j}$ et $\hat{V}_{N_x-j+1} = \overline{\hat{V}_j}$ pour tout $1 \leq j \leq N_x$, on obtient

$$\begin{aligned} D_\kappa(\rho, \rho) &= \int_\Gamma \rho V_\rho = \sum_{j=1}^{N_x} \hat{\rho}_j \overline{\hat{V}_j} = \hat{\rho}_{N_x+1} \overline{\hat{V}_{N_x+1}} + \sum_{j=N_x+2}^{N_x} \left(\hat{\rho}_j \overline{\hat{V}_j} + \overline{\hat{\rho}_j} \hat{V}_j \right) \\ &= 4\pi |\hat{\rho}_{N_x+1}|^2 + 2 \sum_{j=N_x+2}^{N_x} \frac{4\pi |\hat{\rho}_j|^2}{\kappa^2 + 4\pi^2(j - N_x - 1)^2} \\ &= \frac{4\pi}{N_x^2} \left(\frac{|\rho^{\text{FFT}}(1)|^2}{\kappa^2} + 2 \sum_{m=2}^{N_x+1} \frac{|\rho^{\text{FFT}}(m)|^2}{\kappa^2 + 4\pi^2(m - 1)^2} \right). \end{aligned}$$

3.1.3 Optimal Damping Algorithm (ODA)

L'Optimal Damping Algorithm (ODA) permet de rechercher le minimum d'une fonction $E(\gamma)$ sur l'ensemble convexe

$$\mathcal{P} = \{ \gamma \in \mathcal{B}(X) \mid \gamma = \gamma^*, 0 \leq \gamma \leq 1, \text{Tr}(\gamma) = N_e \}$$

où $\mathcal{B}(X)$ désigne l'espace des opérateurs bornés sur un espace vectoriel X donné. Notons que \mathcal{P} est l'enveloppe convexe de l'espace $\tilde{\mathcal{P}}$ des projecteurs orthogonaux de rang N_e

$$\tilde{\mathcal{P}} = \{ \gamma \in \mathcal{B}(X) \mid \gamma = \gamma^* = \gamma^2, \text{Tr}(\gamma) = N_e \}.$$

L'algorithme ODA est un algorithme itératif, dont chaque itération consiste en deux étapes. Soit γ^N la valeur de γ à l'itération N . Pour construire γ^{N+1} , on commence par construire un minimiseur $\tilde{\gamma}^{N+1} \in \tilde{\mathcal{P}}$ du problème

$$\tilde{\gamma}^{N+1} = \underset{\tilde{\gamma} \in \tilde{\mathcal{P}}}{\text{arginf}} \frac{d}{dt} (E((1-t)\gamma^N + t\tilde{\gamma})) \Big|_{t=0}. \quad (3.5)$$

Dans notre cas,

$$E(\gamma) = T(\gamma) + V_{\text{ne}}(\gamma) + E_{\text{ee}}(\gamma)$$

et résoudre (3.5) revient en fait à minimiser sur $\tilde{\mathcal{P}}$ la fonction $\gamma \mapsto \text{Tr}(H_{\rho_\gamma} \tilde{\gamma})$. En effet,

$$\begin{aligned} & \frac{d}{dt} (E((1-t)\gamma^N + t\tilde{\gamma})) \Big|_{t=0} \\ &= \frac{d}{dt} \left((1-t)T(\gamma^N) + tT(\tilde{\gamma}) + (1-t)E_{\text{ne}}(\gamma^N) + tE_{\text{ne}}(\tilde{\gamma}) + \frac{A_{\text{ee}}}{2} (D_\kappa(\rho_{\gamma^N}, \rho_{\gamma^N}) \right. \\ & \quad \left. + 2tD_\kappa(\rho_{\gamma^N}, \rho_{\tilde{\gamma}} - \rho_{\gamma^N}) + t^2D_\kappa(\rho_{\tilde{\gamma}} - \rho_{\gamma^N}, \rho_{\tilde{\gamma}} - \rho_{\gamma^N})) \right) \Big|_{t=0}, \\ &= (T(\tilde{\gamma}) - T(\gamma^N) + E_{\text{ne}}(\tilde{\gamma}) - E_{\text{ne}}(\gamma^N) + A_{\text{ee}}D_\kappa(\rho_{\gamma^N}, \rho_{\tilde{\gamma}} - \rho_{\gamma^N}) + tA_{\text{ee}}D_\kappa(\rho_{\tilde{\gamma}} - \rho_{\gamma^N}, \rho_{\tilde{\gamma}} - \rho_{\gamma^N})) \Big|_{t=0}, \\ &= T(\tilde{\gamma}) - T(\gamma^N) + E_{\text{ne}}(\tilde{\gamma}) - E_{\text{ne}}(\gamma^N) + A_{\text{ee}}D_\kappa(\rho_{\gamma^N}, \rho_{\tilde{\gamma}} - \rho_{\gamma^N}). \end{aligned}$$

On recherche donc un minimiseur sur $\tilde{\mathcal{P}}$ de la fonction

$$\tilde{\gamma} \mapsto T(\tilde{\gamma}) + E_{\text{ne}}(\tilde{\gamma}) + A_{\text{ee}}D_{\kappa}(\rho_{\gamma^N}, \rho_{\tilde{\gamma}}) = \text{Tr}(H_{\rho_{\gamma^N}}\tilde{\gamma})$$

où $H_{\rho_{\gamma^N}}$ est l'opérateur Hamiltonien

$$H_{\rho_{\gamma^N}} = -\frac{A_{ke}}{2} \frac{d^2}{dx^2} + V_{\text{nuc}} + V_{\rho_{\gamma^N}}.$$

D'où

$$\tilde{\gamma}^{N+1} = \mathbb{1}_{(-\infty, \epsilon_F^{N+1})}(H_{\rho_{\gamma^N}}) + \delta,$$

où ϵ_F^{N+1} est choisi tel que $\text{Tr}(\tilde{\gamma}^{N+1}) = N_e$ et où δ est un projecteur orthogonal d'image incluse dans $\ker(H_{\rho_{\gamma^N}} - \epsilon_F^{N+1})$.

Cette première étape est effectuée numériquement par la routine décrite à la section (3.1.1). En fait, comme nous allons le voir dans ce qui suit, nous n'aurons pas besoin de calculer (et de stocker en mémoire) $\tilde{\gamma}^{N+1}$, mais seulement l'énergie cinétique \tilde{T}^{N+1} , la densité $\tilde{\rho}^{N+1}(x) = \tilde{\gamma}^{N+1}(x, x)$ et les énergies $\epsilon_{n, \xi(k)}^{N+1}$ correspondantes.

La deuxième étape consiste à trouver le minimiseur de l'énergie $E(\gamma)$ sur le "segment" $[\gamma^N, \tilde{\gamma}^{N+1}]$. Pour ce faire, on introduit la fonction

$$\begin{aligned} f^N(\alpha) &= E((1-\alpha)\gamma^N + \alpha\tilde{\gamma}^{N+1}), \\ &= (1-\alpha)T^N + \alpha\tilde{T}^{N+1} + E_{\text{ne}}((1-\alpha)\rho^N + \alpha\tilde{\rho}^{N+1}) + E_{\text{ee}}((1-\alpha)\rho^N + \alpha\tilde{\rho}^{N+1}), \end{aligned}$$

où $T^N = T(\gamma^N)$ et $\tilde{T}^{N+1} = T(\tilde{\gamma}^{N+1})$, et on recherche α^{opt} tel que

$$\alpha^{\text{opt}} = \text{argmin} \{ f^N(\alpha), \alpha \in [0, 1] \}.$$

Exprimons $f^N(\alpha)$ en fonction de α . Il vient

$$\begin{aligned} f^N(\alpha) &= T^N + \alpha(\tilde{T}^{N+1} - T^N) + E_{\text{ne}}(\tilde{\rho}^N) + \alpha E_{\text{ne}}(\tilde{\rho}^{N+1} - \rho^N) \\ &\quad + E_{\text{ee}}(\rho^N) + \alpha^2 E_{\text{ee}}(\tilde{\rho}^{N+1} - \rho^N) + \alpha A_{\text{ee}}D_{\kappa}(\rho^N, \tilde{\rho}^{N+1} - \rho^N), \\ &= E^N + \alpha(\tilde{T}^{N+1} - T^N + E_{\text{ne}}(\tilde{\rho}^{N+1} - \rho^N) + A_{\text{ee}}D_{\kappa}(\rho^N, \tilde{\rho}^{N+1} - \rho^N)) \\ &\quad + \alpha^2 E_{\text{ee}}(\tilde{\rho}^{N+1} - \rho^N), \\ &= a^N \alpha^2 + b^N \alpha + E^N, \end{aligned}$$

avec $E^N = E(\gamma^N)$ et

$$a^N = E_{\text{ee}}(\tilde{\rho}^{N+1} - \rho^N),$$

et

$$b^N = \tilde{T}^{N+1} - T^N + E_{\text{ne}}(\tilde{\rho}^{N+1} - \rho^N) + A_{\text{ee}}D_{\kappa}(\rho^N, \tilde{\rho}^{N+1} - \rho^N).$$

On peut montrer facilement que

$$b^N = \frac{1}{N_k} \sum_{n=1}^{N_e} \sum_{k=1}^{N_k} \epsilon_{n, \xi(k)}^{N+1} - 2E^N + T^N + E_{\text{ne}}^N,$$

où $E_{\text{ne}}^N = E_{\text{ne}}(\rho^N)$.

Pour initialiser l'algorithme ODA, on prend comme potentiel de départ $V^0 = V_{\text{nuc}}$, pour lequel on dispose de la formule explicite (3.2). En procédant comme à la section 3.1.1, on obtient des approximations ρ^0 , T^0 , E_{ne}^0 et E^0 de la densité, de l'énergie cinétique, de l'énergie d'interaction noyaux-électrons et de l'énergie totale de l'état fondamental associé au potentiel de champ moyen V_0 .

Nous passons à présent à la description des itérations.

- ρ^N étant donné sur les points de la grille $(x(l))_{1 \leq l \leq N_x}$, on calcule sa transformée de Fourier rapide, qui permet d'obtenir facilement $(V^N)^{\text{FFT}}$ (cf. section 3.1.2). On construit alors le potentiel de champ moyen $(V_{\text{mf}}^N)^{\text{FFT}} = V_{\text{nuc}}^{\text{FFT}} + (V^N)^{\text{FFT}}$.
- On assemble ensuite la matrice de l'Hamiltonien de champ moyen dans la base $(e_j)_{1 \leq j \leq N_x}$ de X_{N_j} pour chaque point k de la grille $(\xi(k))_{1 \leq k \leq N_k}$. En diagonalisant ces matrices, on obtient les valeurs propres $(\epsilon_{n,\xi(k)}^{N+1})$, et on peut calculer \tilde{T}^{N+1} , $\tilde{\rho}^{N+1}$, et $\tilde{E}_{\text{ne}}^{N+1}$.
- On calcule ensuite α^{N+1} comme expliqué ci-dessus
- On pose alors

$$\begin{aligned} \rho^{N+1} &= (1 - \alpha^{\text{opt}})\rho^N + \alpha^{\text{opt}}\tilde{\rho}^{N+1}, \\ T^{N+1} &= (1 - \alpha^{\text{opt}})T^N + \alpha^{\text{opt}}\tilde{T}^{N+1}, \\ E_{\text{ne}}^{N+1} &= (1 - \alpha^{\text{opt}})E_{\text{ne}}^{N+1} + \alpha^{\text{opt}}\tilde{E}_{\text{ne}}^{N+1}, \\ E^{N+1} &= a^N(\alpha^{N+1})^2 + b^N\alpha^{N+1} + E^N. \end{aligned}$$

- Enfin, on décide d'un critère d'arrêt. Dans la version actuelle du code il s'agit d'un nombre d'itérations maximal, qui a été choisi en fonction de la convergence obtenue suite à divers essais. Il serait cependant préférable de fixer une tolérance $\epsilon > 0$ et renvoyer à la première étape si $\|\rho^{N+1} - \rho^N\| \geq \epsilon$ pour une certaine norme.

3.2 Résolution du problème de supercellule

Considérons la supercellule $\Lambda_{N_k} = [0, N_k)$. Pour résoudre numériquement le problème de supercellule, on introduit la grille uniforme en espace $(x^{\text{sc}}(l))_{1 \leq l \leq N_x N_k}$ où

$$x^{\text{sc}}(l) = \frac{l-1}{N_x}$$

et le sous-espace Y_{N_x, N_k} de $L^2_{\text{per}}(\Lambda_{N_k})$ défini par

$$Y_{N_x, N_k} = \text{Span} \{f_\mu(x)\}_{1 \leq l \leq N_x N_k},$$

avec

$$f_\mu(x) = \frac{1}{\sqrt{N_k}} e^{2i\pi(\mu - N_x \tilde{N}_k)x/N_k}.$$

Avec cette convention de normalisation, la famille $(f_\mu(x))_{1 \leq l \leq N_x N_k}$ est une base orthonormale de Y_{N_x, N_k} .

3.2.1 Calcul de la densité électronique fondamentale associée à un potentiel de champ moyen donné

Soit \mathcal{W} un potentiel de champ moyen $N_k\mathbb{Z}$ -périodique et

$$H^{\mathcal{W},\text{sc}} = -\frac{1}{2} \frac{d^2}{dx^2} + \mathcal{W}$$

l'Hamiltonien associé (il s'agit d'un opérateur auto-adjoint sur $L^2_{\text{per}}(\Lambda_{N_k})$).

Pour calculer numériquement l'état fondamental de l'opérateur $H^{\mathcal{W},\text{sc}}$, on diagonalise la matrice de cet opérateur dans la base $(f_\mu(x))_{1 \leq \mu \leq N_x N_k}$. En pratique, les valeurs de \mathcal{W} sont données sur la grille d'espace $(x^{\text{sc}}(l))_{1 \leq l \leq N_x N_k}$, et on obtient

$$\begin{aligned} [H^{\mathcal{W},\text{sc}}]_{\mu_1\mu_2} &= \langle H^{\mathcal{W},\text{sc}} f_{\mu_1}, f_{\mu_2} \rangle_{L^2_{\text{per}}(\Lambda_{N_k})}, \\ &= 2\pi^2 \frac{(\mu_1 - N_x \tilde{N}_k)^2}{N_k^2} \delta_{\mu_1\mu_2} + \frac{1}{N_k} \int_{\Lambda_{N_k}} \mathcal{W}(x) e^{2i\pi(\mu_2 - \mu_1)x/N_k} dx, \\ &\simeq 2\pi^2 \frac{(\mu_1 - N_x \tilde{N}_k)^2}{N_k^2} \delta_{\mu_1\mu_2} + \frac{1}{N_x N_k} \sum_{l=1}^{N_x N_k} W(l) e^{2i\pi(\mu_2 - \mu_1)(l-1)/(N_x N_k)}, \end{aligned}$$

où l'on a posé $W(l) = \mathcal{W}(x^{\text{sc}}(l)) = \mathcal{W}\left(\frac{l-1}{N_x}\right)$.

Comme dans le problème sur la cellule unité, nous pouvons ensuite exprimer les coefficients de la matrice $[H^{\mathcal{W},\text{sc}}]$ en fonction des valeurs par transformée de Fourier rapide de la suite $(W(l))_{1 \leq l \leq N_x N_k}$. Nous obtenons ainsi

$$[H^{\mathcal{W},\text{sc}}]_{\mu_1\mu_2} = \begin{cases} \frac{1}{N_x N_k} \overline{W^{\text{FFT}}(\mu_2 - \mu_1 + 1)} & \text{si } 1 \leq \mu_1 < \mu_2 \leq N_x N_k, \\ 2\pi^2 \frac{(\mu - N_x \tilde{N}_k)^2}{N_k^2} + \frac{1}{N_x N_k} \overline{W^{\text{FFT}}(1)} & \text{si } 1 \leq \mu_1 = \mu = \mu_2 \leq N_x N_k, \\ \frac{1}{N_x N_k} W^{\text{FFT}}(\mu_1 - \mu_2 + 1) & \text{si } 1 \leq \mu_2 < \mu_1 \leq N_x N_k. \end{cases}$$

Une approximation de la matrice densité de l'état fondamental est obtenue par la formule suivante

$$\begin{aligned} (\gamma^0)^{\text{sc}}(x, x') &= \sum_{n=1}^{N_e N_k} \psi_n^{\text{sc}}(x) \overline{\psi_n^{\text{sc}}(x')}, \\ &= \sum_{n=1}^{N_e N_k} \sum_{\mu, \mu'=1}^{N_x N_k} C_\mu^n \overline{C_{\mu'}^n} f_\mu(x) \overline{f_{\mu'}(x')}, \\ &= \frac{1}{N_k} \sum_{n=1}^{N_e N_k} \sum_{\mu, \mu'=1}^{N_x N_k} C_\mu^n \overline{C_{\mu'}^n} e^{2i\pi(\mu - N_x \tilde{N}_k)x/N_k} e^{-2i\pi(\mu' - N_x \tilde{N}_k)x'/N_k}. \end{aligned}$$

Dans les expressions ci-dessus,

$$\psi_n^{\text{sc}}(x) = \sum_{\mu=1}^{N_x N_k} C_\mu^n f_\mu(x)$$

et $(C^n)_{1 \leq n \leq N_x N_k}$ désigne une base orthonormale de vecteurs propres de la matrice de $H^{\mathcal{W}, \text{sc}}$ associée aux valeurs propres $\epsilon_{N_x N_k}^{\text{sc}}$ rangées par ordre croissant :

$$H^{\mathcal{W}, \text{sc}} C^N = \epsilon_n^{\text{sc}} C^n, \quad [C^n]^* C^{n'} = \delta_{nn'}, \quad \epsilon_1^{\text{sc}} \leq \epsilon_2^{\text{sc}} \leq \dots \leq \epsilon_{N_x N_k}^{\text{sc}}.$$

L'énergie cinétique et la densité de l'état fondamental s'expriment alors en fonction des C_μ^n sous la forme

$$\begin{aligned} T_{\text{GS}}^{\text{sc}} &= \text{Tr}_{L^2_{\text{per}}(\Lambda_{N_k})} \left(-\frac{1}{2} \Delta(\gamma^0)^{\text{sc}} \right) = 2\pi^2 \sum_{n=1}^{N_e N_k} \sum_{\mu=1}^{N_x N_k} \frac{(\mu - N_x \tilde{N}_k)^2}{N_k^2} |C_\mu^n|^2 \\ \rho^{\text{sc}}(l) &= (\gamma^0)^{\text{sc}}(x^{\text{sc}}(l), x^{\text{sc}}(l)), \\ &= \sum_{n=1}^{N_e N_k} |\psi_n^{\text{sc}}(x^{\text{sc}}(l))|^2, \\ &= \frac{1}{N_k} \sum_{n=1}^{N_e N_k} \left| \sum_{\mu=1}^{N_x N_k} C_\mu^n e^{2i\pi(\mu - N_x \tilde{N}_k)(l-1)/(N_x N_k)} \right|^2, \\ &= \frac{1}{N_k} \sum_{n=1}^{N_e N_k} \left| \sum_{\mu=1}^{N_x N_k} C_\mu^n e^{2i\pi(\mu-1)(l-1)/(N_x N_k)} \right|^2, \\ &= N_k N_x^2 \sum_{n=1}^{N_e N_k} |[C^n]^{\text{IFFT}}(l)|^2. \end{aligned}$$

Remarque 2. Si le potentiel \mathcal{W} a la propriété supplémentaire d'être \mathbb{Z} -périodique, la matrice densité du problème de supercellule peut se récrire en fonction des $(C^{n,k})$ obtenus en résolvant le problème de cellule

$$\begin{aligned} (\gamma^0)^{\text{sc}}(x, x') &= \gamma^0(x, x'), \\ &= \frac{1}{N_k} \sum_{n=1}^{N_e} \sum_{k=1}^{N_k} \sum_{j, j'=1}^{N_x} C_j^{n,k} \overline{C_{j'}^{n,k}} e^{2i\pi(j - N_j - 1 + (k - \tilde{N}_k)/N_k)x} e^{-2i\pi(j' - N_j - 1 + (k - \tilde{N}_k)/N_k)x'}, \\ &= \frac{1}{N_k} \sum_{n=1}^{N_e} \sum_{k=1}^{N_k} \sum_{j, j'=1}^{N_x} C_j^{n,k} \overline{C_{j'}^{n,k}} e^{2i\pi((j-1)N_k + k - N_x \tilde{N}_k)x/N_k} e^{2i\pi((j'-1)N_k + k - N_x \tilde{N}_k)x'/N_k}, \\ &= \sum_{n=1}^{N_e} \sum_{k=1}^{N_k} \sum_{j, j'=1}^{N_x} C_j^{n,k} \overline{C_{j'}^{n,k}} f_{(j-1)N_k + k}(x) \overline{f_{(j'-1)N_k + k}(x')}. \end{aligned}$$

3.2.2 Calcul du potentiel de Yukawa associé à une densité donnée

Soit ρ une densité de charge $N_k \mathbb{Z}$ -périodique, appartenant à l'espace Y_{N_x, N_k} et représentée en pratique par la suite $(\rho(l))_{1 \leq l \leq N_x N_k}$ de ses valeurs en les points $(x^{\text{sc}}(l))_{1 \leq l \leq N_x N_k}$.

Le potentiel V_ρ^{sc} est l'unique solution dans $H_{\text{per}}^1(\Lambda_{N_k})$ de l'équation de Poisson-Boltzmann linéarisée

$$-V'' + \kappa^2 V = 4\pi\rho.$$

Comme

$$\rho(x) = \sum_{\mu=1}^{N_x N_k} \hat{\rho}_\mu f_\mu(x),$$

il vient

$$V_\rho^{\text{sc}}(x) = \sum_{\mu=1}^{N_x N_k} \hat{V}_\mu f_\mu(x),$$

où

$$\hat{V}_\mu = 4\pi \frac{\hat{\rho}_\mu}{\kappa^2 + 4\pi^2(\mu - N_x \tilde{N}_k)^2 / N_k^2}. \quad (3.6)$$

Posons $V(l) = V_\rho^{\text{sc}}(x^{\text{sc}}(l))$. Comme nous l'avons vu à la section précédente, nous aurons besoin, pour assembler la matrice de l'Hamiltonien de champ moyen associé à ρ , des valeurs $(V^{\text{FFT}}(m))_{1 \leq m \leq N_x N_k}$ de la transformée de Fourier discrète de la suite $(V(l))_{1 \leq l \leq N_x N_k}$.

Soit $(\rho^{\text{FFT}}(m))_{1 \leq m \leq N_x N_k}$ la transformée de Fourier discrète de la suite $(r(l))_{1 \leq l \leq N_x N_k}$ des valeurs de ρ en les points de la grille. On a

$$\begin{aligned} \hat{\rho}_\mu &= \int_{\Gamma} \rho(x) \overline{f_\mu(x)}, \\ &\simeq \frac{1}{N_x N_k^{3/2}} \sum_{l=1}^{N_x N_k} \rho \left(\frac{l-1}{N_x} \right) e^{-2i\pi(\mu - N_x \tilde{N}_k)(l-1)/(N_x N_k)}, \\ &= \sum_{l=1}^{N_x N_k} r(l) e^{-2i\pi(\mu - N_x \tilde{N}_k)(l-1)/(N_x N_k)}. \end{aligned}$$

Donc

$$\forall N_x \tilde{N}_k \leq \mu \leq N_x N_k, \quad \hat{\rho}_\mu = \frac{1}{N_x N_k^{3/2}} \rho^{\text{FFT}}(\mu - N_x \tilde{N}_k + 1)$$

et de même

$$\forall N_x \tilde{N}_k \leq \mu \leq N_x N_k, \quad \hat{V}_\mu = \frac{1}{N_x N_k^{3/2}} V^{\text{FFT}}(\mu - N_x \tilde{N}_k + 1),$$

ce qui entraîne en utilisant (3.6)

$$\forall N_x \tilde{N}_k \leq \mu \leq N_x N_k, \quad V^{\text{FFT}}(\mu - N_x \tilde{N}_k + 1) = 4\pi \frac{\rho^{\text{FFT}}(\mu - N_x \tilde{N}_k + 1)}{\kappa^2 + 4\pi^2(\mu - N_x \tilde{N}_k)^2 / N_k^2}$$

et donc

$$\forall 1 \leq m \leq N_x \tilde{N}_k + 1, \quad V^{\text{FFT}}(m) = 4\pi \frac{\rho^{\text{FFT}}(m)}{\kappa^2 + 4\pi^2(m-1)^2 / N_k^2}. \quad (3.7)$$

Comme le potentiel V_ρ^{sc} est réel, on a ensuite pour tout $1 \leq m \leq N_x N_k - 1$,

$$\begin{aligned} V^{\text{FFT}}(N_x N_k - m + 1) &= \sum_{l=1}^{N_x N_k} V(l) e^{-2i\pi(N_x N_k - m)(l-1)/(N_x N_k)}, \\ &= \sum_{l=1}^{N_x N_k} V(l) e^{2i\pi m(l-1)/(N_x N_k)}, \\ &= \overline{\sum_{l=1}^{N_x N_k} V(l) e^{-2i\pi m(l-1)/(N_x N_k)}}, \\ &= \overline{V^{\text{FFT}}(m + 1)}. \end{aligned}$$

D'où

$$\forall N_x \tilde{N}_k + 2 \leq m \leq N_x N_k, \quad V^{\text{FFT}}(m) = \overline{V^{\text{FFT}}(N_x N_k - m + 2)}. \quad (3.8)$$

Pour résumer, en ayant des pas de discrétisation fixés et les valeurs de la densité en les points de la grille d'espace, on obtient les valeurs "FFT" du potentiel \mathcal{W} , en deux étapes : d'abord le calcul de la transformée FFT de ρ puis le calcul de W^{FFT} par les formules (3.7) et (3.8).

3.2.3 Algorithme ODA pour le problème de supercellule

La résolution du problème de supercellule s'effectue en utilisant une adaptation de l'algorithme ODA permettant de résoudre le problème de cellule (cf. section 3.1.3). Les modifications à apporter étant mineures, nous ne les détaillons pas ici.

3.3 Résolution numérique du modèle de défaut

Pour obtenir des approximations variationnelles de notre modèle de défaut, nous allons devoir construire des sous-espaces de dimension finie des espaces $\mathfrak{H}_- = \text{Ran}(\gamma_{\text{per}}^0)$ (espace occupé) et $\mathfrak{H}_+ = \text{Ran}(1 - \gamma_{\text{per}}^0)$ (espace virtuel), engendrés par des fonctions localisées autour du défaut. Dans la section 3.3.1, nous expliquons comment construire de tels sous-espaces à l'aide de fonctions de Wannier maximale-ment localisées, et dans la section 3.3.2, nous montrons comment calculer numériquement de telles fonctions en dimension 1. Nous décrivons à la section 3.3.3 le problème de dimension finie résultant de l'approximation variationnelle de notre modèle de défaut, et nous présentons quelques résultats numériques à la section 3.3.4.

3.3.1 Fonctions de Wannier

Soit ρ_{per}^0 la densité de l'état fondamental du problème périodique et

$$H_{\text{per}}^0 = -\frac{1}{2} \frac{d^2}{dx^2} + V_{\text{nuc}} + V_{\rho_{\text{per}}^0}.$$

Soit $(\psi_{n,\xi})_{n \geq 1, \xi \in \Gamma^*}$ les orbitales de Bloch de l'Hamiltonien H_{per}^0 , i.e.

$$-\frac{1}{2} \frac{d^2 \psi_{n,\xi}}{dx^2} + V_{\text{nuc}} \psi_{n,\xi} + V_{\rho_{\text{per}}^0} \psi_{n,\xi} = \epsilon_{n,\xi} \psi_{n,\xi}, \quad \psi_{n,\xi} \in L_{\xi}^2(\Gamma),$$

$$\int_{\Gamma} \psi_{m,\xi} \psi_{n,\xi}^* = \delta_{mn}, \quad \epsilon_{1,\xi} \leq \epsilon_{2,\xi} \leq \dots$$

Soit $n \in \mathbb{N}^*$. On définit la fonction de Wannier ‘‘mère’’ χ_n par

$$\chi_n(x) = \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \psi_{n,\xi}(x) d\xi.$$

On pose ensuite pour tout $r \in \mathcal{R}$,

$$\chi_{n,r}(x) = \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \psi_{n,\xi}(x) e^{-i\xi r} d\xi.$$

Les fonctions $(\chi_{n,r})_{r \in \mathcal{R}}$ vérifient

$$\forall r \in \mathcal{R}, \quad \chi_{n,r}(x) = \chi_n(x - r) \quad \text{p.p.}$$

Par ailleurs, $(\chi_{n,r})_{r \in \mathcal{R}}$ forme une famille orthonormée. En effet,

$$\begin{aligned} \langle \chi_{n,r}, \chi_{n,r'} \rangle_{L^2(\mathbb{R}^3)} &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \langle (\chi_{n,r})_{\xi}, (\chi_{n,r'})_{\xi} \rangle_{L_{\xi}^2(\Gamma)} d\xi, \\ &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \int_{\Gamma} \overline{\psi_{n,\xi}(x) e^{-i\xi r}} \psi_{n,\xi}(x) e^{-i\xi r'} d\xi dx, \\ &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \langle \psi_{n,\xi}, \psi_{n,\xi} \rangle_{L_{\xi}^2(\Gamma)} e^{-i\xi(r'-r)} d\xi \\ &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} e^{-i\xi(r'-r)} d\xi \\ &= \delta_{r,r'}. \end{aligned}$$

On peut montrer que $(\chi_{n,r})_{r \in \mathcal{R}}$ constitue en fait une base hilbertienne de $\text{Ran}(\Pi_n)$ où Π_n est le projecteur spectral sur la n -ième bande de H_{per}^0 .

Il en résulte que $(\chi_{n,r})_{1 \leq n \leq N_e, r \in \mathcal{R}}$ et $(\chi_{n,r})_{n > N_e, r \in \mathcal{R}}$ constituent des bases hilbertiennes de \mathfrak{H}_- et \mathfrak{H}_+ respectivement.

Numériquement, on a intérêt à ce que les fonctions de base soient les plus localisées possible (de façon à ce que les matrices hamiltoniennes soient les plus creuses possibles). C'est la raison pour laquelle on fait appel au procédé de localisation suivant.

On commence par remarquer que, les fonctions $\psi_{n,\xi}$ étant définies à une phase globale près, on peut remplacer $\psi_{n,\xi}$ par $\psi_{n,\xi}^{\alpha_n} = \psi_{n,\xi} e^{i\alpha_n(\xi)}$ où α_n est une fonction de Γ^* à

valeurs dans \mathbb{R} . Dans le cas particulier où l'on prend $\alpha_n(\xi) = \xi r_0$ avec $r_0 \in \mathcal{R}$, on obtient

$$\begin{aligned}\chi_n^{\alpha_n} &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \psi_{n,\xi}^{\alpha_n}(x) d\xi, \\ &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \psi_{n,\xi}(x) e^{i\alpha_n(\xi)} d\xi, \\ &= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \psi_{n,\xi}(x) e^{i\xi r_0} d\xi, \\ &= \chi_{n,-r_0}(x) = \chi_n(x + r_0).\end{aligned}$$

La famille $(\chi_{n,r})_{r \in \mathcal{R}}$ de fonctions de Wannier reste donc inchangée : la fonction mère a simplement été translatée. En revanche, pour un α_n quelconque la forme de la fonction mère est modifiée.

Pour avoir des fonctions de Wannier les plus localisées possible, il faut choisir la fonction $\alpha_n(\xi)$ de manière à minimiser un critère de délocalisation.

Dans [64], Marzari et Vanderbilt ont proposé comme critère de délocalisation la variance

$$\Omega(\alpha_n) = \langle \chi_n^{\alpha_n} | |x|^2 | \chi_n^{\alpha_n} \rangle - |\langle \chi_n^{\alpha_n} | x | \chi_n^{\alpha_n} \rangle|^2$$

où x désigne l'opérateur position dans $L^2(\mathbb{R}^3)$. Les fonctions de Wannier obtenues grâce à ce critère sont appelées *Maximally Localized Wannier Functions (MLWF)*.

La construction ci-dessus s'effectue bande par bande. Pour obtenir des bases de \mathfrak{H}_- et \mathfrak{H}_+ encore mieux localisées, on peut procéder comme suit. Soit $(B_n)_{n \in \mathcal{N}}$ un ensemble fini de bandes. Nous pouvons associer à n'importe quelle famille $(U(\xi))_{\xi \in \Gamma^*}$ de matrices unitaires de taille $|\mathcal{N}| \times |\mathcal{N}|$ les fonctions de Wannier généralisées définies par

$$\forall n \in \mathcal{N}, \quad \chi_n^U(x) = \frac{1}{|\Gamma^*|} \int_{\Gamma^*} \sum_{m \in \mathcal{N}} U_{nm}(\xi) \psi_{m,\xi}(x) d\xi.$$

Alors $\chi_n^U \in \oplus_{m \in \mathcal{N}} \text{Ran}(\Pi_m)$ et $(\chi_n^U(\cdot - r))_{n \in \mathcal{N}, r \in \mathbb{Z}}$ forme une base hilbertienne de $\oplus_{m \in \mathcal{N}} \text{Ran}(\Pi_m)$. La famille $(U(\xi))_{\xi \in \Gamma^*}$ peut être choisie de manière à minimiser un critère de délocalisation et on obtient ainsi une base de *Maximally Localized (Generalized) Wannier Functions*.

3.3.2 Construction de MLWF en dimension 1

Pour construire des MLWF pour notre problème, nous suivons l'approche décrite dans [67]. Détaillons cette approche pour la construction de MLWF relatives à l'espace occupé.

Nous partons du résultat d'un calcul de supercellule pour le cristal parfait donnant une approximation

$$\gamma^{\text{sc}}(x, x') = \sum_{\mu, \mu'=1}^{N_x N_k} P_{\mu\mu'} f_\mu(x) f_{\mu'}(x')$$

de la matrice densité γ_{per}^0 . Notons que γ^{sc} peut être obtenu de manière plus économique en temps de calcul par un calcul sur la cellule unité avec une discrétisation uniforme de taille N_k de la zone de Brillouin. Notons également que

$$\text{Tr}_{L_{\text{per}}^2(\Lambda_{N_k})}(\gamma^{\text{sc}}) = N_e N_k := N_0.$$

La méthode proposée dans [67] consiste à diagonaliser l'opérateur

$$\Xi = \gamma^{\text{sc}} \sin\left(\frac{2\pi x}{N_k}\right) \gamma^{\text{sc}}.$$

Cet opérateur est en effet une approximation de l'opérateur $\gamma_{\text{per}}^0 x \gamma_{\text{per}}^0$ dont on peut montrer que les fonctions propres sont des MLWF relatives à l'espace occupé.

Les coefficients de la matrice de Ξ dans la base $(f_\mu)_{1 \leq \mu \leq N_x N_k}$ sont donnés par

$$\Xi_{\mu, \mu'} = \sum_{\mu_1, \mu_2=1}^{N_x N_k} P_{\mu \mu_1} S_{\mu_1 \mu_2} P_{\mu_2 \mu'}$$

où

$$\begin{aligned} S_{\mu_1 \mu_2} &= \int_{\Lambda_{N_k}} \sin\left(\frac{2\pi x}{N_k}\right) \overline{f_{\mu_1}(x)} f_{\mu_2}(x) dx, \\ &= \frac{1}{N_k} \int_{\Lambda_{N_k}} \sin\left(\frac{2\pi x}{N_k}\right) e^{2i\pi(\mu_2 - \mu_1)x/N_k} dx, \\ &= \frac{i}{N_k} \int_{\Lambda_{N_k}} \sin\left(\frac{2\pi x}{N_k}\right) \sin\left(\frac{2\pi(\mu_2 - \mu_1)x}{N_k}\right) dx, \\ &= i \int_{\Gamma} \sin(2\pi y) \sin(2\pi(\mu_2 - \mu_1)y) dy, \\ &= \frac{i}{2} \int_{\Gamma} (\cos(2\pi y(\mu_2 - \mu_1 - 1)) - \cos(2\pi y(\mu_2 - \mu_1 + 1))). \end{aligned}$$

Les intégrales précédentes sont non nulles si et seulement si $\mu_2 - \mu_1 - 1 = 0$ ou $\mu_2 - \mu_1 + 1 = 0$. D'où

$$S_{\mu_1 \mu_2} = \begin{cases} i/2 & \text{si } \mu_2 - \mu_1 = 1 \\ -i/2 & \text{si } \mu_2 - \mu_1 = -1. \end{cases}$$

Il en découle que

$$\Xi_{\mu, \mu'} = \frac{i}{2} \sum_{\mu_1=1}^{N_x N_k - 1} (P_{\mu \mu_1} P_{\mu_1 + 1, \mu'} - P_{\mu, \mu_1 + 1} P_{\mu_1 \mu'}).$$

On diagonalise ensuite cette matrice et on note $(M^n)_{1 \leq n \leq N_e}$ un système de N_e vecteurs propres de Ξ associés aux N_e plus petites valeurs propres. Les fonctions

$$\chi_n^{\text{LW}}(x) = \sum_{\mu=1}^{N_x N_k} M_\mu^n f_\mu(x)$$

seront des approximations des N_e MLWF de l'espace occupé localisées autour du point $3\tilde{N}_k/2$, le minimum de la fonction $\sin\left(\frac{2\pi x}{N_k}\right)$.

Nous pouvons aussi calculer facilement les valeurs de ces fonctions aux points de la grille d'espace par FFT :

$$\begin{aligned}\chi_n^{\text{LW}}(x(l)) &= \frac{1}{N_k^{1/2}} \sum_{\mu=1}^{N_x N_k} M_\mu^n e^{2i\pi(\mu - N_x \tilde{N}_k)(l-1)/(N_x N_k)}, \\ &= \frac{1}{N_k^{1/2}} \left(\sum_{\mu=1}^{N_x N_k} M_\mu^n e^{2i\pi(\mu-1)(l-1)/(N_x N_k)} \right) e^{-2i\pi(N_x \tilde{N}_k - 1)(l-1)/(N_x N_k)}, \\ &= N_x N_k^{1/2} [M^n]^{\text{FFT}}(l) e^{-2i\pi(N_x \tilde{N}_k - 1)(l-1)/(N_x N_k)}.\end{aligned}$$

Ensuite, on en déduit par translation une base de MLWF de l'espace occupé : pour tout $r \in \mathbb{Z}$, on pose

$$\begin{aligned}\chi_{n,r}^{\text{LW}}(x) &= \chi_n^{\text{LW}}(x - r) \\ &= \sum_{\mu=1}^{N_x N_k} M_\mu^n f_\mu(x - r) \\ &= \sum_{\mu=1}^{N_x N_k} M_\mu^n e^{-2i\pi(\mu - N_x \tilde{N}_k)r/N_k} f_\mu(x) \\ &= \sum_{\mu=1}^{N_x N_k} M_\mu^{n,r} f_\mu(x),\end{aligned}$$

où $M_\mu^{n,r} = M_\mu^n e^{-2i\pi(\mu - N_x \tilde{N}_k)r/N_k}$.

En sélectionnant dans la base des $(\chi_{n,r}^{\text{LW}})$ les m_- fonctions les plus proches du défaut, on pourra construire un espace d'approximation de \mathfrak{H}_- .

En remplaçant γ^{sc} par une approximation du projecteur spectral sur les N_a bandes virtuelles de plus basse énergie, on pourra construire numériquement une base de MLWF relatives à ces bandes. Et de même, on obtiendra des espaces d'approximation de \mathfrak{H}_+ en sélectionnant parmi ces MLWF, celles qui sont les plus proches du défaut.

3.3.3 Approximation variationnelle du modèle de défaut

Considérons une perturbation $\nu(x)$ de la charge nucléaire périodique du cristal parfait et intéressons-nous au problème de minimisation avec contrainte de charge, i.e. à la transposition au modèle 1D dont il est question dans ce chapitre du modèle (2.25) étudié d'un point de vue théorique à la section 2.3.2.

Pour discrétiser ce problème, il faut prendre en compte la contrainte que la matrice Q doit appartenir à l'espace \mathcal{K} . Pour ce faire, on construit comme expliqué à la section précédente, des espaces d'approximation V_-^h et V_+^h de \mathfrak{H}_- et \mathfrak{H}_+ respectivement,

engendrés par des MLWF. Soit m_{\pm} la dimension de V_{\pm}^h , $N_b = m_- + m_+$, et $V^h = V_-^h \oplus V_+^h$.

Soit $(\varphi_1, \dots, \varphi_{m_-})$ (resp. $(\varphi_{m_-+1}, \dots, \varphi_{N_b})$) une base non nécessairement orthonormée de V_-^h (resp. de V_+^h). Quand on discrétise le problème de minimisation avec contrainte de charge dans cette base, on restreint l'espace variationnel aux opérateurs Q de rang fini de la forme

$$Q = \sum_{i,j=1}^{N_b} Q_{ij}^h |\varphi_i\rangle \langle \varphi_j|, \quad (3.9)$$

où la matrice Q^h appartient à l'ensemble convexe

$$\mathcal{Q}^h = \{Q^h \in \mathbb{C}^{N_b \times N_b}, (Q^h)^* = Q^h, -\gamma^0 \leq Q^h \leq S^{-1} - \gamma^0, \text{Tr}(SQ^h) = q\}.$$

Dans l'expression ci-dessus, S désigne la matrice de recouvrement

$$S = \begin{pmatrix} S_{m_-} & 0 \\ 0 & S_{m_+} \end{pmatrix}$$

où $[S_{m_-}]_{ij} = (\varphi_i, \varphi_j)_{L^2}$ et $[S_{m_+}]_{ij} = (\varphi_{m_-+i}, \varphi_{m_-+j})_{L^2}$ et γ^0 désigne la matrice qui représente γ_{per}^0 dans la base $(\varphi_i)_{1 \leq i \leq N_b}$:

$$\gamma^0 = \begin{pmatrix} S_{m_-}^{-1} & 0 \\ 0 & 0 \end{pmatrix} = S^{-1/2} D_0 S^{-1/2} \quad \text{avec } D_0 = \begin{pmatrix} I_{m_-} & 0 \\ 0 & 0 \end{pmatrix}.$$

Le fait que S soit diagonale par blocs vient du fait que les espaces V_-^h et V_+^h sont orthogonaux. De même, la matrice de H_{per}^0 dans cette base est de la forme

$$H^0 = \begin{pmatrix} H_-^0 & 0 \\ 0 & H_+^0 \end{pmatrix}$$

La densité et l'énergie d'un opérateur Q de la forme (3.9) sont données respectivement par

$$\rho_Q(x) = \sum_{i,j=1}^{N_b} Q_{ij}^h \varphi_i(x) \overline{\varphi_j(x)}$$

et

$$E(Q) = \text{Tr}(\overline{H}Q^h) + \frac{1}{2} \text{Tr}(J(Q^h)Q^h),$$

avec $[\overline{H}]_{ij} = H^0 + D_{\kappa}(\nu, \overline{\varphi_i} \varphi_j)$ et $[J(Q^h)]_{ij} = D_{\kappa}(\rho_{Q^h}, \overline{\varphi_i} \varphi_j)$.

Afin de pouvoir résoudre ce problème par l'algorithme ODA présenté ci-dessus, nous effectuons le changement de variable

$$D = D_0 + S^{1/2} Q^h S^{1/2}.$$

Pour $Q^h \in \mathcal{Q}^h$, D appartient à l'ensemble

$$\{D \in \mathbb{C}^{N_b \times N_b}, D^* = D, 0 \leq D \leq I_{N_b}, \text{Tr}(D) = m_- + q\}.$$

L'expression de l'énergie en fonction de D est alors

$$\begin{aligned} \epsilon(D) &= E(Q^h), \\ &= E(S^{-1/2}(D - D_0)S^{-1/2}), \\ &= E(S^{-1/2}DS^{-1/2} - \gamma^0), \\ &= \text{Tr}(\overline{H}(S^{-1/2}DS^{-1/2} - \gamma^0)) + \frac{1}{2}\text{Tr}(J(S^{-1/2}DS^{-1/2} - \gamma^0)(S^{-1/2}DS^{-1/2} - \gamma^0)), \\ &= \text{Tr}(\overline{H}(S^{-1/2}DS^{-1/2}) - \text{Tr}(\overline{H}\gamma^0) + \frac{1}{2}\text{Tr}(J(S^{-1/2}DS^{-1/2})S^{-1/2}DS^{-1/2}) \\ &\quad - \text{Tr}(J(\gamma^0)S^{-1/2}DS^{-1/2}) + \frac{1}{2}\text{Tr}(J(\gamma^0)\gamma^0)), \\ &= \text{Tr}(\tilde{H}D) + \frac{1}{2}\text{Tr}(S^{-1/2}J(S^{-1/2}DS^{-1/2})S^{-1/2}D) - E_0, \end{aligned}$$

avec

$$\tilde{H} = S^{-1/2}(\overline{H} - J(\gamma^0))S^{-1/2}$$

et

$$E_0 = \text{Tr}(\overline{H}\gamma^0) - \frac{1}{2}\text{Tr}(J(\gamma^0)\gamma^0).$$

Nous passons à présent à la dernière section de ce chapitre où nous donnons quelques résultats numériques obtenus en appliquant les méthodes et approximations décrites jusqu'à présent.

3.3.4 Quelques tests numériques

Pour l'application numérique, nous avons choisi comme valeurs des paramètres

$$A_{ee} = 10, A_{ke} = 1 \text{ et } \kappa = 5$$

de manière à ce que les énergies cinétique et potentielle de l'état fondamental soient du même ordre de grandeur.

Nous avons pris différents défauts. Ln premier défaut,

$$\nu = (Z - 1)\delta_{0.25} - Z\delta_0,$$

qui correspond à bouger un noyau et diminuer sa charge d'une unité, a été traité dans notre article [58]. Nous montrons au Chapitre 5 les résultats obtenus pour $Z = 2$, $N_e = 2$, $N_a = 2$ et $N_b = 28$, comparés à un calcul habituel effectué avec une base de 1224 ondes planes. Outre la figure montrant la densité ρ_Q nous avons exhibé le tracé des modules de MLWF mères correspondant aux deux bandes occupées (bandes de valence) et aux deux bandes virtuelles (bandes de conduction) de plus basse énergie.

Nous prenons à présent $Z = 60$, $N_e = 2$, $N_a = 4$ et $N_b = 90$. Dans la figure ci-dessous, le premier cadre montre le potentiel nucléaire du défaut et le second la différence

entre les densités des états fondamentaux avec et sans défaut, obtenue par deux méthodes différentes : la résolution de notre modèle de défaut (trait continu) et un calcul de supercellule (en pointillés).

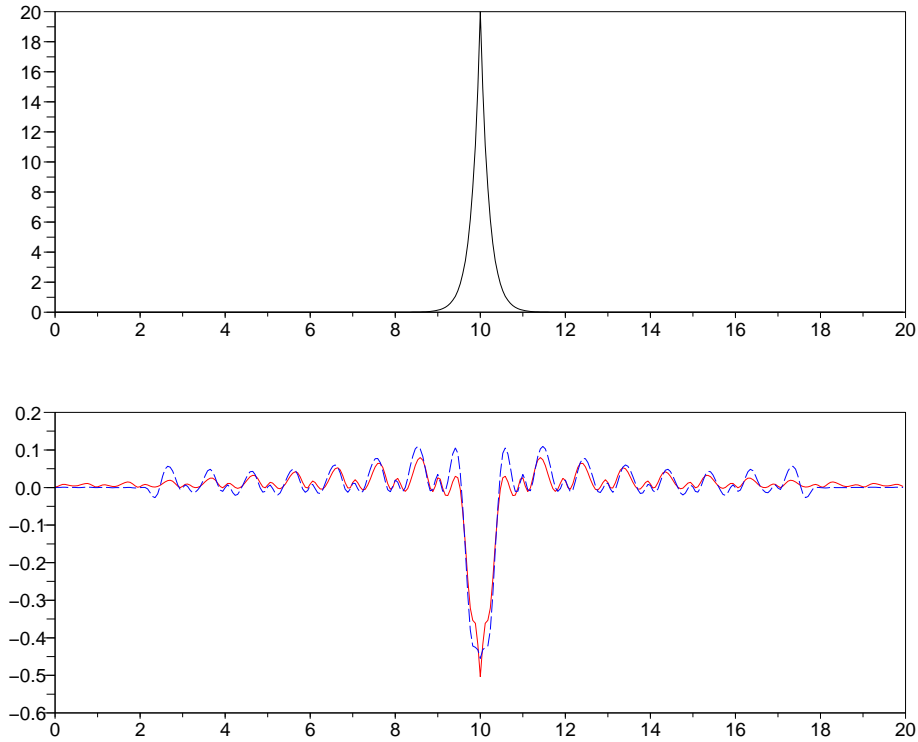


FIG. 3.1 – Tracé des différences de potentiels (dessin du haut) et de densités (dessin du bas) de la mer de Fermi entre l'état perturbé et l'état au repos.

Les résultats obtenus sont satisfaisants, même avec un nombre de fonctions de Wannier largement inférieur au nombre d'ondes planes nécessaires pour un calcul de supercellule. Pour améliorer la précision de la résolution numérique de notre modèle de défaut, on pourra compléter la description de l'espace virtuel en "augmentant" la base de MLWF en projetant sur \mathfrak{H}_+ des orbitales atomiques adaptées au défaut.

Chapitre 4

Local defects in periodic crystals

"Crystals are like people; it's the defects in them that make them interesting."

Sir Charles Frank

In this chapter, we are going to develop and detail the theory exposed in [7], based on formal analogies between the Fermi sea of a perturbed crystal and the Dirac sea in Quantum Electrodynamics in the presence of an external electrostatic field. Recently, the latter model was extensively studied by Hainzl, Lewin, Séré and Solovej in the Hartree-Fock approximation [22–25], based on ideas from Chaix and Iracane [13] (see also [1, 14]). Using and adapting these methods, we were able to propose a new mathematical approach for the self-consistent description of a crystal in the presence of local defects.

We focused on the *reduced Hartree-Fock* (rHF) model in which the so-called *exchange term* is neglected. To further simplify the mathematical formulas, we will not explicitly take the spin variable into account and will assume that the host crystal is cubic with a single atom of charge Z per unit cell. The arguments below can be easily extended to the general case.

In this study, the main object of interest is the so-called *density matrix* of the electrons. This is a self-adjoint operator $0 \leq \gamma \leq 1$ acting on the one-body space $L^2(\mathbb{R}^3)$. When γ has a finite rank, it models a finite number of electrons. In the periodic case, the ground state density matrix γ_{per}^0 has an infinite rank (it describes infinitely many electrons) and commutes with the translations of the lattice. We will see in the sequel that the ground state density matrix of a crystal with a local defect can be written as $\gamma = \gamma_{\text{per}}^0 + Q$, where Q is a compact perturbation of the periodic density matrix γ_{per}^0 of the reference perfect crystal.

In each of the above three cases (finite number of electrons, perfect crystal, defective crystal), the ground state density matrix can be obtained by minimizing some nonlinear energy functional on a set of admissible density matrices. In the case of a crystal with a local defect, the perturbation Q is a minimizer of some nonlinear minimization problem set in the whole space \mathbb{R}^3 , with a possible lack of compactness at infinity. The main unusual feature compared to standard variational problems is that Q is a self-adjoint operator of infinite rank. This was already the case in [22–25].

This chapter is organized as follows. We will first recall the definition of the reduced Hartree-Fock model for a finite number of electrons. The second section will be devoted to the definition of the model for the infinite periodic crystal, following mainly [11, 12] (but providing some additional material compared to what was done in [11, 12]). Then, we will describe the model for the crystal with local defects which takes the perfect crystal as reference and finally prove that this model is the thermodynamic limit of the supercell model.

Almost all the proofs will be self-consistent, using some estimates and some general results gathered in the previous chapter and in the appendix. Unless otherwise stated the operators in the proofs are considered as operators on $L^2(\mathbb{R}^3)$.

4.1 The reduced Hartree-Fock model for N electrons

We start by recalling the definition of the reduced Hartree-Fock model [46] for a finite number of electrons. Note that the *reduced* Hartree-Fock model should not be confused with the *restricted* Hartree-Fock model commonly used in numerical simulations (see e.g. [16]). We consider a system containing N nonrelativistic quantum electrons and a set of nuclei having a density of charge ρ_{nuc} . If for instance there are K nuclei of charges $z_1, \dots, z_K \in \mathbb{N} \setminus \{0\}$ located at $R_1, \dots, R_K \in \mathbb{R}^3$, then

$$\rho_{\text{nuc}}(x) := \sum_{k=1}^K z_k m_k(x - R_k),$$

where m_1, \dots, m_K are positive measures on \mathbb{R}^3 of total mass one. Point-like nuclei would correspond to $m_k = \delta$ (the Dirac measure) but for convenience we shall deal with smeared nuclei in the sequel, i.e. we assume that for all $k = 1 \dots K$, m_k is a smooth nonnegative function such that $\int_{\mathbb{R}^3} m_k = 1$.

The energy of the whole system in the reduced Hartree-Fock model reads [12, 46]

$$\mathcal{E}_{\rho_{\text{nuc}}}^{\text{rHF}}(\gamma) = \text{Tr} \left(-\frac{1}{2} \Delta \gamma \right) + \frac{1}{2} D(\rho_\gamma - \rho_{\text{nuc}}, \rho_\gamma - \rho_{\text{nuc}}). \quad (4.1)$$

As mentioned above, we have chosen a system of units such that $\hbar = m = e = \frac{1}{4\pi\epsilon_0} = 1$ where m and e are respectively the mass and the charge of an electron, \hbar is the reduced Planck constant and ϵ_0 is the dielectric permittivity of the vacuum. The first term in the right-hand side of (4.1) is the kinetic energy of the electrons and $D(\cdot, \cdot)$ is the classical Coulomb interaction, which reads for f and g in $L^{6/5}(\mathbb{R}^3)$ as

$$D(f, g) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{f(x)g(y)}{|x-y|} dx dy = 4\pi \int_{\mathbb{R}^3} \frac{\widehat{f}(k)\widehat{g}(k)}{|k|^2} dk \quad (4.2)$$

where \widehat{f} denotes the Fourier transform of f . In this mean-field model, the state of the N electrons is described by the one-body density matrix γ , which is an element of the following class

$$\mathcal{P}^N = \left\{ \gamma \in \mathcal{B}(L^2(\mathbb{R}^3)) \mid 0 \leq \gamma \leq 1, \text{Tr}(\gamma) = N, \text{Tr} \left(\sqrt{-\Delta} \gamma \sqrt{-\Delta} \right) < \infty \right\}.$$

We recall that $\mathcal{B}(\mathfrak{H})$ denotes the space of bounded self-adjoint operators acting on the Hilbert space \mathfrak{H} . We also define $\text{Tr}((-\Delta)\gamma) := \text{Tr}(\sqrt{-\Delta}\gamma\sqrt{-\Delta})$ which makes sense when $\gamma \in \mathcal{P}^N$. The set \mathcal{P}^N is the closed convex hull of the set of orthogonal projectors of rank N acting on $L^2(\mathbb{R}^3)$ and having a finite kinetic energy. Each such projector $\gamma = \sum_{i=1}^N |\varphi_i\rangle\langle\varphi_i|$ is the density matrix of a Hartree-Fock state

$$\Psi = \varphi_1 \wedge \dots \wedge \varphi_N \quad (4.3)$$

in the usual N -body space of fermionic wavefunctions with finite kinetic energy $\bigwedge_{i=1}^N H^1(\mathbb{R}^3)$.

Any $\gamma \in \mathcal{P}^N$ is non-negative, self-adjoint and has a finite trace (equal to N). Consequently, it is a compact operator and therefore it can be diagonalized in some orthonormal basis $(\varphi_i)_{i \in \mathbb{N} \setminus \{0\}}$ of $L^2(\mathbb{R}^3)$

$$\gamma = \sum_{i=1}^{\infty} n_i |\varphi_i\rangle \langle \varphi_i|.$$

Since $0 \leq \gamma \leq 1$ and $\text{Tr}(\gamma) = N$, the eigenvalues n_i of γ are all in the range $[0, 1]$ and sum to N . Lastly the condition $\text{Tr}(\sqrt{-\Delta}\gamma\sqrt{-\Delta})$ implies that $\varphi_i \in H^1(\mathbb{R}^3)$ for all $i \in \mathbb{N} \setminus \{0\}$.

The function ρ_γ appearing in (4.1) is the density associated with the operator γ , defined by $\rho_\gamma(x) = \gamma(x, x)$ where $\gamma(x, y)$ is the kernel of the trace class operator γ . We have

$$\rho_\gamma(x) = \gamma(x, x) = \sum_{i=1}^{\infty} n_i \varphi_i(x)^2 \geq 0.$$

All the terms are defined since $0 \leq n_i \leq 1$ and $\sum_{i=1}^{\infty} n_i = \text{Tr}(\gamma) = N$. For all $\gamma \in \mathcal{P}^N$, one has $\rho_\gamma \geq 0$ and $\sqrt{\rho_\gamma} \in H^1(\mathbb{R}^3)$. As $H^1(\mathbb{R}^3) \hookrightarrow L^2 \cap L^6(\mathbb{R}^3)$ and $\sqrt{\rho_\gamma} \in L^2 \cap L^6(\mathbb{R}^3)$, it holds $\rho_\gamma \in L^1 \cap L^3(\mathbb{R}^3) \hookrightarrow L^{6/5}(\mathbb{R}^3)$, hence the last term of (4.1) is well-defined.

It can be proved (see the appendix of [46]) that if $N \leq \sum_{k=1}^K z_k$ (neutral or positively charged systems), the variational problem

$$I_{\text{rHF}}(\rho_{\text{nuc}}, N) = \inf \{ \mathcal{E}_{\rho_{\text{nuc}}}^{\text{rHF}}(\gamma), \quad \gamma \in \mathcal{P}^N \} \quad (4.4)$$

has a minimizer γ and that the corresponding minimizing density ρ_γ is unique.

The Hartree-Fock model [36] is the variational approximation of the time-independent Schrödinger equation obtained by restricting the set of fermionic wavefunctions under consideration to the subset of functions of the form (4.3). The HF functional reads

$$\mathcal{E}_{\rho_{\text{nuc}}}^{\text{HF}}(\gamma) = \mathcal{E}_{\rho_{\text{nuc}}}^{\text{rHF}}(\gamma) - \frac{1}{2} \iint_{\mathbb{R}^6} \frac{|\gamma(x, y)|^2}{|x - y|} dx dy, \quad (4.5)$$

the last term being called the *exchange energy*. As the Hartree-Fock energy functional is nonconvex, there is little hope to obtain rigorous thermodynamic limits in this setting, at least with current state-of-the-art techniques. For this reason, the exchange term is often neglected in mathematical studies.

4.2 The reduced Hartree-Fock model for a perfect crystal

In this work, the nuclei are clamped, and we optimize only over the state of the electrons. In other words, we do not consider the geometry optimization problem.

Instead, we are interested in the change of the electronic state of the crystal when a local defect is introduced in the originally periodic nuclear configuration. To this end, we shall rely heavily on the rHF model for the infinite perfect crystal (with no defect) which was studied by Catto, Le Bris and Lions in [11, 12]. The latter can be obtained as the thermodynamical limit of the rHF model for finite systems which was introduced in the previous section. This will be explained in Section 4.4 below.

Let $\Gamma = [-1/2, 1/2]^3$ be the unit cell. We denote by $\Gamma^* = [-\pi, \pi]^3$ the first Brillouin zone of the lattice, and by τ_k the translation operator on $L^2_{\text{loc}}(\mathbb{R}^3)$ defined by $\tau_k u(x) = u(x - k)$. We then introduce

$$\mathcal{P}_{\text{per}} = \left\{ \gamma \in \mathcal{B}(L^2(\mathbb{R}^3)) \mid 0 \leq \gamma \leq 1, \forall k \in \mathbb{Z}^3, \tau_k \gamma = \gamma \tau_k, \right. \\ \left. \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}((1 - \Delta_\xi)^{1/2} \gamma_\xi (1 - \Delta_\xi)^{1/2}) d\xi < \infty \right\}$$

where $(\gamma_\xi)_{\xi \in \Gamma^*}$ is the Bloch waves decomposition of γ we have presented in the previous chapter (see also [12, 42]) :

$$\gamma = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \gamma_\xi d\xi, \quad \gamma_\xi \in \mathcal{B}(L^2_\xi(\Gamma)),$$

$$L^2_\xi(\Gamma) = \{u \in L^2_{\text{loc}}(\mathbb{R}^3) \mid \tau_k u = e^{-ik \cdot \xi} u, \forall k \in \mathbb{Z}^3\},$$

which corresponds to the decomposition in fibers $L^2(\mathbb{R}^3) = \int_{\Gamma^*}^\oplus d\xi L^2_\xi(\Gamma)$. For any $\gamma \in \mathcal{P}_{\text{per}}$, we denote by $\gamma_\xi(x, y)$ the integral kernel of γ_ξ . The density of γ is then the nonnegative \mathbb{Z}^3 -periodic function of $L^1_{\text{loc}}(\mathbb{R}^3) \cap L^3_{\text{loc}}(\mathbb{R}^3)$ defined as

$$\rho_\gamma(x) := \frac{1}{(2\pi)^3} \int_{\Gamma^*} \gamma_\xi(x, x) d\xi.$$

Let us underline two points : first, for any $\gamma \in \mathcal{P}_{\text{per}}^Z$,

$$\frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}(\gamma_\xi) d\xi = \int_\Gamma \rho_\gamma(x) dx,$$

which is the number of electrons per unit cell. Later we shall add the constraint that the system is neutral and restrict ourselves to states $\gamma \in \mathcal{P}_{\text{per}}$ satisfying

$$\int_\Gamma \rho_\gamma(x) dx = Z$$

where Z is the nuclear charge per unit cell.

The second point is that the periodicity condition upon the operator γ translates onto its kernel as follows :

$$\gamma(x - k, y - k) = \gamma(x, y) \quad \forall k \in \mathbb{Z}^3, \forall (x, y) \in \mathbb{R}^3 \times \mathbb{R}^3.$$

We also introduce the \mathbb{Z}^3 -periodic Green kernel of the Poisson interaction [37], denoted by G_1 and uniquely defined by

$$\begin{cases} -\Delta G_1 = 4\pi \left(\sum_{k \in \mathbb{Z}^3} \delta_k - 1 \right) \\ G_1 \text{ } \mathbb{Z}^3\text{-periodic, } \min_{\mathbb{R}^3} G_1 = 0. \end{cases}$$

The Fourier expansion of G_1 is

$$G_1(x) = c + \sum_{k \in 2\pi\mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} e^{ik \cdot x}$$

with $c = \int_{\Gamma} G_1 > 0$. The electrostatic potential associated with a \mathbb{Z}^3 -periodic density $\rho \in L^1_{\text{per}}(\Gamma) \cap L^3_{\text{per}}(\Gamma)$ is the \mathbb{Z}^3 -periodic function defined as

$$(\rho \star_{\Gamma} G_1)(x) := \int_{\Gamma} G_1(x - y) \rho(y) dy.$$

We introduce here the spaces $L^p_{\text{per}}(\Gamma) = \{u \in L^p_{\text{loc}}(\mathbb{R}^3) \mid u \text{ } \mathbb{Z}^3\text{-periodic}\}$.

We also set for any \mathbb{Z}^3 -periodic functions f and g

$$D_{G_1}(f, g) := \int_{\Gamma} \int_{\Gamma} G_1(x - y) f(x) g(y) dx dy.$$

In the sequel, we will denote by χ_I the characteristic function of the set $I \subset \mathbb{R}$ and by $\chi_I(A)$ the spectral projector on I of the self-adjoint operator A .

We recall that the periodic density of the nuclei is given by

$$\mu_{\text{per}}(x) = \sum_{R \in \mathbb{Z}^3} Z m(x - R).$$

We have assumed for simplicity that m is a nonnegative function of $C_0^\infty(\mathbb{R}^3)$ with support in Γ , and that $\int_{\mathbb{R}^3} m(x) dx = 1$. Hence $\int_{\Gamma} \mu_{\text{per}}(x) dx = Z$, the total nuclear charge in each unit cell. The periodic rHF energy is then defined for $\gamma \in \mathcal{P}_{\text{per}}$ as

$$\boxed{\mathcal{E}_{\text{per}}^0(\gamma) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_{\xi}(\Gamma)} \left(-\frac{1}{2} \Delta \gamma_{\xi} \right) + \frac{1}{2} D_{G_1}(\rho_{\gamma} - \mu_{\text{per}}, \rho_{\gamma} - \mu_{\text{per}})}. \quad (4.6)$$

Introducing

$$\mathcal{P}_{\text{per}}^Z := \left\{ \gamma \in \mathcal{P}_{\text{per}} \mid \int_{\Gamma} \rho_{\gamma} = Z \right\}, \quad (4.7)$$

the periodic rHF ground state energy (per unit cell) is given by

$$\boxed{I_{\text{per}}^0 = \inf \{ \mathcal{E}_{\text{per}}^0(\gamma), \gamma \in \mathcal{P}_{\text{per}}^Z \}}. \quad (4.8)$$

It was proved by Catto, Le Bris and Lions in [12] that there exists a minimizer $\gamma_{\text{per}}^0 \in \mathcal{P}_{\text{per}}^Z$ to the minimization problem (4.8), and that all the minimizers of (4.8) share the same density $\rho_{\gamma_{\text{per}}^0}$. We have the following result

Theorem 1 (Definition of the periodic rHF minimizer). *Let $Z \in \mathbb{N} \setminus \{0\}$. The minimization problem (4.8) admits a unique minimizer γ_{per}^0 . Denoting by*

$$H_{\text{per}}^0 := -\frac{\Delta}{2} + (\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1, \quad (4.9)$$

the corresponding periodic mean-field hamiltonian, γ_{per}^0 is solution to the self-consistent equation

$$\gamma_{\text{per}}^0 = \chi_{(-\infty, \epsilon_F]}(H_{\text{per}}^0), \quad (4.10)$$

where ϵ_F is a Lagrange multiplier called Fermi level, which can be interpreted as a chemical potential.

Additionally, for any $\epsilon_F \in \mathbb{R}$ such that (4.10) holds, γ_{per}^0 is the unique minimizer on \mathcal{P}_{per} of the energy functional

$$\gamma \mapsto \mathcal{E}_{\text{per}}^0(\gamma) - \epsilon_F \int_{\Gamma} \rho_{\gamma}.$$

Proof. Our proof uses classical ideas for Hartree-Fock theories. See [38, Section 4] for a very similar setting. Let us consider a minimizer γ_{per}^0 of I_{per}^0 (it is known to exist by [12, Thm 2.1]). First we note that the periodic potential $V_{\text{per}} := (\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1$ is in $L_{\text{loc}}^2(\mathbb{R}^3)$. Indeed, V_{per} is solution to the Poisson equation

$$-\Delta u = f, \quad (4.11)$$

with $f = 4\pi(\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}})$. As $\rho_{\gamma_{\text{per}}^0} \in L_{\text{loc}}^1(\mathbb{R}^3)$ and $\mu_{\text{per}} \in L_{\text{loc}}^1(\mathbb{R}^3)$, f is in $L_{\text{per}}^1(\mathbb{R}^3)$. Any solution to (4.11) therefore is in $W_{\text{loc}}^{2,1}(\mathbb{R}^3)$ [20]. As $W^{2,1} \hookrightarrow L^q$ for all $1 \leq q \leq 3$ [20], one has in particular $V_{\text{per}} \in L_{\text{loc}}^2(\mathbb{R}^3)$.

Thus V_{per} defines a Δ -bounded operator on $L^2(\mathbb{R}^3)$ with relative bound zero, (see Section 6.1) and therefore $H_{\text{per}}^0 = -\Delta/2 + V_{\text{per}}$ is self-adjoint on $L^2(\mathbb{R}^3)$ with domain $D(-\Delta) = H^2(\mathbb{R}^3)$ and form domain $H^1(\mathbb{R}^3)$. Besides, the spectrum of H_{per}^0 is purely absolutely continuous, composed of bands as stated in [47, Thm 1-2] and [42, Thm XIII.100]. This is going to be an essential property for the proof of the uniqueness of γ_{per}^0 . Let $(\lambda_k(\xi))_{k \geq 1}$ denote the non-decreasing sequence of the eigenvalues of $(H_{\text{per}}^0)_{\xi}$. Then

$$\sigma(H_{\text{per}}^0) = \bigcup_{k \geq 1} \lambda_k(\Gamma^*), \quad H_{\text{per}}^0 = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (H_{\text{per}}^0)_{\xi} d\xi.$$

The Bloch eigenvalues $\lambda_k(\xi)$, $k \geq 1$, $\xi \in \Gamma^*$ are known to be real analytic in each fixed direction and cannot be constant with respect to the variable ξ . We are going to prove that a consequence of this property is that the non-decreasing function

$$C : \kappa \mapsto \sum_{k \geq 1} |\{\xi \in \Gamma^* \mid \lambda_k(\xi) \leq \kappa\}|$$

is continuous on \mathbb{R} . To this aim, we begin by noting that the sum contained in the function is a finite one. The operator $H_{\text{per}}^0 = -\Delta/2 + V_{\text{per}}$ is bounded below by $-\Delta/2 + K$ for some real constant K . Consequently, the Bloch eigenvalues of H_{per}^0 are bounded, up to a multiplicative and an additive constants, by

those of the Laplacian, considered as an operator on $L^2_{\text{per}}(\Gamma)$, and it is known that $\sum_{k \geq 1} |\{\xi \in \Gamma^* \mid \tilde{\lambda}_k(\xi) \leq \kappa\}|$, $\tilde{\lambda}_k$ being here the Bloch eigenvalues of the Laplacian, is a finite sum. Now we take a decreasing sequence $(\kappa_n)_{n \in \mathbb{N}}$ which converges to κ . Using that the sum in $C(\kappa)$ is finite, it is sufficient to prove the continuity only for one given k , $k \geq 1$. We can write that

$$C_k(\kappa_n) = |\{\xi \in \Gamma^* \mid \lambda_k(\xi) \leq \kappa\}| + |A_n|$$

with $A_n = \{\xi \in \Gamma^* \mid \kappa < \lambda_k(\xi) \leq \kappa_n\}$. $(A_n)_{n \in \mathbb{N}}$ is a non-increasing sequence of ensembles and $\bigcap_{n \in \mathbb{N}} A_n = \emptyset$. Therefore $(|A_n|)_{n \in \mathbb{N}}$ goes to zero, which proves the right-continuity of C_k . To prove the left-continuity we consider an increasing sequence $(\kappa_n)_{n \in \mathbb{N}}$ which converges to κ . We have

$$C_k(\kappa) = C_k(\kappa_n) + |B_n|,$$

with $B_n = \{\xi \in \Gamma^* \mid \kappa_n < \lambda_k(\xi) \leq \kappa\}$. $(B_n)_{n \in \mathbb{N}}$ is a non-increasing sequence of ensembles such that $\bigcap_{n \in \mathbb{N}} B_n = \{\xi \in \Gamma^* \mid \lambda_k(\xi) = \kappa\}$, this set having a Lebesgue measure equal to zero by [47, Lemma 2]. The left-continuity of C_k follows. As for the non-decreasingness of C , it is obvious.

The operator H_{per}^0 being bounded from below, we have $C \equiv 0$ on $(-\infty, \inf \lambda_k(\Gamma^*))$ and it is known [47, Lemma A-2] that $\lim_{\kappa \rightarrow \infty} C(\kappa) = \infty$. We can thus choose a chemical potential ϵ_F such that

$$Z = C(\epsilon_F) = \sum_{k \geq 1} |\{\xi \in \Gamma^* \mid \lambda_k(\xi) \leq \epsilon_F\}|. \quad (4.12)$$

Considering a variation $(1-t)\gamma_{\text{per}}^0 + t\gamma$ for any $\gamma \in \mathcal{P}_{\text{per}}^Z$ and $t \in [0, 1]$, we have

$$\begin{aligned} \mathcal{E}_{\text{per}}^0((1-t)\gamma_{\text{per}}^0 + t\gamma) &= \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left(-\frac{1}{2} \Delta \left((\gamma_{\text{per}}^0)_\xi + t(\gamma_\xi - (\gamma_{\text{per}}^0)_\xi) \right) \right) \\ &\quad + \frac{1}{2} D_{G_1} \left(\rho_{\gamma_{\text{per}}^0} + t(\rho_\gamma - \rho_{\gamma_{\text{per}}^0}) - \mu_{\text{per}}, \rho_{\gamma_{\text{per}}^0} + t(\rho_\gamma - \rho_{\gamma_{\text{per}}^0}) - \mu_{\text{per}} \right) \\ &= \mathcal{E}_{\text{per}}^0(\gamma_{\text{per}}^0) + t \left(\frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left(-\frac{1}{2} \Delta (\gamma_\xi - (\gamma_{\text{per}}^0)_\xi) \right) \right) \\ &\quad + t D_{G_1} \left(\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}, \rho_\gamma - \rho_{\gamma_{\text{per}}^0} \right) + \frac{t^2}{2} D_{G_1} \left(\rho_\gamma - \rho_{\gamma_{\text{per}}^0}, \rho_\gamma - \rho_{\gamma_{\text{per}}^0} \right). \end{aligned}$$

For any t this expression is greater than or equal to the minimum $\mathcal{E}_{\text{per}}^0(\gamma_{\text{per}}^0)$. As $t \in [0, 1]$, the term of first order in t has to be nonnegative. This term reads $\frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left((H_{\text{per}}^0)_\xi (\gamma_\xi - (\gamma_{\text{per}}^0)_\xi) \right) d\xi$ where H_{per}^0 is the mean-field operator defined in (4.9). Therefore γ_{per}^0 minimizes the following linear functional

$$\gamma \in \mathcal{P}_{\text{per}}^Z \mapsto \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left((H_{\text{per}}^0)_\xi \gamma_\xi \right) d\xi.$$

We subtract the chemical potential ϵ_F defined above and introduce the free energy functional

$$\gamma \in \mathcal{P}_{\text{per}} \mapsto F(\gamma) := \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} \left((H_{\text{per}}^0 - \epsilon_F)_\xi \gamma_\xi \right) d\xi.$$

4.2. THE REDUCED HARTREE-FOCK MODEL FOR A PERFECT CRYSTAL 69

Notice that since $\frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}(\gamma_\xi) d\xi = Z$ for any $\gamma \in \mathcal{P}_{\text{per}}^Z$, then γ_{per}^0 also minimizes F on $\mathcal{P}_{\text{per}}^Z$. For any $\xi \in \Gamma^*$, we can find orthonormal functions $e_k(\xi, \cdot) \in L^2_\xi(\Gamma)$ such that

$$(H_{\text{per}}^0)_\xi = \sum_{k \geq 1} \lambda_k(\xi) |e_k(\xi, \cdot)\rangle \langle e_k(\xi, \cdot)|, \quad (4.13)$$

each function $(\xi, x) \mapsto e_k(\xi, x)$ being measurable on $\Gamma^* \times \Gamma$. Let us now define $\gamma^0 \in \mathcal{P}_{\text{per}}$ by

$$(\gamma^0)_\xi(x, y) = \sum_{k \geq 1} \delta_k(\xi) e_k(\xi, x) \overline{e_k(\xi, y)}, \quad \delta_k(\xi) = \begin{cases} 1 & \text{if } \lambda_k(\xi) \leq \epsilon_F \\ 0 & \text{if } \lambda_k(\xi) > \epsilon_F. \end{cases}$$

Saying this differently $\gamma^0 = \chi_{(-\infty, \epsilon_F]}(H_{\text{per}}^0)$. Notice ϵ_F was chosen to ensure $\gamma^0 \in \mathcal{P}_{\text{per}}^Z$.

We now prove that γ^0 is the unique minimizer of the function F defined above, on the set \mathcal{P}_{per} without a charge constraint. Since $\gamma^0 \in \mathcal{P}_{\text{per}}^Z$, this will prove that $\gamma_{\text{per}}^0 = \gamma^0$ and that γ_{per}^0 is the unique minimizer of F on \mathcal{P}_{per} . We write

$$\begin{aligned} F(\gamma) - F(\gamma^0) &= \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2_\xi(\Gamma)}((H_{\text{per}}^0 - \epsilon_F)_\xi(\gamma - \gamma^0)_\xi) d\xi \\ &= \sum_{k \geq 1} \frac{1}{(2\pi)^3} \int_{\Gamma^*} (\lambda_k(\xi) - \epsilon_F) (\langle \gamma_\xi e_k(\xi, \cdot), e_k(\xi, \cdot) \rangle_\xi - \delta_k(\xi)) d\xi \end{aligned}$$

where $\langle \cdot, \cdot \rangle_\xi$ is the usual inner product of $L^2_\xi(\Gamma)$. Since $0 \leq \gamma \leq 1$ in $L^2(\mathbb{R}^3)$, we have that $0 \leq \gamma_\xi \leq 1$ on $L^2_\xi(\Gamma)$ (see Section 2.2). As a consequence, $\langle \gamma_\xi e_k(\xi, \cdot), e_k(\xi, \cdot) \rangle \in [0, 1]$, for almost every $\xi \in \Gamma^*$. Hence, using the definition of $\delta_k(\xi)$,

$$F(\gamma) - F(\gamma^0) = \sum_{k \geq 1} \frac{1}{(2\pi)^3} \int_{\Gamma^*} |\lambda_k(\xi) - \epsilon_F| \times |\langle \gamma_\xi e_k(\xi, \cdot), e_k(\xi, \cdot) \rangle - \delta_k(\xi)| d\xi \geq 0.$$

This shows that γ^0 minimizes F on $\mathcal{P}_{\text{per}}^Z$. If now $F(\gamma) = F(\gamma^0)$, then necessarily $\langle \gamma_\xi e_k(\xi, \cdot), e_k(\xi, \cdot) \rangle = \delta_k(\xi)$ for almost every $\xi \in \Gamma^*$ and any $k \geq 1$, the set $\{\xi \in \Gamma^* \mid \exists k, \lambda_k(\xi) = \epsilon_F\}$ having a Lebesgue measure equal to zero as we saw before. Using now that the operators γ_ξ and $(1 - \gamma)_\xi$ are nonnegative, we infer that $\gamma_\xi e_k(\xi, \cdot) = \delta_k(\xi) e_k(\xi, \cdot)$ for all $k \geq 1$ and almost all $\xi \in \Gamma^*$. Hence $\gamma = \gamma^0$ and γ^0 is the unique minimizer of F . In particular $\gamma_{\text{per}}^0 = \gamma^0$, i.e. γ_{per}^0 solves the self-consistent equation (4.10).

Consider now another minimizer γ of the energy $\mathcal{E}_{\text{per}}^0$ on $\mathcal{P}_{\text{per}}^Z$, we recall that $\rho_\gamma = \rho_{\gamma_{\text{per}}^0}$ as was shown in [12]. Hence the operators H_{per}^0 and γ^0 defined above do not depend on the chosen minimizer. The above argument applied to γ shows that $\gamma = \gamma^0 = \gamma_{\text{per}}^0$, i.e. γ_{per}^0 is unique. \square

Let us notice that Theorem 1 contains three main results that were not mentioned in [12] : first γ_{per}^0 is unique, second it is a projector, and third it satisfies Equation (4.10). These three properties are crucial for a proper construction of the model for the crystal with a defect.

The projector γ_{per}^0 represents the state of the *Fermi sea*, i.e. of the infinite system of all the electrons in the periodic crystal. Of course, it is an infinite rank projector, meaning that

$$\gamma_{\text{per}}^0 = \sum_k |\varphi_k\rangle\langle\varphi_k|$$

should be interpreted as the one-body matrix of a formal infinite Slater determinant

$$\Psi = \varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_k \wedge \cdots .$$

The fact that γ_{per}^0 is additionally a spectral projector associated with the continuous spectrum of an operator leads to the obvious analogy with the *Dirac sea* which is the projector on the negative spectral subspace of the Dirac operator [22–25].

Most of our results will hold true for insulators only. When necessary, we shall take $Z \in \mathbb{N} \setminus \{0\}$ and make the following assumption :

(A1) *There is a gap between the Z -th and the $(Z + 1)$ -st band, i.e. $\Sigma_Z^+ < \Sigma_{Z+1}^-$, where Σ_Z^+ and Σ_{Z+1}^- are respectively the maximum and the minimum of the Z -th and the $(Z + 1)$ -st bands of H_{per}^0 .*

We emphasize that Assumption **(A1)** is a condition on the solution γ_{per}^0 of the *nonlinear* problem (4.8). Note that under **(A1)**, one has $\gamma_{\text{per}}^0 = \chi_{(-\infty, \epsilon_F]}(H_{\text{per}}^0)$ for any $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$.

4.3 The reduced Hartree-Fock model for a crystal with a defect

In this section, we define the reduced Hartree-Fock model describing the behavior of the Fermi sea and possibly of a finite number of bound electrons (or holes) close to a local defect. Our model is an obvious transposition of the Bogoliubov-Dirac-Fock model which was proposed by Chaix and Iracane [13] to describe the polarized Dirac sea (and a finite number of relativistic electrons) in the presence of an external potential. Our mathematical definition of the reduced energy functional follows mainly ideas from [22, 23]. We shall prove in Section 4.4 that this model can be obtained as the thermodynamic limit of the so-called supercell model. An analogous result was proved in [25] for the Bogoliubov-Dirac-Fock (BDF) model.

Assume that the periodic nuclear density μ_{per} defined in (2.15) is replaced by a locally perturbed nuclear density $\mu_{\text{per}} + \nu$. The defect ν can model a vacancy, an interstitial atom, or an impurity, with possible local rearrangement of the neighboring atoms. The main idea underlying the model is to define a *finite* energy by subtracting the *infinite* energy of the periodic Fermi sea γ_{per}^0 defined in the previous section, from the *infinite* energy of the perturbed system under consideration. For the BDF

model, this was proposed first in [25]. Formally, one obtains for a test state γ

$$\begin{aligned} \mathcal{E}_{\mu_{\text{per}}+\nu}^{\text{rHF}}(\gamma) - \mathcal{E}_{\mu_{\text{per}}+\nu}^{\text{rHF}}(\gamma_{\text{per}}^0) &= \text{Tr} \left(H_{\text{per}}^0(\gamma - \gamma_{\text{per}}^0) \right) \\ &- \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\nu(x)\rho_{[\gamma-\gamma_{\text{per}}^0]}(y)}{|x-y|} dx dy + \frac{1}{2} \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{\rho_{[\gamma-\gamma_{\text{per}}^0]}(x)\rho_{[\gamma-\gamma_{\text{per}}^0]}(y)}{|x-y|} dx dy. \end{aligned} \quad (4.14)$$

Of course the two terms in the left-hand side of (4.14) are not well-defined because μ_{per} is periodic and because γ and γ_{per}^0 have infinite ranks, but we shall be able to give a mathematical meaning to the right-hand side, exploiting the fact that $Q := \gamma - \gamma_{\text{per}}^0$ induces a small perturbation of the reference state γ_{per}^0 . The formal computation (4.14) will be justified by means of thermodynamic limit arguments in Section 4.4.

4.3.1 Definition of the reduced Hartree-Fock energy of a defect

We now define properly the reduced Hartree-Fock energy of the Fermi sea in the presence of the defect ν . We denote by \mathfrak{S}_p the Schatten class of operators Q acting on $L^2(\mathbb{R}^3)$ having a finite p trace, i.e. such that $\text{Tr}(|Q|^p) < \infty$. Note that \mathfrak{S}_1 is the space of trace-class operators, and that \mathfrak{S}_2 is the space of Hilbert-Schmidt operators. Let Π be an orthogonal projector on $L^2(\mathbb{R}^3)$ such that both Π and $1 - \Pi$ have infinite ranks. A self-adjoint compact operator Q is said to be Π -trace class ($Q \in \mathfrak{S}_1^\Pi$) when $Q \in \mathfrak{S}_2$ and $\Pi Q \Pi, (1 - \Pi)Q(1 - \Pi) \in \mathfrak{S}_1$. Its Π -trace is then defined as $\text{Tr}_\Pi(Q) = \text{Tr}(\Pi Q \Pi + (1 - \Pi)Q(1 - \Pi))$. Notice that if $Q \in \mathfrak{S}_1$, then $Q \in \mathfrak{S}_1^\Pi$ for any Π and $\text{Tr}_\Pi(Q) = \text{Tr}(Q)$. See [22, Section 2.1] for general properties related to this definition. In the following, we use the shorthand notation

$$Q^{--} := \gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0, \quad Q^{++} := (1 - \gamma_{\text{per}}^0)Q(1 - \gamma_{\text{per}}^0),$$

$$\mathfrak{S}_1^0 := \mathfrak{S}_1^{\gamma_{\text{per}}^0} = \{Q \in \mathfrak{S}_2 \mid Q^{++} \in \mathfrak{S}_1, Q^{--} \in \mathfrak{S}_1\} \quad \text{and} \quad \text{Tr}_0(Q) := \text{Tr}_{\gamma_{\text{per}}^0}(Q).$$

We also introduce the Banach space

$$\mathcal{Q} = \{Q \in \mathfrak{S}_1^0 \mid Q^* = Q, |\nabla|Q \in \mathfrak{S}_2, |\nabla|Q^{++}|\nabla| \in \mathfrak{S}_1, |\nabla|Q^{--}|\nabla| \in \mathfrak{S}_1\},$$

endowed with its natural norm

$$\begin{aligned} \|Q\|_{\mathcal{Q}} &:= \|Q\|_{\mathfrak{S}_2} + \|Q^{++}\|_{\mathfrak{S}_1} + \|Q^{--}\|_{\mathfrak{S}_1} \\ &+ \||\nabla|Q\|_{\mathfrak{S}_2} + \||\nabla|Q^{++}|\nabla|\|_{\mathfrak{S}_1} + \||\nabla|Q^{--}|\nabla|\|_{\mathfrak{S}_1}. \end{aligned} \quad (4.15)$$

The convex set on which the energy will be defined is

$$\mathcal{K} := \{Q \in \mathcal{Q} \mid -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0\}. \quad (4.16)$$

Notice that \mathcal{K} is the closed convex hull of states $Q \in \mathcal{Q}$ of the special form $Q = \gamma - \gamma_{\text{per}}^0$, γ being an orthogonal projector on $L^2(\mathbb{R}^3)$. Besides, the number $\text{Tr}_0(Q)$

can be interpreted as the charge of the system measured with respect to that of the unperturbed Fermi sea. As we saw in Section 2.3, it can be proved (see also [22, Lemma 2]) that $\text{Tr}_0(Q)$ is always an integer if Q is a Hilbert-Schmidt operator of the special form $Q = \gamma - \gamma_{\text{per}}^0$, with γ an orthogonal projector. Additionally, in this case, $\text{Tr}_0(Q) = 0$ when $\|Q\| < 1$.

We are now going to show (see also [1, 22]) that the constraint $-\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0$ in (4.16) is equivalent to the inequality

$$Q^2 \leq Q^{++} - Q^{--} \quad (4.17)$$

and implies in particular that $Q^{++} \geq 0$ and $Q^{--} \leq 0$ for any $Q \in \mathcal{K}$.

$$\begin{aligned} -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0 &\iff 0 \leq Q + \gamma_{\text{per}}^0 \leq 1, \\ &\iff (Q + \gamma_{\text{per}}^0)^2 \leq Q + \gamma_{\text{per}}^0, \\ &\iff Q^2 + Q\gamma_{\text{per}}^0 + \gamma_{\text{per}}^0 Q + \gamma_{\text{per}}^0 \leq Q + \gamma_{\text{per}}^0, \\ &\iff Q^2 \leq Q - Q\gamma_{\text{per}}^0 - \gamma_{\text{per}}^0 Q, \end{aligned}$$

and we have $Q^{++} - Q^{--} = (1 - \gamma_{\text{per}}^0)Q(1 - \gamma_{\text{per}}^0) - \gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0 = Q - Q\gamma_{\text{per}}^0 - \gamma_{\text{per}}^0 Q$. In addition,

$$Q \leq 1 - \gamma_{\text{per}}^0 \implies Q^{--} = \gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0 \leq \gamma_{\text{per}}^0 (1 - \gamma_{\text{per}}^0) \gamma_{\text{per}}^0 = 0$$

and

$$-\gamma_{\text{per}}^0 \leq Q \implies 0 = -(1 - \gamma_{\text{per}}^0)\gamma_{\text{per}}^0(1 - \gamma_{\text{per}}^0) \leq (1 - \gamma_{\text{per}}^0)Q(1 - \gamma_{\text{per}}^0) = Q^{++}.$$

In order to define properly the energy of Q , we need to associate a density ρ_Q with any state $Q \in \mathcal{K}$. We shall see that ρ_Q can in fact be defined for any $Q \in \mathcal{Q}$. This is not obvious *a priori* since \mathcal{Q} does not only contain trace-class operators. Additionally we need to check that the last two terms of (4.14) are well-defined. For this purpose, we introduce the so-called Coulomb space

$$\mathcal{C} := \{\rho \in \mathcal{S}'(\mathbb{R}^3) \mid \hat{\rho} \in L^1_{\text{loc}}(\mathbb{R}^3), D(\rho, \rho) < \infty\}$$

where $D(f, g) = 4\pi \int_{\mathbb{R}^3} |k|^{-2} \widehat{f}(k) \widehat{g}(k) dk$ was already defined by (4.2). Endowed with this inner product, \mathcal{C} is a Hilbert space. Indeed, let us take some Cauchy sequence (ρ_n) in \mathcal{C} . Then $\frac{\hat{\rho}_n(k)}{|k|}$ is Cauchy in $L^2(\mathbb{R}^3)$, therefore converges to some $g \in L^2(\mathbb{R}^3)$ in $L^2(\mathbb{R}^3)$. Let $\rho = \mathcal{F}^{-1}(|k|g)$. Clearly, $\rho \in \mathcal{S}'(\mathbb{R}^3)$ and since $g \in L^2(\mathbb{R}^3)$, $\rho \in \mathcal{C}$. Moreover,

$$\begin{aligned} \|\rho_n - \rho\|_{\mathcal{C}}^2 &= 4\pi \int_{\mathbb{R}^3} \left| \frac{\hat{\rho}_n(k) - \hat{\rho}(k)}{|k|} \right|^2 dk, \\ &= 4\pi \int_{\mathbb{R}^3} \left| \frac{\hat{\rho}_n(k)}{|k|} - g(k) \right|^2 dk, \\ &\rightarrow 0 \quad \text{when } n \rightarrow +\infty. \end{aligned}$$

This proves the completeness of \mathcal{C} . We then introduce the Beppo-Levi space $\tilde{\mathcal{C}} := \{V \in L^6(\mathbb{R}^3) \mid \nabla V \in L^2(\mathbb{R}^3)\}$. Endowed with the inner product

$$(V, W)_{\tilde{\mathcal{C}}} = \int_{\mathbb{R}^3} \nabla V \cdot \nabla W,$$

$\tilde{\mathcal{C}}$ is a Hilbert space. Note that the condition over the gradient can be rewritten $\int_{\mathbb{R}^3} |k|^2 |\hat{V}(k)|^2 dk < \infty$. For $V \in \tilde{\mathcal{C}}$, we now introduce the linear form

$$\begin{aligned} \Phi_V : \rho \in \mathcal{C} &\rightarrow \mathbb{C}, \\ \rho &\mapsto \Phi_V(\rho) = \int_{\mathbb{R}^3} \hat{\rho}(k) \overline{\hat{V}(k)} dk, \end{aligned}$$

which is well defined since

$$\int_{\mathbb{R}^3} |\hat{\rho}(k) \overline{\hat{V}(k)}| dk \leq \left(\int_{\mathbb{R}^3} \frac{|\hat{\rho}(k)|^2}{|k|^2} dk \right)^{\frac{1}{2}} \left(\int_{\mathbb{R}^3} |k|^2 |\hat{V}(k)|^2 dk \right)^{\frac{1}{2}} = \|\rho\|_{\mathcal{C}} \|V\|_{\tilde{\mathcal{C}}} < +\infty.$$

The above inequality also proves that $\|\Phi_V\|_{\mathcal{C}'} = \sup_{\rho \in \mathcal{C} \setminus \{0\}} \frac{|\Phi_V(\rho)|}{\|\rho\|_{\mathcal{C}}} \leq \|V\|_{\tilde{\mathcal{C}}}$. Now, we define for $V \in \tilde{\mathcal{C}}$, $\rho_V = \frac{1}{4\pi} \mathcal{F}^{-1}(|k|^2 \hat{V}(k))$. It is easily seen that $\rho_V \in \mathcal{C}$ and that $|\Phi_V(\rho_V)| = \|V\|_{\tilde{\mathcal{C}}} \|\rho_V\|_{\mathcal{C}}$. Therefore $\|\Phi_V\|_{\mathcal{C}'} = \|V\|_{\tilde{\mathcal{C}}}$. To prove that Φ is a bijection between \mathcal{C}' and $\tilde{\mathcal{C}}$, we now prove that for each $\rho \in \mathcal{C}$, there exists $V \in \tilde{\mathcal{C}}$ such that $\Phi_V = D(\rho, \cdot)$. Taking $V_\rho = \mathcal{F}^{-1}(|k|^{-2} \hat{\rho})$, we indeed have

$$\begin{aligned} \forall \rho' \in \mathcal{C}, \Phi_{V_\rho}(\rho') &= \int_{\mathbb{R}^3} \hat{\rho}'(k) \overline{\hat{V}_\rho(k)} dk \\ &= 4\pi \int_{\mathbb{R}^3} \frac{\hat{\rho}'(k) \hat{\rho}(k)}{|k|^2} dk \\ &= D(\rho_V, \rho'). \end{aligned}$$

The isometry Φ allows us to identify the Beppo-Levi space with \mathcal{C}' .

We now use a duality argument to define ρ_Q :

Proposition 1 (Definition of the density ρ_Q for $Q \in \mathcal{Q}$). *Assume that $Q \in \mathcal{Q}$. Then $QV \in \mathfrak{S}_1^0$ for any $V = V_1 + V_2 \in \mathcal{C}' + (L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3))$ and moreover there exists a constant C (independent of Q and V) such that*

$$|\mathrm{Tr}_0(QV)| \leq C \|Q\|_{\mathcal{Q}} (\|V_1\|_{\mathcal{C}'} + \|V_2\|_{L^2(\mathbb{R}^3)}). \quad (4.18)$$

Thus the linear form $V \in \mathcal{C}' + (L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)) \mapsto \mathrm{Tr}_0(QV)$ can be continuously extended to $\mathcal{C}' + L^2(\mathbb{R}^3)$ and there exists a uniquely defined function $\rho_Q \in \mathcal{C} \cap L^2(\mathbb{R}^3)$ such that

$$\forall V = V_1 + V_2 \in \mathcal{C}' + (L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)), \quad \langle \rho_Q, V_1 \rangle_{\mathcal{C}, \mathcal{C}'} + \int_{\mathbb{R}^3} \rho_Q V_2 = \mathrm{Tr}_0(QV).$$

The linear map $Q \in \mathcal{Q} \mapsto \rho_Q \in \mathcal{C} \cap L^2(\mathbb{R}^3)$ is continuous :

$$\|\rho_Q\|_{\mathcal{C}} + \|\rho_Q\|_{L^2(\mathbb{R}^3)} \leq C \|Q\|_{\mathcal{Q}}.$$

Eventually when $Q \in \mathfrak{S}_1 \subset \mathfrak{S}_1^0$, then $\rho_Q(x) = Q(x, x)$ where $Q(x, y)$ is the integral kernel of Q .

Proof. Let $V = V_1 + V_2$ where $V_1 \in \mathcal{C}'$ and $V_2 \in L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$, and $Q \in \mathcal{Q}$. Denoting by $(\gamma_{\text{per}}^0)^\perp = 1 - \gamma_{\text{per}}^0$, we can notice that

$$(QV)^{++} = (\gamma_{\text{per}}^0)^\perp(QV)(\gamma_{\text{per}}^0)^\perp$$

and that

$$\begin{aligned} & Q^{++}V(\gamma_{\text{per}}^0)^\perp + Q^{+-}[\gamma_{\text{per}}^0, V](\gamma_{\text{per}}^0)^\perp \\ &= (\gamma_{\text{per}}^0)^\perp Q(\gamma_{\text{per}}^0)^\perp V(\gamma_{\text{per}}^0)^\perp + (\gamma_{\text{per}}^0)^\perp Q\gamma_{\text{per}}^0 (\gamma_{\text{per}}^0 V - V\gamma_{\text{per}}^0)(\gamma_{\text{per}}^0)^\perp, \\ &= (\gamma_{\text{per}}^0)^\perp Q(\gamma_{\text{per}}^0)^\perp V(\gamma_{\text{per}}^0)^\perp + (\gamma_{\text{per}}^0)^\perp Q\gamma_{\text{per}}^0 V(\gamma_{\text{per}}^0)^\perp, \\ &= (\gamma_{\text{per}}^0)^\perp Q(\gamma_{\text{per}}^0 + (\gamma_{\text{per}}^0)^\perp)V(\gamma_{\text{per}}^0)^\perp, \\ &= (\gamma_{\text{per}}^0)^\perp(QV)(\gamma_{\text{per}}^0)^\perp. \end{aligned}$$

Similarly,

$$\begin{aligned} & Q^{--}V\gamma_{\text{per}}^0 - Q^{-+}[\gamma_{\text{per}}^0, V]\gamma_{\text{per}}^0 \\ &= \gamma_{\text{per}}^0 Q\gamma_{\text{per}}^0 V\gamma_{\text{per}}^0 - \gamma_{\text{per}}^0 Q(\gamma_{\text{per}}^0)^\perp (\gamma_{\text{per}}^0 V - V\gamma_{\text{per}}^0)\gamma_{\text{per}}^0, \\ &= \gamma_{\text{per}}^0 Q\gamma_{\text{per}}^0 V\gamma_{\text{per}}^0 + \gamma_{\text{per}}^0 Q(\gamma_{\text{per}}^0)^\perp V\gamma_{\text{per}}^0, \\ &= \gamma_{\text{per}}^0(QV)\gamma_{\text{per}}^0, \end{aligned}$$

so that

$$(QV)^{--} = Q^{--}V\gamma_{\text{per}}^0 - Q^{-+}[\gamma_{\text{per}}^0, V]\gamma_{\text{per}}^0. \quad (4.19)$$

We only deal with the $(QV)^{--}$ term, the argument being the same for $(QV)^{++}$.

First we write $Q^{--}V\gamma_{\text{per}}^0 = (1 + |\nabla|)^{-1}(1 + |\nabla|)Q^{--}(1 + |\nabla|)(1 + |\nabla|)^{-1}V\gamma_{\text{per}}^0$ and notice that $(1 + |\nabla|)^{-1}V$ is bounded since $V_2 \in L^\infty(\mathbb{R}^3)$ by assumption and

$$\|(1 + |\nabla|)^{-1}V_1\|_{\mathfrak{S}_6} \leq C \|V_1\|_{L^6} \leq C' \|\nabla V_1\|_{L^2} = C' \|V_1\|_{\mathcal{C}'}$$

by the Kato-Simon-Seiler inequality (6.3) and the critical Sobolev embedding of $H^1(\mathbb{R}^3)$ in $L^6(\mathbb{R}^3)$. As $(1 + |\nabla|)^{-1}$ and γ_{per}^0 are also bounded, and $(1 + |\nabla|)Q^{--}(1 + |\nabla|)$ is trace-class, this proves that $Q^{--}V\gamma_{\text{per}}^0$ is a trace-class operator. Thus the following is true :

$$\begin{aligned} |\text{Tr}(Q^{--}V\gamma_{\text{per}}^0)| &= |\text{Tr}(Q^{--}V)| \\ &= |\text{Tr}((1 + |\nabla|)Q^{--}(1 + |\nabla|)(1 + |\nabla|)^{-1}V(1 + |\nabla|)^{-1})| \\ &\leq \|Q\|_{\mathcal{Q}} \|(1 + |\nabla|)^{-1}V(1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty}. \end{aligned}$$

Then

$$\begin{aligned} \|(1 + |\nabla|)^{-1}V_1(1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty} &\leq \|(1 + |\nabla|)^{-1}V_1\|_{\mathfrak{S}_6} \|(1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty} \\ &\leq C \|V_1\|_{L^6} \leq C' \|V_1\|_{\mathcal{C}'}, \end{aligned}$$

and

$$\begin{aligned} \|(1 + |\nabla|)^{-1}V_2(1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty} &\leq \|(1 + |\nabla|)^{-1}|V_2|(1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty} \\ &\leq \left\| \left((1 + |\nabla|)^{-1}|V_2|^{1/2} \right)^2 \right\|_{\mathfrak{S}_2} \\ &\leq \left\| (1 + |\nabla|)^{-1}|V_2|^{1/2} \right\|_{\mathfrak{S}_4}^2 \\ &\leq C \left\| (V_2)^{\frac{1}{2}} \right\|_{L^4}^2 = C \|V_2\|_{L^2}. \end{aligned}$$

In the last inequality, we have used the Kato-Simon-Seiler inequality. Hence,

$$|\mathrm{Tr}(Q^{--}V\gamma_{\mathrm{per}}^0)| \leq C \|Q\|_{\mathcal{Q}} (\|V_1\|_{\mathcal{C}'} + \|V_2\|_{L^2}).$$

For the second term of (4.19), we just use Lemma 12 which tells us that

$$Q^{-+}[\gamma_{\mathrm{per}}^0, V]\gamma_{\mathrm{per}}^0 \in \mathfrak{S}_1$$

since $Q^{-+} \in \mathfrak{S}_2$ and $[\gamma_{\mathrm{per}}^0, V] \in \mathfrak{S}_2$. Additionally

$$|\mathrm{Tr}(Q^{-+}[\gamma_{\mathrm{per}}^0, V]\gamma_{\mathrm{per}}^0)| \leq \|Q^{-+}[\gamma_{\mathrm{per}}^0, V]\gamma_{\mathrm{per}}^0\|_{\mathfrak{S}_1} \leq C \|Q^{-+}\|_{\mathfrak{S}_2} (\|V_1\|_{\mathcal{C}'} + \|V_2\|_{L^2}).$$

By the Hahn-Banach Theorem [6], the linear form $V \in \mathcal{C}' + (L^2 \cap L^\infty) \mapsto \mathrm{Tr}_0(QV)$ can be continuously extended to $\mathcal{C}' + L^2$. By density of $L^2 \cap L^\infty$ in L^2 , this form is unique.

$\mathcal{C} \cap L^2$ is reflexive so it can be identified to its bidual. Its dual is $\mathcal{C}' + L^2$. $\mathrm{Tr}_0(Q)$ belonging to $(\mathcal{C}' + L^2)'$, we can now identify it to a unique function $\rho_Q \in \mathcal{C} \cap L^2$. We now compute the norm of that function, which is by definition

$$\begin{aligned} \|\rho_Q\|_{\mathcal{C}} + \|\rho_Q\|_{L^2} &= \sup_{V \in \mathcal{C}' + L^2} \frac{|\mathrm{Tr}_0(QV)|}{\|V\|_{\mathcal{C}' + L^2}}, \\ &= \sup_{V_1 \in \mathcal{C}', V_2 \in L^2} \frac{|\mathrm{Tr}_0(Q(V_1 + V_2))|}{\|V_1\|_{\mathcal{C}'} + \|V_2\|_{L^2}}, \\ &\leq C \|Q\|_{\mathcal{Q}}, \end{aligned}$$

by (4.18).

If $Q \in \mathfrak{S}_1$, then $\forall V \in \mathcal{C}' + L^2$, $\mathrm{Tr}_0(QV) = \mathrm{Tr}(QV) = \int_{\mathbb{R}^3} Q(x, x)V(x) dx$, which is also equal by the definition of ρ_Q , to $\int_{\mathbb{R}^3} \rho_Q(x)V(x) dx$. This implies $Q(x, x) = \rho_Q(x)$. \square

Assuming that **(A1)** holds true, we are now in position to give a rigorous sense to the right-hand side of (4.14) for $\gamma - \gamma_{\mathrm{per}}^0 = Q \in \mathcal{K}$. In the sequel, we use the following *notation* for any $Q \in \mathcal{Q}$:

$$\mathrm{Tr}_0(H_{\mathrm{per}}^0 Q) := \mathrm{Tr}(|H_{\mathrm{per}}^0 - \kappa|^{1/2}(Q^{++} - Q^{--})|H_{\mathrm{per}}^0 - \kappa|^{1/2}) + \kappa \mathrm{Tr}_0(Q) \quad (4.20)$$

where κ is an arbitrary real number in the gap $(\Sigma_Z^+, \Sigma_{Z+1}^-)$ (this expression will be proved to be independent of κ , see Corollary 1 below). Then we define the energy of any state $Q \in \mathcal{K}$ as

$$\boxed{\mathcal{E}^\nu(Q) := \mathrm{Tr}_0(H_{\mathrm{per}}^0 Q) - D(\rho_Q, \nu) + \frac{1}{2}D(\rho_Q, \rho_Q)}. \quad (4.21)$$

The function ν is an external density of charge representing the nuclear charge of the defect. For the sake of simplicity, we shall assume here that $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) \subset \mathcal{C}$, although some of our results are true with a weaker assumption. We shall need the following

Lemma 1. *Assume that (A1) holds true. For any fixed κ in the gap $(\Sigma_Z^+, \Sigma_{Z+1}^-)$, there exist two constants $c_1, c_2 > 0$ such that*

$$c_1(1 - \Delta) \leq |H_{\text{per}}^0 - \kappa| \leq c_2(1 - \Delta) \quad (4.22)$$

as operators on $L^2(\mathbb{R}^3)$. In particular

$$\| |H_{\text{per}}^0 - \kappa|^{1/2} (1 - \Delta)^{-1/2} \| \leq \sqrt{c_2}, \quad \| |H_{\text{per}}^0 - \kappa|^{-1/2} (1 - \Delta)^{1/2} \| \leq 1/\sqrt{c_1}.$$

Similarly, $(H_{\text{per}}^0 - \kappa)(1 - \Delta)^{-1}$ and its inverse are bounded operators.

Proof. Recall that $H_{\text{per}}^0 = -\Delta/2 + V_{\text{per}}$ with $V_{\text{per}} = (\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1 \in L^\infty(\mathbb{R}^3)$. Indeed, V_{per} verifies the Poisson equation with $4\pi(\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}})$ as the right-hand side member of the equation. As both $\rho_{\gamma_{\text{per}}^0}$ and μ_{per} are in L^2_{per} , V_{per} is in H^2_{per} , and there is a Sobolev injection between that space and $C^{\frac{1}{2}}(\mathbb{R}^3)$ [6]. In particular, V_{per} is continuous and therefore bounded. Thus $|H_{\text{per}}^0 - \kappa| \geq H_{\text{per}}^0 - \kappa \geq -\Delta/2 - C$ for some large enough constant C . On the other hand, as $\kappa \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, there exists $\alpha > 0$ such that $|H_{\text{per}}^0 - \kappa| \geq \alpha$. This implies that

$$|H_{\text{per}}^0 - \kappa| \geq \max(-\Delta/2 - C, \alpha) \geq c_1(1 - \Delta)$$

for some constant $c_1 > 0$. The proof of the upper bound in (4.22) is straightforward.

Then $(-\Delta/2 + c)^{-1}(H_{\text{per}}^0 - \kappa + c) = 1 + (-\Delta/2 + c)^{-1}(V_{\text{per}} - \kappa)$ is a bounded (and therefore continuous) invertible operator for c large enough, since

$$\| (-\Delta/2 + c)^{-1}(V_{\text{per}} - \kappa) \| \leq \frac{\|V_{\text{per}}\|_{L^\infty} + |\kappa|}{c}.$$

Thus $(H_{\text{per}}^0 - \kappa + c)^{-1}(-\Delta/2 + c)$ is bounded for a well-chosen $c \gg 1$ [6], which clearly implies that

$$(H_{\text{per}}^0 - \kappa)^{-1}(-\Delta + 1) = \frac{H_{\text{per}}^0 - \kappa + c}{H_{\text{per}}^0 - \kappa} (H_{\text{per}}^0 - \kappa + c)^{-1}(-\Delta/2 + c) \frac{-\Delta + 1}{-\Delta/2 + c}$$

is also bounded, together with its inverse. \square

By the definition of \mathcal{Q} and Lemma 1, it is clear that the right-hand side of (4.20) is a well-defined quantity for any $Q \in \mathcal{Q}$ and any $\kappa \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. Together with Proposition 1 which tells us that $\rho_Q \in \mathcal{C}$ for any $Q \in \mathcal{Q}$, we deduce that (4.21) is a well-defined functional.

We shall need the following space of more regular operators

$$\mathcal{Q}_r := \{Q \in \mathcal{Q} \mid (-\Delta)Q^2(-\Delta) \in \mathfrak{S}_1, (-\Delta)(Q^{++} - Q^{--})(-\Delta) \in \mathfrak{S}_1\} \quad (4.23)$$

and the associated convex set

$$\mathcal{K}_r := \mathcal{K} \cap \mathcal{Q}_r.$$

The following result will be useful :

Lemma 2. *The space \mathcal{Q}_r (resp. the convex set \mathcal{K}_r) is dense in \mathcal{Q} (resp. in \mathcal{K}) for the topology of \mathcal{Q} .*

Proof. Let $Q \in \mathcal{Q}$. For $\epsilon > 0$, we introduce the following regularization operator

$$R_\epsilon := (1 + \epsilon |H_{\text{per}}^0 - \Sigma|)^{-1} \quad (4.24)$$

where we have chosen Σ as the center of the gap

$$\Sigma := \frac{\Sigma_Z^+ + \Sigma_{Z+1}^-}{2}.$$

and set

$$Q_\epsilon := R_\epsilon Q R_\epsilon.$$

Notice first that $Q_\epsilon \in \mathcal{Q}_r$: to show that, we introduce the notation $B(z) := (z - H_{\text{per}}^0)^{-1}(-\Delta + 1)$ and we obtain

$$(1 - \Delta)R_\epsilon = (1 - \Delta)(H_{\text{per}}^0 - \Sigma)^{-1} \frac{H_{\text{per}}^0 - \Sigma}{1 + \epsilon |H_{\text{per}}^0 - \Sigma|} = -B(\Sigma)^* \frac{H_{\text{per}}^0 - \Sigma}{1 + \epsilon |H_{\text{per}}^0 - \Sigma|}$$

which shows that $\|(1 - \Delta)R_\epsilon\| \leq \epsilon^{-1} \|B(\Sigma)^*\|$. Similarly, $\|R_\epsilon(1 - \Delta)\| \leq \epsilon^{-1} \|B(\Sigma)\|$. As R_ϵ commutes with γ_{per}^0 , we infer

$$(1 - \Delta)Q_\epsilon^{--}(1 - \Delta) = (1 - \Delta)R_\epsilon Q^{--} R_\epsilon (1 - \Delta) \in \mathfrak{S}_1.$$

Likewise, $(1 - \Delta)Q_\epsilon^{++}(1 - \Delta) \in \mathfrak{S}_1$. It remains to prove that $(-\Delta)Q_\epsilon^2(-\Delta) \in \mathfrak{S}_1$. We have

$$\begin{aligned} \|(-\Delta)Q_\epsilon^2(-\Delta)\|_{\mathfrak{S}_1} &= \|(-\Delta)R_\epsilon Q R_\epsilon^2 Q R_\epsilon(-\Delta)\|_{\mathfrak{S}_1}, \\ &\leq \|(-\Delta)R_\epsilon\|_{\mathfrak{S}_\infty} \|Q\|_{\mathfrak{S}_2} \|R_\epsilon^2\|_{\mathfrak{S}_\infty} \|Q\|_{\mathfrak{S}_2} \|(-\Delta)R_\epsilon\|_{\mathfrak{S}_\infty}. \end{aligned}$$

As $(-\Delta)R_\epsilon$ and R_ϵ^2 are bounded and $Q \in \mathfrak{S}_2$, the right-hand side is bounded and $(-\Delta)Q_\epsilon^2(-\Delta) \in \mathfrak{S}_1$. Then we show that $Q_\epsilon \in \mathcal{K}_r \subset \mathcal{K}$ when $Q \in \mathcal{K}$. To prove this, we use the fact that $-\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0$ is equivalent to $Q^2 \leq Q^{++} - Q^{--}$ (see Section 4.3.1). As $\|R_\epsilon\| \leq 1$, we obtain

$$(Q_\epsilon)^2 = R_\epsilon Q (R_\epsilon)^2 Q R_\epsilon \leq R_\epsilon Q^2 R_\epsilon \leq R_\epsilon (Q^{++} - Q^{--}) R_\epsilon = (Q_\epsilon)^{++} - (Q_\epsilon)^{--} \quad (4.25)$$

where we have used that $(R_\epsilon)^2 \leq 1$ and that γ_{per}^0 commutes with R_ϵ . Hence, it only remains to prove that $Q_\epsilon \rightarrow Q$ for the \mathcal{Q} -topology as $\epsilon \rightarrow 0$, for any fixed $Q \in \mathcal{Q}$. We shall need the

Lemma 3. *For any $1 \leq p < \infty$ and any fixed $Q \in \mathfrak{S}_p$, one has*

$$\lim_{\epsilon \rightarrow 0} \|R_\epsilon Q - Q\|_{\mathfrak{S}_p} = 0. \quad (4.26)$$

Proof. Notice that

$$R_\epsilon - 1 = -\frac{\epsilon |H_{\text{per}}^0 - \Sigma|}{1 + \epsilon |H_{\text{per}}^0 - \Sigma|}$$

satisfies $\|R_\epsilon - 1\| \leq 1$. Hence $\|(R_\epsilon - 1)Q\|_{\mathfrak{S}_p} \leq \|Q\|_{\mathfrak{S}_p}$. The space of “smooth” finite rank operators is dense in \mathfrak{S}_p for any $1 \leq p < \infty$. To show that, let us take $Q \in \mathfrak{S}_p$. One can always expand Q as

$$Q = \sum_{i=1}^{+\infty} \lambda_i |f_i\rangle \langle f_i|,$$

where $(f_i)_{i \in \mathbb{N}}$ forms an orthonormal basis of $L^2(\mathbb{R}^3)$ and where the sequence of the eigenvalues of Q (λ_i) is in ℓ^p . Then we introduce

$$Q_N = \sum_{i=1}^N \lambda_i |f_i\rangle \langle f_i|,$$

and

$$Q_{N\epsilon} = \sum_{i=1}^N \lambda_i |f_{i\epsilon}\rangle \langle f_{i\epsilon}|,$$

the $f_{i\epsilon}$ being in $H^2(\mathbb{R}^3)$, such that $\langle f_{i\epsilon}, f_{j\epsilon} \rangle = \delta_{ij}$ and converging when $\epsilon \rightarrow 0$ to f_i in $L^2(\mathbb{R}^3)$.

We are going to show that $Q_{N\epsilon} \rightarrow Q$ in \mathfrak{S}_p . First,

$$\|Q - Q_N\|_{\mathfrak{S}_p} = \sum_{i=N+1}^{+\infty} |\lambda_i|^p,$$

which converges to 0 when N goes to infinity. Then, using the fact that the norms $\|\cdot\|_{\mathfrak{S}_1}$ and $\|\cdot\|_{\mathfrak{S}_2}$ are equivalent on the set of the operators on $L^2(\mathbb{R}^3)$ of rank lower

than or equal to 2, we obtain that there exists a constant $c \in \mathbb{R}_+$ such that

$$\begin{aligned}
\|Q_N - Q_{N\epsilon}\|_{\mathfrak{S}_1} &= \left\| \sum_{i=1}^N \lambda_i (|f_i\rangle\langle f_i| - |f_{i\epsilon}\rangle\langle f_{i\epsilon}|) \right\|_{\mathfrak{S}_1}, \\
&\leq 2N \left(\sup_{1 \leq i \leq N} |\lambda_i| \right) \sup_{1 \leq i \leq N} \| |f_i\rangle\langle f_i| - |f_{i\epsilon}\rangle\langle f_{i\epsilon}| \|_{\mathfrak{S}_1}, \\
&\leq 2cN \left(\sup_{1 \leq i \leq N} |\lambda_i| \right) \sup_{1 \leq i \leq N} \| |f_i\rangle\langle f_i| - |f_{i\epsilon}\rangle\langle f_{i\epsilon}| \|_{\mathfrak{S}_2}, \\
&= 2cN \left(\sup_{1 \leq i \leq N} |\lambda_i| \right) \sup_{1 \leq i \leq N} \left(\iint_{\mathbb{R}^3 \times \mathbb{R}^3} (f_i(x)f_i(y) - f_{i\epsilon}(x)f_{i\epsilon}(y))^2 dx dy \right)^{\frac{1}{2}}, \\
&= 2cN \left(\sup_{1 \leq i \leq N} |\lambda_i| \right) \sup_{1 \leq i \leq N} \left(\iint_{\mathbb{R}^3 \times \mathbb{R}^3} (f_i(x)(f_i(y) - f_{i\epsilon}(y)) \right. \\
&\quad \left. + f_{i\epsilon}(y)(f_i(x) - f_{i\epsilon}(x)))^2 dx dy \right)^{\frac{1}{2}}, \\
&\leq 2cN \left(\sup_{1 \leq i \leq N} |\lambda_i| \right) \sup_{1 \leq i \leq N} \left(2 \iint_{\mathbb{R}^3 \times \mathbb{R}^3} f_i(x)^2 (f_i(y) - f_{i\epsilon}(y))^2 dx dy \right. \\
&\quad \left. + 2 \iint_{\mathbb{R}^3 \times \mathbb{R}^3} f_{i\epsilon}(y)^2 (f_i(x) - f_{i\epsilon}(x))^2 dx dy \right)^{\frac{1}{2}}, \\
&= 2cN \left(\sup_{1 \leq i \leq N} |\lambda_i| \right) \sup_{1 \leq i \leq N} \left(2 \|f_i - f_{i\epsilon}\|_{L^2(\mathbb{R}^3)}^2 (\|f_i\|_{L^2(\mathbb{R}^3)}^2 + \|f_{i\epsilon}\|_{L^2(\mathbb{R}^3)}^2) \right)^{\frac{1}{2}},
\end{aligned}$$

the term $\|f_i - f_{i\epsilon}\|_{L^2(\mathbb{R}^3)}$ converging to 0 with ϵ for every $1 \leq i \leq N$. This demonstrates the convergence of $Q_{N\epsilon}$ to Q_N in \mathfrak{S}_1 , hence in \mathfrak{S}_p , when ϵ goes to zero. As a consequence, $Q_{N\epsilon}$ converges to Q in \mathfrak{S}_p when ϵ goes to zero and N to infinity. Thanks to the above result, and by linearity, it suffices to prove (4.26) for $Q = |f\rangle\langle f|$ with $f \in H^2(\mathbb{R}^3)$. Then

$$\|(R_\epsilon - 1)|f\rangle\langle f|\|_{\mathfrak{S}_1} \leq \|(R_\epsilon - 1)f\|_{L^2} \|f\|_{L^2} \leq \epsilon \| |H_{\text{per}}^0 - \Sigma| f \|_{L^2} \|f\|_{L^2}$$

which converges to 0 as $\epsilon \rightarrow 0$ and controls $\|(R_\epsilon - 1)|f\rangle\langle f|\|_{\mathfrak{S}_p}$ for $1 \leq p < \infty$.

We are now in position to complete the proof of Lemma 2. First, by (4.105) (see Section 4.5, proof of Lemma 11),

$$\begin{aligned}
[R_\epsilon, |\nabla|] &= [(1 + \epsilon |H_{\text{per}}^0 - \Sigma|)^{-1}, |\nabla|], \\
&= \left[\frac{1}{\epsilon} \left(\frac{1}{\epsilon} + |H_{\text{per}}^0 - \Sigma| \right)^{-1}, |\nabla| \right], \\
&= \epsilon R_\epsilon [|H_{\text{per}}^0 - \Sigma|, |\nabla|] R_\epsilon
\end{aligned}$$

and therefore by Lemma 13 there exists a constant $C > 0$ such that

$$\| [R_\epsilon, |\nabla|] \| \leq C\epsilon. \quad (4.27)$$

Hence, $\lim_{\epsilon \rightarrow 0} \| [R_\epsilon, |\nabla|] \| = 0$. Let us prove that

$$\begin{aligned}
|\nabla|(R_\epsilon Q R_\epsilon - Q)^{-} |\nabla| &= |\nabla|((R_\epsilon - 1)Q^{-} R_\epsilon + Q^{-}(R_\epsilon - 1)) |\nabla| \\
&= [|\nabla|, R_\epsilon] Q^{-} [R_\epsilon, |\nabla|] + [|\nabla|, R_\epsilon] Q^{-} |\nabla| R_\epsilon \\
&\quad + (R_\epsilon - 1) |\nabla| Q^{-} [R_\epsilon, |\nabla|] + |\nabla| Q^{-} [R_\epsilon, |\nabla|] \\
&\quad + (R_\epsilon - 1) |\nabla| Q^{-} |\nabla| R_\epsilon + |\nabla| Q^{-} |\nabla| (R_\epsilon - 1).
\end{aligned}$$

By developing all the terms of the right-hand side expression

$$\begin{aligned}
[|\nabla|, R_\epsilon]Q^{--}[R_\epsilon, |\nabla|] &= (|\nabla|R_\epsilon - R_\epsilon|\nabla|)Q^{--}(R_\epsilon|\nabla| - |\nabla|R_\epsilon) \\
&= (|\nabla|R_\epsilon Q^{--} - R_\epsilon|\nabla|Q^{--})(R_\epsilon|\nabla| - |\nabla|R_\epsilon) \\
&= |\nabla|R_\epsilon Q^{--}R_\epsilon|\nabla| - |\nabla|R_\epsilon Q^{--}|\nabla|R_\epsilon \\
&\quad - R_\epsilon|\nabla|Q^{--}R_\epsilon|\nabla| + R_\epsilon|\nabla|Q^{--}|\nabla|R_\epsilon,
\end{aligned}$$

$$[|\nabla|, R_\epsilon]Q^{--}[R_\epsilon, |\nabla|] + [|\nabla|, R_\epsilon]Q^{--}|\nabla|R_\epsilon = |\nabla|R_\epsilon Q^{--}R_\epsilon|\nabla| - R_\epsilon|\nabla|Q^{--}R_\epsilon|\nabla|,$$

$$(R_\epsilon - 1)|\nabla|Q^{--}[R_\epsilon, |\nabla|] + |\nabla|Q^{--}[R_\epsilon, |\nabla|] = R_\epsilon|\nabla|Q^{--}R_\epsilon|\nabla| - R_\epsilon|\nabla|Q^{--}|\nabla|R_\epsilon,$$

$$\begin{aligned}
[|\nabla|, R_\epsilon]Q^{--}[R_\epsilon, |\nabla|] + [|\nabla|, R_\epsilon]Q^{--}|\nabla|R_\epsilon + (R_\epsilon - 1)|\nabla|Q^{--}[R_\epsilon, |\nabla|] + |\nabla|Q^{--}[R_\epsilon, |\nabla|] \\
= |\nabla|R_\epsilon Q^{--}R_\epsilon|\nabla| - R_\epsilon|\nabla|Q^{--}|\nabla|R_\epsilon,
\end{aligned}$$

$$(R_\epsilon - 1)|\nabla|Q^{--}|\nabla|R_\epsilon + |\nabla|Q^{--}|\nabla|(R_\epsilon - 1) = R_\epsilon|\nabla|Q^{--}|\nabla|R_\epsilon - |\nabla|Q^{--}|\nabla|.$$

Thus,

$$\begin{aligned}
[|\nabla|, R_\epsilon]Q^{--}[R_\epsilon, |\nabla|] + [|\nabla|, R_\epsilon]Q^{--}|\nabla|R_\epsilon + (R_\epsilon - 1)|\nabla|Q^{--}[R_\epsilon, |\nabla|] + |\nabla|Q^{--}[R_\epsilon, |\nabla|] \\
+ (R_\epsilon - 1)|\nabla|Q^{--}|\nabla|R_\epsilon + |\nabla|Q^{--}|\nabla|(R_\epsilon - 1) \\
= |\nabla|R_\epsilon Q^{--}R_\epsilon|\nabla| - |\nabla|Q^{--}|\nabla|, \\
= |\nabla|((R_\epsilon - 1)Q^{--}R_\epsilon + Q^{--}(R_\epsilon - 1))|\nabla|.
\end{aligned}$$

Applying either (4.27) or Lemma 3 to each term of the previous expression allows to conclude that

$$\lim_{\epsilon \rightarrow 0} \| |\nabla|(Q_\epsilon^{--} - Q^{--})|\nabla| \|_{\mathfrak{S}_1} = 0.$$

The proof is the same for the term $\| |\nabla|(Q_\epsilon^{++} - Q^{++})|\nabla| \|_{\mathfrak{S}_1}$.

Then, we have

$$\begin{aligned}
Q_\epsilon - Q &= R_\epsilon Q R_\epsilon - Q, \\
&= R_\epsilon Q (R_\epsilon - 1) + (R_\epsilon - 1)Q, \\
Q_\epsilon^{++} - Q^{++} &= (R_\epsilon Q R_\epsilon)^{++} - Q^{++}, \\
&= R_\epsilon Q^{++} R_\epsilon - Q^{++}, \\
&= R_\epsilon Q^{++} (R_\epsilon - 1) + (R_\epsilon - 1)Q^{++}, \\
Q_\epsilon^{--} - Q^{--} &= (R_\epsilon Q R_\epsilon)^{--} - Q^{--}, \\
&= R_\epsilon Q^{--} R_\epsilon - Q^{--}, \\
&= R_\epsilon Q^{--} (R_\epsilon - 1) + (R_\epsilon - 1)Q^{--}.
\end{aligned}$$

As R_ϵ is bounded and $Q \in \mathfrak{S}_2$, $Q^{++}, Q^{--} \in \mathfrak{S}_1$, and $(R_\epsilon - 1) \rightarrow 0$ when $\epsilon \rightarrow 0$, we obtain

$$\lim_{\epsilon \rightarrow 0} (\|Q_\epsilon - Q\|_{\mathfrak{S}_2} + \|(Q_\epsilon^{++} - Q^{++})\|_{\mathfrak{S}_1} + \|(Q_\epsilon^{--} - Q^{--})\|_{\mathfrak{S}_1}) = 0.$$

Finally,

$$\begin{aligned} |\nabla|(Q_\epsilon - Q) &= |\nabla|R_\epsilon Q R_\epsilon - |\nabla|Q, \\ &= [|\nabla|, R_\epsilon]Q R_\epsilon + R_\epsilon |\nabla|Q R_\epsilon - |\nabla|Q, \\ &= [|\nabla|, R_\epsilon]Q R_\epsilon + R_\epsilon |\nabla|Q (R_\epsilon - 1) + (R_\epsilon - 1)|\nabla|Q, \end{aligned}$$

which also converges to zero in \mathfrak{S}_2 for the same reasons as above.

Thus

$$\lim_{\epsilon \rightarrow 0} \|Q_\epsilon - Q\|_{\mathfrak{Q}} = 0.$$

□

Corollary 1. *Assume that (A1) holds true. When $Q \in \mathcal{Q}_r$, then $H_{\text{per}}^0 Q \in \mathfrak{S}_1^0$. For any $Q \in \mathfrak{Q}$, the expression (4.20) for $\text{Tr}_0(H_{\text{per}}^0 Q)$ does not depend on $\kappa \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. If $Q \in \mathcal{K}$, then*

$$\begin{aligned} 0 &\leq c_1 \text{Tr}((1 - \Delta)^{1/2} Q^2 (1 - \Delta)^{1/2}) \\ &\leq c_1 \text{Tr}((1 - \Delta)^{1/2} (Q^{++} - Q^{--}) (1 - \Delta)^{1/2}) \\ &\leq \text{Tr}_0(H_{\text{per}}^0 Q) - \kappa \text{Tr}_0(Q) \\ &\leq c_2 \text{Tr}((1 - \Delta)^{1/2} (Q^{++} - Q^{--}) (1 - \Delta)^{1/2}) \end{aligned} \quad (4.28)$$

where c_1 and c_2 are given by Lemma 1.

Proof. Let $Q \in \mathcal{Q}_r$ and $\kappa \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. Then $((H_{\text{per}}^0 - \kappa)Q)^{++} = |H_{\text{per}}^0 - \kappa|Q^{++} = |H_{\text{per}}^0 - \kappa|(1 - \Delta)^{-1}(1 - \Delta)Q^{++} \in \mathfrak{S}_1$ by Lemma 1 and the definition of \mathcal{Q}_r . A similar argument for $((H_{\text{per}}^0 - \kappa)Q)^{--}$ proves that $H_{\text{per}}^0 Q \in \mathfrak{S}_1^0$. Then for any $Q \in \mathcal{Q}_r$, (4.28) is a straightforward consequence of (4.22) and (4.17). We conclude using the density of \mathcal{Q}_r in \mathfrak{Q} and the density of \mathcal{K}_r in \mathcal{K} . The following is an adaptation of [22, Thm 1] :

Corollary 2. *Let $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, $Z \in \mathbb{N} \setminus \{0\}$ and assume that (A1) holds. For any $\kappa \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, one has for some $d_1, d_2 > 0$*

$$\begin{aligned} \forall Q \in \mathcal{K}, \quad \mathcal{E}^\nu(Q) - \kappa \text{Tr}_0(Q) &\geq d_1 \left(\|Q^{++}\|_{\mathfrak{S}_1} + \|Q^{--}\|_{\mathfrak{S}_1} \right. \\ &\quad \left. + \||\nabla|Q^{++}|\nabla|\|_{\mathfrak{S}_1} + \||\nabla|Q^{--}|\nabla|\|_{\mathfrak{S}_1} \right) + d_2 (\|Q\|_{\mathfrak{S}_2}^2 + \||\nabla|Q\|_{\mathfrak{S}_2}^2) - \frac{1}{2}D(\nu, \nu) \end{aligned} \quad (4.29)$$

Hence $\mathcal{E}^\nu - \kappa \text{Tr}_0$ is bounded from below and coercive on \mathcal{K} . Additionally, when $\nu \equiv 0$, $Q \mapsto \mathcal{E}^0(Q) - \kappa \text{Tr}_0(Q)$ is nonnegative, 0 being its unique minimizer.

Proof. Inequality (4.29) is a straightforward consequence of (4.28) and the fact that $D(\cdot, \cdot)$ defines an inner product on \mathcal{C} :

$$\begin{aligned} \mathcal{E}^\nu(Q) - \kappa \text{Tr}_0(Q) &= \text{Tr}_0((H_{\text{per}}^0 - \kappa)Q) - D(\rho_Q, \nu) + \frac{1}{2}D(\rho_Q, \rho_Q), \\ &= \text{Tr}(|H_{\text{per}}^0 - \kappa|^{\frac{1}{2}}(Q^{++} - Q^{--})|H_{\text{per}}^0 - \kappa|^{\frac{1}{2}}) - D(\rho_Q, \nu) + \frac{1}{2}D(\rho_Q, \rho_Q), \\ &\geq c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) - D(\rho_Q, \nu) + \frac{1}{2}D(\rho_Q, \rho_Q). \end{aligned}$$

Besides,

$$-D(\rho_Q, \nu) + \frac{1}{2}D(\rho_Q, \rho_Q) \geq -\frac{1}{2}D(\nu, \nu),$$

So it remains to prove that

$$\begin{aligned} c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) &\geq d_1 \left(\|Q^{++}\|_{\mathfrak{S}_1} + \|Q^{--}\|_{\mathfrak{S}_1} + \|\nabla|Q^{++}|\nabla\|_{\mathfrak{S}_1} \right. \\ &\quad \left. + \|\nabla|Q^{--}|\nabla\|_{\mathfrak{S}_1} \right) + d_2 (\|Q\|_{\mathfrak{S}_2}^2 + \|\nabla|Q|\nabla\|_{\mathfrak{S}_2}^2) \end{aligned} \quad (4.30)$$

Applying $(1 - \Delta)^{\frac{1}{2}}$ is equivalent, up to a multiplicative constant, to multiplying the k -th Fourier mode by $(1 + |k|^2)^{\frac{1}{2}}$, which is greater than $|k|$, the symbol of the operator $|\nabla|$.

Thus

$$\begin{aligned} c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) &\geq c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^2)(1 - \Delta)^{\frac{1}{2}}) \geq c_3 \text{Tr}(Q^2) = c_3 \|Q\|_{\mathfrak{S}_2}^2, \\ c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) &\geq c_4 \text{Tr}(|\nabla|Q^{++}|\nabla|) = c_4 \|\nabla|Q^{++}|\nabla\|_{\mathfrak{S}_1}, \\ c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) &\geq c_4 \text{Tr}(|\nabla|Q^{--}|\nabla|) = c_4 \|\nabla|Q^{--}|\nabla\|_{\mathfrak{S}_1}, \\ c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}Q^2(1 - \Delta)^{\frac{1}{2}}) &\geq c_5 \text{Tr}(|\nabla|^2 Q^2) = c_5 \|\nabla|Q|\nabla\|_{\mathfrak{S}_2}^2, \\ c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) &\geq c_6 \text{Tr}(Q^{++}) = \|Q^{++}\|_{\mathfrak{S}_1}, \\ c_1 \text{Tr}((1 - \Delta)^{\frac{1}{2}}(Q^{++} - Q^{--})(1 - \Delta)^{\frac{1}{2}}) &\geq c_6 \text{Tr}(-Q^{--}) = \|Q^{--}\|_{\mathfrak{S}_1}, \end{aligned}$$

so by taking $d_1 = \min(c_4, c_6)$ and $d_2 = \min(c_3, c_5)$, we obtain (4.30).

If $\nu \equiv 0$, by (4.29), it is easily seen that $Q \mapsto \mathcal{E}^0(Q) - \kappa \text{Tr}_0(Q)$ is nonnegative, and equals to 0, if and only if $\|Q\|_{\mathcal{K}} = 0$, ie. $Q = 0$.

Remark 1. The energy $\mathcal{E}^\nu(Q)$ measures the energy of a state $\gamma = \gamma_{\text{per}}^0 + Q$ with respect to that of γ_{per}^0 . Thus the last statement of Corollary 2 is another way of expressing the fact that γ_{per}^0 is the state of lowest energy of the periodic system when there is no defect.

4.3.2 Existence of minimizers with a chemical potential

In view of Corollary 2, it is natural to introduce the following minimization problem

$$\boxed{E_{\epsilon_F}^\nu := \inf\{\mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}_0(Q), Q \in \mathcal{K}\} > -\infty} \quad (4.31)$$

for any Fermi level $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. The following result will be proved following ideas from [23] :

Theorem 2 (Existence of minimizers with a chemical potential). *Let $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, $Z \in \mathbb{N} \setminus \{0\}$ and assume that **(A1)** holds. Then for any $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, there exists a minimizer $\bar{Q} \in \mathcal{K}$ for (4.31). Problem (4.31) may have*

several minimizers, but they all share the same density $\bar{\rho} = \rho_{\bar{Q}}$. Any minimizer \bar{Q} of (4.31) satisfies the self-consistent equation

$$\begin{cases} \bar{Q} = \chi_{(-\infty, \epsilon_F)}(H_{\bar{Q}}) - \gamma_{\text{per}}^0 + \delta, \\ H_{\bar{Q}} = H_{\text{per}}^0 + (\rho_{\bar{Q}} - \nu) * |\cdot|^{-1} \end{cases} \quad (4.32)$$

where δ is a finite rank self-adjoint operator satisfying $0 \leq \delta \leq 1$ and $\text{Ran}(\delta) \subseteq \ker(H_{\bar{Q}} - \epsilon_F)$.

Remark 2. Before proving this theorem, let us show that every $V \in L^6(\mathbb{R}^3)$, and therefore *a fortiori* every $V \in \mathcal{C}'$, is a compact perturbation of H_{per}^0 .

To this aim, let us take a sequence $u_n \in H^2(\mathbb{R}^3)$ that converges weakly to u^* in $H^2(\mathbb{R}^3)$. Without restriction, we can assume that (u_n) converges almost everywhere in \mathbb{R}^3 . We then set $v_n = Vu_n$. We are going to prove that (v_n) converges to $v^* = Vu^*$ in $L^2(\mathbb{R}^3)$.

Let $\epsilon > 0$ and $K_\epsilon = \{x \in \mathbb{R}^3, |V(x)| \leq c_\epsilon = \frac{\epsilon}{M}\}$, where we can choose M such that $\|u_n - u^*\|_{L^2(\mathbb{R}^3)} \leq M$ and $\|u_n\|_{L^\infty} \leq M$ since (u_n) is bounded in $H^2(\mathbb{R}^3)$. We have $|\mathbb{R}^3 \setminus K_\epsilon| < \infty$; otherwise V would not be in $L^6(\mathbb{R}^3)$. Thus

$$\begin{aligned} \|v_n - v^*\|_{L^2(\mathbb{R}^3)}^2 &= \int_{\mathbb{R}^3} V^2(x)(u_n(x) - u^*(x))^2 dx, \\ &= \int_{K_\epsilon} V^2(x)(u_n(x) - u^*(x))^2 dx + \int_{\mathbb{R}^3 \setminus K_\epsilon} V^2(x)(u_n(x) - u^*(x))^2 dx, \\ &\leq \frac{\epsilon^2}{M^2} \int_{K_\epsilon} ((u_n - u^*)(x))^2 dx + \int_{\mathbb{R}^3 \setminus K_\epsilon} V^2(x)((u_n - u^*)(x))^2 dx. \end{aligned}$$

The first term is bounded by ϵ^2 . For the second term, we have for almost all $x \in \mathbb{R}^3 \setminus K_\epsilon$

$$V^2(x)((u_n - u^*)(x))^2 \leq 4M^2V^2(x)$$

together with

$$\int_{\mathbb{R}^3 \setminus K_\epsilon} V^2 \leq \|V\|_{L^6(\mathbb{R}^3)}^2 |\mathbb{R}^3 \setminus K_\epsilon|^{2/3} < \infty.$$

Besides, $V^2(x)((u_n - u^*)(x))^2 \rightarrow 0$ almost everywhere on $\mathbb{R}^3 \setminus K_\epsilon$. By dominated convergence, we have $\int_{\mathbb{R}^3 \setminus K_\epsilon} V^2(x)((u_n - u^*)(x))^2 dx \rightarrow 0$ when n goes to infinity. For n large enough, we thus have

$$\|v_n - v^*\|_{L^2(\mathbb{R}^3)}^2 \leq 2\epsilon^2,$$

yielding the convergence of (v_n) to v^* , hence the fact that V is a compact perturbation of H_{per}^0 .

In particular, $(\rho_{\bar{Q}} - \nu) * |\cdot|^{-1}$ is a compact perturbation of H_{per}^0 , implying that $H_{\bar{Q}}$ is self-adjoint on $L^2(\mathbb{R}^3)$ with domain $D(H_{\text{per}}^0) = D(-\Delta) = H^2(\mathbb{R}^3)$ and that $\sigma_{\text{ess}}(H_{\bar{Q}}) = \sigma(H_{\text{per}}^0)$. Thus the discrete spectrum of $H_{\bar{Q}}$ is either empty or composed of isolated eigenvalues of finite multiplicity, possibly accumulating at the ends of the bands.

Proof of Theorem 2. Let $(Q_n)_{n \in \mathbb{N}}$ be a minimizing sequence for

$$\mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}_0(Q), Q \in \mathcal{K} \quad (4.33)$$

It follows from Corollary 2 that $(Q_n)_{n \in \mathbb{N}}$ is bounded in \mathcal{Q} . By Proposition 1, $(\rho_{Q_n})_{n \in \mathbb{N}}$ is bounded in $\mathcal{C} \cap L^2(\mathbb{R}^3)$. Thus, up to extraction, we can assume that there exists \bar{Q} in the convex set \mathcal{K} such that

- i) $Q_n \rightharpoonup \bar{Q}$ and $|\nabla|Q_n \rightharpoonup |\nabla|\bar{Q}$ weakly in \mathfrak{S}_2 ;
- ii) $|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2} \rightharpoonup |H_{\text{per}}^0 - \epsilon_F|^{1/2} \bar{Q}^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2}$,
 $|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{--} |H_{\text{per}}^0 - \epsilon_F|^{1/2} \rightharpoonup |H_{\text{per}}^0 - \epsilon_F|^{1/2} \bar{Q}^{--} |H_{\text{per}}^0 - \epsilon_F|^{1/2}$
for the weak-* topology of \mathfrak{S}_1 ;
- iii) $\rho_{Q_n} \rightharpoonup \rho_{\bar{Q}}$ in $\mathcal{C} \cap L^2(\mathbb{R}^3)$.

Let us justify and detail the above three assertions.

i) First, both (Q_n) and $(|\nabla|Q_n)$ weakly converge in \mathfrak{S}_2 (up to extraction) because of the boundedness of (Q_n) in \mathcal{Q} . So there exist \bar{Q} and \bar{G} such that $Q_n \rightharpoonup \bar{Q}$ and $|\nabla|Q_n \rightharpoonup \bar{G}$ weakly in \mathfrak{S}_2 . Let Φ be an integral operator on $L^2(\mathbb{R}^3)$ with kernel in $\mathcal{C}_0^\infty(\mathbb{R}^3 \times \mathbb{R}^3)$. It holds

$$\begin{aligned} (|\nabla|Q_n, \Phi)_{\mathfrak{S}_2} &= (Q_n, |\nabla|\Phi)_{\mathfrak{S}_2}, \\ &\rightarrow (\bar{Q}, |\nabla|\Phi)_{\mathfrak{S}_2}. \end{aligned}$$

Hence $\forall \Phi \in \mathfrak{S}_2$ with kernel in $\mathcal{C}_0^\infty(\mathbb{R}^3 \times \mathbb{R}^3)$,

$$(\bar{G}, \Phi)_{\mathfrak{S}_2} = (\bar{Q}, |\nabla|\Phi)_{\mathfrak{S}_2}.$$

This implies that $|\nabla|\bar{Q} = \bar{G} \in \mathfrak{S}_2$ and that $|\nabla|Q_n \rightharpoonup |\nabla|\bar{Q}$ in \mathfrak{S}_2 .

ii) Recall that \mathfrak{S}_1 is the dual of the space of compact operators \mathfrak{S}_∞ . Thus here $A_n \rightharpoonup A$ for the weak-* topology of \mathfrak{S}_1 means $\text{Tr}(A_n K) \rightarrow \text{Tr}(AK)$ for any compact operator K . As $|H_{\text{per}}^0 - \epsilon_F|(1-\Delta)^{-1}$ is a bounded operator and as $(1-\Delta)Q_n^{++}$ is bounded in \mathfrak{S}_1 , $|H_{\text{per}}^0 - \epsilon_F|^{1/2}(1-\Delta)^{-1/2}(1-\Delta)^{1/2}Q_n^{++}(1-\Delta)^{1/2}(1-\Delta)^{-1/2}|H_{\text{per}}^0 - \epsilon_F|^{1/2}$ is bounded in \mathfrak{S}_1 and therefore there exists $P \in \mathfrak{S}_1$ such that, up to extraction

$$|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2} \rightharpoonup P$$

for the weak-* topology of \mathfrak{S}_1 . To show that $P = |H_{\text{per}}^0 - \epsilon_F|^{1/2} \bar{Q}^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2}$, let us consider $\varphi \in L^2(\mathbb{R}^3)$, $\psi \in L^2(\mathbb{R}^3)$ and the compact operator $K = |H_{\text{per}}^0 - \epsilon_F|^{-1/2} |\varphi\rangle\langle\psi| |H_{\text{per}}^0 - \epsilon_F|^{-1/2}$, for which on the one hand

$$\begin{aligned} \text{Tr}(|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2} K) \\ = \langle\psi| Q_n^{++} |\varphi\rangle \rightarrow \langle\psi| \bar{Q}^{++} |\varphi\rangle \end{aligned}$$

and on the other hand

$$\begin{aligned} \text{Tr}(|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2} K) \\ \rightarrow \text{Tr}(PK) \\ = \langle\psi| |H_{\text{per}}^0 - \epsilon_F|^{-1/2} P |H_{\text{per}}^0 - \epsilon_F|^{-1/2} |\varphi\rangle. \end{aligned}$$

It follows that $\bar{Q}^{++} = |H_{\text{per}}^0 - \epsilon_F|^{-1/2} P |H_{\text{per}}^0 - \epsilon_F|^{-1/2}$, or in other words that

$$P = |H_{\text{per}}^0 - \epsilon_F|^{1/2} \bar{Q}^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2}.$$

The proof is the same for Q^{--} .

iii) As we saw before, ρ_{Q_n} is bounded in $\mathcal{C} \cap L^2(\mathbb{R}^3)$ and therefore converges up to extraction weakly in $\mathcal{C} \cap L^2(\mathbb{R}^3)$ to some $\rho_{\bar{Q}}$, which is in fact $\rho_{\bar{Q}}$ thanks to the fact that the map $Q \mapsto \rho_Q$ is linear and continuous from \mathcal{Q} to $\mathcal{C} \cap L^2(\mathbb{R}^3)$.

As $Q_n \rightharpoonup \bar{Q} \in \mathfrak{S}_2$, we have $(Q_n, F) \rightarrow (\bar{Q}, F)$, for all operator $F := |f\rangle\langle f|$ with $f \in L^2(\mathbb{R}^3)$, which means

$$\langle f | Q_n | f \rangle \rightarrow \langle f | \bar{Q} | f \rangle.$$

As $Q_n \in \mathcal{K}$, $-\gamma_{\text{per}}^0 \leq Q_n \leq 1 - \gamma_{\text{per}}^0$, so

$$-\langle f | \gamma_{\text{per}}^0 | f \rangle \leq \langle f | Q_n | f \rangle \leq \langle f | 1 - \gamma_{\text{per}}^0 | f \rangle.$$

Passing to the limit, we obtain

$$\forall f \in L^2(\mathbb{R}^3), \quad -\langle f | \gamma_{\text{per}}^0 | f \rangle \leq \langle f | \bar{Q} | f \rangle \leq \langle f | 1 - \gamma_{\text{per}}^0 | f \rangle,$$

which implies that $\bar{Q} \in \mathcal{K}$.

Then, as $D(\cdot, \cdot)$ defines an inner product on \mathcal{C} , $\tilde{D}(f) = D(f, f)$ is convex (and even α -convex) on \mathcal{C} .

The quadratic form \tilde{D} being convex and continuous on \mathcal{C} , it is lower semicontinuous for the weak topology and we have :

$$D(\rho_{\bar{Q}} - \nu, \rho_{\bar{Q}} - \nu) \leq \liminf_{n \rightarrow +\infty} D(\rho_{Q_n} - \nu, \rho_{Q_n} - \nu).$$

Now since $Q_n^{++} \geq 0$, $|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2}$ is also a nonnegative operator for any n . Thus Fatou's Lemma [45] yields

$$\text{Tr}(|H_{\text{per}}^0 - \epsilon_F|^{1/2} \bar{Q}^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2}) \leq \liminf_{n \rightarrow \infty} \text{Tr}(|H_{\text{per}}^0 - \epsilon_F|^{1/2} Q_n^{++} |H_{\text{per}}^0 - \epsilon_F|^{1/2}).$$

The same argument for the term involving $-\bar{Q}^{--} \geq 0$ yields

$$\mathcal{E}^\nu(\bar{Q}) - \epsilon_F \text{Tr}_0(\bar{Q}) \leq \liminf_{n \rightarrow \infty} (\mathcal{E}^\nu(Q_n) - \epsilon_F \text{Tr}_0(Q_n)) = E_{\epsilon_F}^\nu,$$

i.e. $\bar{Q} \in \mathcal{K}$ is a minimizer.

Let us show that \bar{Q} satisfies the self-consistent equation

$$\bar{Q} = \chi_{(-\infty, \epsilon_F)}(H_{\bar{Q}}) - \gamma_{\text{per}}^0 + \delta. \quad (4.34)$$

For this purpose, we first make use of the convexity of \mathcal{K} and write that for any $Q \in \mathcal{K}$ and any $t \in [0, 1]$,

$$\mathcal{E}^\nu((1-t)\bar{Q} + tQ) - \epsilon_F \text{Tr}_0((1-t)\bar{Q} + tQ) \geq \mathcal{E}^\nu(\bar{Q}) - \epsilon_F \text{Tr}_0(\bar{Q}).$$

We have

$$\begin{aligned} \mathcal{E}^\nu((1-t)\bar{Q} + tQ) - \mathcal{E}^\nu(\bar{Q}) &= \text{Tr}_0(H_{\text{per}}^0((1-t)\bar{Q} + tQ)) - \text{Tr}_0(H_{\text{per}}^0\bar{Q}) \\ &\quad - D(\rho_{(1-t)\bar{Q}+tQ}, \nu) + D(\rho_{\bar{Q}}, \nu) + \frac{1}{2}D(\rho_{(1-t)\bar{Q}+tQ}, \rho_{(1-t)\bar{Q}+tQ}) - \frac{1}{2}D(\rho_{\bar{Q}}, \rho_{\bar{Q}}), \end{aligned}$$

Let us rewrite the first two terms of the above expression as

$$\text{Tr}_0(H_{\text{per}}^0((1-t)\bar{Q} + tQ)) - \text{Tr}_0(H_{\text{per}}^0\bar{Q}) = t(\text{Tr}_0(H_{\text{per}}^0Q) - \text{Tr}_0(H_{\text{per}}^0\bar{Q})).$$

The next two terms read as

$$\begin{aligned} -D(\rho_{(1-t)\bar{Q}+tQ}, \nu) + D(\rho_{\bar{Q}}, \nu) &= -D((1-t)\rho_{\bar{Q}} + t\rho_Q, \nu) + D(\rho_{\bar{Q}}, \nu), \\ &= t(D(\rho_{\bar{Q}}, \nu) - D(\rho_Q, \nu)), \end{aligned}$$

and the last two terms as

$$\begin{aligned} D(\rho_{(1-t)\bar{Q}+tQ}, \rho_{(1-t)\bar{Q}+tQ}) - D(\rho_{\bar{Q}}, \rho_{\bar{Q}}) &= (1-t)^2D(\rho_{\bar{Q}}, \rho_{\bar{Q}}) + 2t(1-t)D(\rho_{\bar{Q}}, \rho_Q) \\ &\quad + t^2D(\rho_Q, \rho_Q) - D(\rho_{\bar{Q}}, \rho_{\bar{Q}}), \\ &= t^2(D(\rho_{\bar{Q}}, \rho_{\bar{Q}}) + D(\rho_Q, \rho_Q) - 2D(\rho_{\bar{Q}}, \rho_Q)) \\ &\quad + 2t(D(\rho_{\bar{Q}}, \rho_Q) - D(\rho_{\bar{Q}}, \rho_{\bar{Q}})). \end{aligned}$$

Thus we obtain

$$\begin{aligned} \mathcal{E}^\nu((1-t)\bar{Q} + tQ) - \epsilon_F \text{Tr}_0((1-t)\bar{Q} + tQ) - (\mathcal{E}^\nu(\bar{Q}) - \epsilon_F \text{Tr}_0(\bar{Q})) \\ &= t(\text{Tr}_0(H_{\text{per}}^0Q) - \text{Tr}_0(H_{\text{per}}^0\bar{Q}) - \epsilon_F(\text{Tr}_0(Q) - \text{Tr}_0(\bar{Q}))) \\ &\quad + t(D(\rho_{\bar{Q}}, \nu) - D(\rho_Q, \nu) + D(\rho_{\bar{Q}}, \rho_Q) - D(\rho_{\bar{Q}}, \rho_{\bar{Q}})) \\ &\quad + \frac{t^2}{2}(D(\rho_{\bar{Q}}, \rho_{\bar{Q}}) + D(\rho_Q, \rho_Q) - 2D(\rho_{\bar{Q}}, \rho_Q)). \end{aligned}$$

This expression being nonnegative for all $Q \in \mathcal{K}$ and all $t \in [0, 1]$, we obtain that for all $Q \in \mathcal{K}$

$$\begin{aligned} \text{Tr}_0(H_{\text{per}}^0Q) - \text{Tr}_0(H_{\text{per}}^0\bar{Q}) - \epsilon_F(\text{Tr}_0(Q) - \text{Tr}_0(\bar{Q})) \\ + D(\rho_{\bar{Q}}, \nu) - D(\rho_Q, \nu) + D(\rho_{\bar{Q}}, \rho_Q) - D(\rho_{\bar{Q}}, \rho_{\bar{Q}}) \geq 0. \end{aligned} \quad (4.35)$$

Denoting by

$$F(Q) := \text{Tr}_0((H_{\text{per}}^0 - \epsilon_F)Q) + D(\rho_{\bar{Q}} - \nu, \rho_Q)$$

(4.35) also reads

$$\begin{aligned} F(Q) - F(\bar{Q}) &= \text{Tr}_0(H_{\text{per}}^0Q) - \text{Tr}_0(H_{\text{per}}^0\bar{Q}) - \epsilon_F(\text{Tr}_0(Q) - \text{Tr}_0(\bar{Q})) \\ &\quad + D(\rho_{\bar{Q}}, \rho_Q) - D(\rho_{\bar{Q}}, \rho_{\bar{Q}}) - D(\rho_Q, \nu) + D(\rho_{\bar{Q}}, \nu) \geq 0. \end{aligned}$$

As $D(\rho_{\bar{Q}} - \nu, \rho_Q) = \text{Tr}_0(((\rho_{\bar{Q}} - \nu) * |\cdot|^{-1})Q)$ we readily obtain that \bar{Q} is a minimizer on \mathcal{K} of

$$F(Q) = \text{Tr}_0((H_{\bar{Q}} - \epsilon_F)Q).$$

Let $\bar{Q}' = \chi_{(-\infty, \epsilon_F)}(H_{\bar{Q}}) - \gamma_{\text{per}}^0$. Using the same arguments as in the proof of Proposition 2 below, we have $\bar{Q}' \in \mathcal{K}$. Any $Q \in \mathcal{K}$ is of the form $Q = \gamma - \gamma_{\text{per}}^0$ with $0 \leq \gamma \leq 1$, and it holds

$$\begin{aligned} \text{Tr}_0((H_{\bar{Q}} - \epsilon_F)Q) - \text{Tr}_0((H_{\bar{Q}} - \epsilon_F)\bar{Q}') &= \text{Tr}_0((H_{\bar{Q}} - \epsilon_F)(Q - \bar{Q}')) \\ &= \text{Tr}_0((H_{\bar{Q}} - \epsilon_F)(\gamma - \chi_{(-\infty, \epsilon_F)}(H_{\bar{Q}}))). \end{aligned}$$

Therefore,

$$\text{Tr}_0((H_{\bar{Q}} - \epsilon_F)Q) - \text{Tr}_0((H_{\bar{Q}} - \epsilon_F)\bar{Q}') \geq 0$$

with inequality if and only if

$$\gamma = \chi_{(-\infty, \epsilon_F)}(H_{\bar{Q}}) + \delta$$

with $0 \leq \delta \leq 1$ and $\text{Ran}(\delta) \subseteq \ker(H_{\bar{Q}} - \epsilon_F)$. Hence, \bar{Q} is solution to the self-consistent field equation (4.32).

The uniqueness of the density remains to be proved. Let us assume that there exist two minimizers Q_1 and Q_2 with different densities ρ_{Q_1} and ρ_{Q_2} . Then, if we compute the energy of $\frac{Q_1+Q_2}{2}$, we get, using the strict convexity of D ,

$$\begin{aligned} &\mathcal{E}^\nu \left(\frac{Q_1 + Q_2}{2} \right) - \epsilon_F \text{Tr}_0 \left(\frac{Q_1 + Q_2}{2} \right) \\ &= \text{Tr}_0 \left(H_{\text{per}}^0 \left(\frac{Q_1 + Q_2}{2} \right) \right) - D \left(\rho_{\frac{Q_1+Q_2}{2}}, \nu \right) + \frac{1}{2} D \left(\rho_{\frac{Q_1+Q_2}{2}}, \rho_{\frac{Q_1+Q_2}{2}} \right) - \epsilon_F \text{Tr}_0 \left(\frac{Q_1 + Q_2}{2} \right), \\ &= \frac{1}{2} \text{Tr}_0(H_{\text{per}}^0 Q_1) + \frac{1}{2} \text{Tr}_0(H_{\text{per}}^0 Q_2) - \frac{1}{2} D(\rho_{Q_1}, \nu) - \frac{1}{2} D(\rho_{Q_2}, \nu) \\ &\quad + \frac{1}{2} D \left(\frac{1}{2} \rho_{Q_1} + \frac{1}{2} \rho_{Q_2}, \frac{1}{2} \rho_{Q_1} + \frac{1}{2} \rho_{Q_2} \right) - \epsilon_F \frac{1}{2} \text{Tr}_0(Q_1) - \epsilon_F \frac{1}{2} \text{Tr}_0(Q_2), \\ &< \frac{1}{2} \text{Tr}_0(H_{\text{per}}^0 Q_1) + \frac{1}{2} \text{Tr}_0(H_{\text{per}}^0 Q_2) - \frac{1}{2} D(\rho_{Q_1}, \nu) - \frac{1}{2} D(\rho_{Q_2}, \nu) \\ &\quad + \frac{1}{4} D(\rho_{Q_1}, \rho_{Q_1}) + \frac{1}{4} D(\rho_{Q_2}, \rho_{Q_2}) - \epsilon_F \frac{1}{2} \text{Tr}_0(Q_1) - \epsilon_F \frac{1}{2} \text{Tr}_0(Q_2), \\ &= \frac{1}{2} \left(\mathcal{E}^\nu(Q_1) - \epsilon_F \text{Tr}_0(Q_1) \right) + \frac{1}{2} \left(\mathcal{E}^\nu(Q_2) - \epsilon_F \text{Tr}_0(Q_2) \right) \\ &= E_{\epsilon_F}^\nu, \end{aligned}$$

which is in contradiction with the fact that $E_{\epsilon_F}^\nu$ is the minimum. \square

Recall that the charge of the minimizing state \bar{Q} obtained in Theorem 2 is defined as $\text{Tr}_0(\bar{Q})$. Similarly to [22, 23], it can be proved by perturbation theory that for any fixed $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, there exists a constant $C(\epsilon_F) > 0$ such that when $D(\nu, \nu) \leq C(\epsilon_F)$, one has $\ker(H_{\bar{Q}} - \epsilon_F) = \{0\}$ and $\text{Tr}_0(\bar{Q}) = 0$, i.e. the minimizer of

the free energy with chemical potential ϵ_F is a neutral perturbation of the periodic Fermi sea.

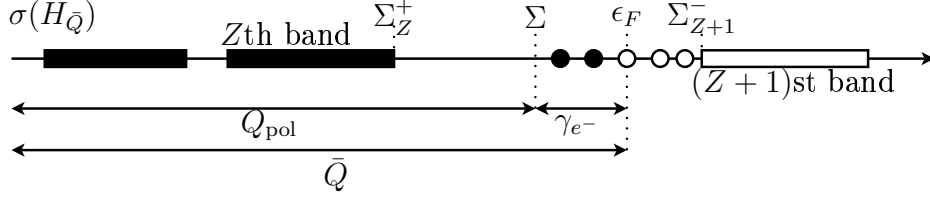


FIG. 4.1 – Decomposition $\bar{Q} = Q_{\text{pol}} + \gamma_{e^-}$ for not too strong a positively charged nuclear defect ($\nu \geq 0$).

For a fixed external density ν and an adequately chosen chemical potential ϵ_F , one can have $\text{Tr}_0(\bar{Q}) \neq 0$ meaning either that electron-hole pairs have been created from the Fermi sea, and/or that the system of lowest energy contains a finite number of bound electrons or holes close to the defect. In the applications, one will usually have for a positively charged nuclear defect ($\nu \geq 0$) that the spectrum of $H_{\bar{Q}}$ contains a sequence of eigenvalues converging to the bottom Σ_{Z+1}^- of the lowest unfilled band (conduction band), and that ϵ_F is chosen such that exactly q eigenvalues are filled, corresponding to q bound electrons :

$$\bar{Q} = (\chi_{(-\infty, \Sigma)}(H_{\bar{Q}}) - \gamma_{\text{per}}^0) + (\chi_{[\Sigma, \epsilon_F]}(H_{\bar{Q}}) + \delta) := Q_{\text{pol}} + \gamma_{e^-}, \quad (4.36)$$

where we recall that Σ is the center of the gap

$$\Sigma := \frac{\Sigma_Z^+ + \Sigma_{Z+1}^-}{2}.$$

For not too strong a defect density ν , one has $\text{Tr}_0(Q_{\text{pol}}) = 0$. Hence $\text{Tr}(\gamma_{e^-}) = q$. Let us assume for simplicity that $\delta = 0$ and that $q \in \mathbb{N} \setminus \{0\}$. Then

$$\gamma_{e^-} = \chi_{[\Sigma, \epsilon_F]}(H_{\bar{Q}}) = \sum_{n=1}^q |\varphi_n\rangle\langle\varphi_n| \quad (4.37)$$

where (φ_n) are eigenfunctions of $H_{\bar{Q}}$ corresponding to its q first eigenvalues in $[\Sigma, \epsilon_F)$:

$$H_{\bar{Q}}\varphi_n = \lambda_n\varphi_n. \quad (4.38)$$

Notice that

$$H_{\bar{Q}} = -\Delta/2 + (\rho_{\gamma_{e^-}} - \nu) * |\cdot|^{-1} + V_{\text{pol}} \quad (4.39)$$

where

$$V_{\text{pol}} = (\rho_{\gamma_{\text{per}}} - \mu_{\text{per}}) \star_{\Gamma} G_1 + \rho_{Q_{\text{pol}}} * |\cdot|^{-1}$$

is the polarization potential created by the self-consistent Fermi sea and seen by the q electrons. Thus the q electrons solve a usual reduced Hartree-Fock equation (4.38) in which the mean-field operator (4.39) additionally contains the self-consistent polarization of the medium.

The interpretation given in the previous paragraph is different if the positive density of charge ν of the defect is strong enough to create an electron-hole pair from the Fermi sea.

We end this section by specifying the regularity of solutions of (4.32).

Proposition 2. *Let $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, $Z \in \mathbb{N} \setminus \{0\}$ and assume that (A1) holds. Any $Q \in \mathcal{K}$ solution of the self-consistent equation (4.32) belongs to \mathcal{K}_r .*

Proof. Let $Q \in \mathcal{Q}$ be of the form

$$Q = \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) - \gamma_{\text{per}}^0 + \delta$$

where $0 \leq \delta \leq 1$ is a finite rank operator with $\text{Ran}(\delta) \subseteq \ker(H_{\text{per}}^0 + V - \epsilon_F)$ and $V = \rho * |\cdot|^{-1}$ for some $\rho \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ (in our case $\rho = \rho_Q - \nu$). As $\rho \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) \subset L^{6/5}(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) \subset \mathcal{C}$, we know (see the previous section) that $V \in \tilde{\mathcal{C}} \subset L^6(\mathbb{R}^3)$. Besides, the Fourier transform of $|\cdot|^{-1}$ being $\frac{1}{(2\pi)^{3/2}} \frac{4\pi}{|k|^2}$, the Fourier transform of $V = \rho * |\cdot|^{-1}$ is $4\pi \frac{\hat{\rho}(k)}{|k|^2}$ with $\hat{\rho} \in L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$. Consequently $V \in C^0(\mathbb{R}^3)$ (and vanishes at infinity). Hence $V \in L^6(\mathbb{R}^3) \cap C^0(\mathbb{R}^3)$ and it follows from Remark 2 that V is a compact perturbation of H_{per}^0 . Since $\ker(H_{\text{per}}^0 + V - \epsilon_F) \subseteq D(H_{\text{per}}^0 + V) = D(H_{\text{per}}^0) = H^2(\mathbb{R}^3)$, it is clear that the finite rank operator δ satisfies $(1 - \Delta)\delta^2(1 - \Delta) \in \mathfrak{S}_1$. Thus, there is no restriction in assuming that $\delta = 0$ and consider the case when

$$Q = \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) - \gamma_{\text{per}}^0.$$

Then we remark that $Q^2 = Q^{++} - Q^{--}$: indeed

$$\begin{aligned} Q^2 &= (\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) - \gamma_{\text{per}}^0)^2, \\ &= \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) + \gamma_{\text{per}}^0 - \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V)\gamma_{\text{per}}^0 - \gamma_{\text{per}}^0\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V), \\ Q^{++} &= (1 - \gamma_{\text{per}}^0)(\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) - \gamma_{\text{per}}^0)(1 - \gamma_{\text{per}}^0), \\ &= (1 - \gamma_{\text{per}}^0)(\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V))(1 - \gamma_{\text{per}}^0), \\ &= \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) + \gamma_{\text{per}}^0\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V)\gamma_{\text{per}}^0 \\ &\quad - \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V)\gamma_{\text{per}}^0 - \gamma_{\text{per}}^0\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V), \\ Q^{--} &= \gamma_{\text{per}}^0\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V)\gamma_{\text{per}}^0 - \gamma_{\text{per}}^0, \end{aligned}$$

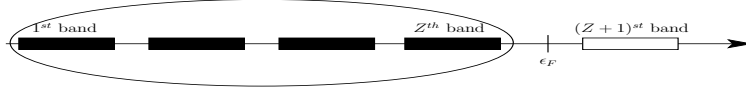
so that

$$\begin{aligned} Q^{++} - Q^{--} &= \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) - \gamma_{\text{per}}^0\chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V) - \chi_{(-\infty, \epsilon_F)}(H_{\text{per}}^0 + V)\gamma_{\text{per}}^0 + \gamma_{\text{per}}^0, \\ &= Q^2. \end{aligned}$$

Hence it only remains to prove that $(1 - \Delta)Q \in \mathfrak{S}_2$.

Let \mathcal{C} be a smooth curve enclosing the whole spectrum of $H_{\text{per}}^0 + V$ below ϵ_F (ie. the Z first bands). Since $V \in L^\infty(\mathbb{R}^3)$ and $|z - H_{\text{per}}^0 - V| \geq c > 0$ uniformly in $z \in \mathcal{C}$, we can mimic the proof of Lemma 1 and find that

$$\sup_{z \in \mathcal{C}} \|(1 - \Delta)(z - H_{\text{per}}^0 - V)^{-1}\| < \infty. \quad (4.40)$$

FIG. 4.2 – Curve enclosing the Z first bands.

We then use Cauchy's formula to write

$$\gamma_{\text{per}}^0 = \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} dz,$$

and iterate the resolvent formula

$$(z - H_{\text{per}}^0 - V)^{-1} = (z - H_{\text{per}}^0)^{-1} + (z - H_{\text{per}}^0 - V)^{-1}V(z - H_{\text{per}}^0)^{-1}$$

to rewrite Q as

$$\begin{aligned} Q &= \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0 - V)^{-1} - (z - H_{\text{per}}^0)^{-1} dz, \\ &= \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0 - V)^{-1}V(z - H_{\text{per}}^0)^{-1} dz, \\ &= \frac{1}{2i\pi} \int_{\mathcal{C}} \left((z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} + (z - H_{\text{per}}^0 - V)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} \right) dz, \\ &= \frac{1}{2i\pi} \int_{\mathcal{C}} \left((z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} + (z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} \right. \\ &\quad \left. + (z - H_{\text{per}}^0 - V)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} \right) dz, \\ &= Q_1 + Q_2 + Q_3, \end{aligned}$$

with

$$\begin{aligned} Q_1 &= \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} dz, \\ Q_2 &= \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} dz, \\ Q_3 &= \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0 - V)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1}V(z - H_{\text{per}}^0)^{-1} dz. \end{aligned}$$

Let us denote $B(z) = (z - H_{\text{per}}^0)^{-1}(1 - \Delta)$. The family of operators $(B(z))_{z \in \mathcal{C}}$ is uniformly bounded (this result is stated in Lemma 10). Consequently

$$\|V(z - H_{\text{per}}^0)^{-1}\|_{\mathfrak{S}_6} \leq \|V(1 - \Delta)^{-1}\|_{\mathfrak{S}_6} \|B(z)^*\| \leq C \|V\|_{L^6}$$

where we have used the Kato-Seiler-Simon inequality

$$\|V(x)f(-i\nabla)\|_{\mathfrak{S}_6} \leq C_6 \|f\|_{L^6(\mathbb{R}^3)} \|V\|_{L^6(\mathbb{R}^3)},$$

with $f(p) = \frac{1}{|p|^2+1}$. Together with (4.40), this result implies that $(1 - \Delta)Q_3 \in \mathfrak{S}_2$.

We now prove that $(1 - \Delta)Q_1 \in \mathfrak{S}_2$. Let us first show that

$$\begin{aligned} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz \\ = \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} (\gamma_{\text{per}}^0)^\perp V (\gamma_{\text{per}}^0)^\perp (z - H_{\text{per}}^0)^{-1} dz = 0. \end{aligned}$$

We have

$$\begin{aligned} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0{}^\perp V \gamma_{\text{per}}^0{}^\perp (z - H_{\text{per}}^0)^{-1} dz \\ = \frac{1}{(2\pi)^6} \sum_{k, k'=Z+1}^{+\infty} \int_{\Gamma^*} \int_{\Gamma^*} \int_{\mathcal{C}} \frac{1}{z - \lambda_k(\xi)} \frac{1}{z - \lambda_{k'}(\xi')} dz |e_{k'}(\xi', \cdot)\rangle \langle e_{k'}(\xi', \cdot)| V |e_k(\xi, \cdot)\rangle \langle e_k(\xi, \cdot)| d\xi' d\xi, \end{aligned}$$

and $\frac{1}{z - \lambda_k(\xi)} \frac{1}{z - \lambda_{k'}(\xi')}$ can be rewritten as

$$\frac{1}{\lambda_{k'}(\xi') - \lambda_k(\xi)} \left(\frac{1}{z - \lambda_{k'}(\xi')} - \frac{1}{z - \lambda_k(\xi)} \right),$$

hence the result, thanks to Cauchy's formula :

$$\frac{1}{2i\pi} \int_{\mathcal{C}} \frac{dz}{z - \lambda} = 1$$

if λ is in the part of the spectrum enclosed by \mathcal{C} . In this case, $\lambda_k(\xi)$ and $\lambda_{k'}(\xi')$ are not in this part of the spectrum as $k, k' \geq Z + 1$ since only the first Z bands are enclosed in \mathcal{C} . Therefore

$$\int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0{}^\perp V \gamma_{\text{per}}^0{}^\perp (z - H_{\text{per}}^0)^{-1} dz = 0.$$

Similarly,

$$\begin{aligned} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz \\ = \frac{1}{(2\pi)^6} \sum_{k, k'=1}^Z \int_{\Gamma^*} \int_{\Gamma^*} \int_{\mathcal{C}} \frac{1}{z - \lambda_k(\xi)} \frac{1}{z - \lambda_{k'}(\xi')} dz |e_{k'}(\xi', \cdot)\rangle \langle e_{k'}(\xi', \cdot)| V |e_k(\xi, \cdot)\rangle \langle e_k(\xi, \cdot)| d\xi' d\xi, \end{aligned}$$

both $\lambda_k(\xi)$ and $\lambda_{k'}(\xi')$ being now in the part of the spectrum enclosed by \mathcal{C} . Hence,

$$\int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz = 0.$$

Then we have

$$\begin{aligned} (1 - \Delta) \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V (\gamma_{\text{per}}^0)^\perp (z - H_{\text{per}}^0)^{-1} dz \\ = \int_{\mathcal{C}} B(z)^* [\gamma_{\text{per}}^0, V] (\gamma_{\text{per}}^0)^\perp (z - H_{\text{per}}^0)^{-1} dz \end{aligned}$$

and

$$\begin{aligned} (1 - \Delta) \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} (\gamma_{\text{per}}^0)^\perp V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz \\ = - \int_{\mathcal{C}} B(z)^* (\gamma_{\text{per}}^0)^\perp [\gamma_{\text{per}}^0, V] (z - H_{\text{per}}^0)^{-1} dz, \end{aligned}$$

which belong to \mathfrak{S}_2 by Lemmas 10 and 12 ($B(z)^* \in \mathfrak{S}_\infty$ and $[\gamma_{\text{per}}^0, V] \in \mathfrak{S}_2$).

Let us now prove that $(1 - \Delta)Q_2$ is also in \mathfrak{S}_2 .

Let us decompose Q_2 into two terms which we will then treat separately :

$$\begin{aligned} Q_2 &= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\ &\quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} (\gamma_{\text{per}}^0)^\perp V (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\ &:= Q_2^0 + Q_2^{0,\perp}. \end{aligned}$$

First,

$$\begin{aligned} Q_2^0 &= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 {}^2 V (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz, \\ &= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 [\gamma_{\text{per}}^0, V] (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\ &\quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz, \\ &= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 [\gamma_{\text{per}}^0, V] (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\ &\quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 {}^2 V (z - H_{\text{per}}^0)^{-1} dz, \\ &= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 [\gamma_{\text{per}}^0, V] (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\ &\quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 [\gamma_{\text{per}}^0, V] (z - H_{\text{per}}^0)^{-1} dz \\ &\quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz. \end{aligned}$$

On the one hand, when multiplying by $(1 - \Delta)$, the first two terms are obviously in \mathfrak{S}_2 since $[\gamma_{\text{per}}^0, V] \in \mathfrak{S}_2$ and the other terms are bounded (same justification as for

Q_1). On the other hand, the third term of Q_2^0 is equal to zero :

$$\begin{aligned}
& \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^0 V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz \\
&= \frac{1}{(2\pi)^9} \sum_{n,m,l=1}^Z \int_{\Gamma^*} \int_{\Gamma^*} \int_{\Gamma^*} \int_{\mathcal{C}} \frac{1}{z - \lambda_n(\xi)} \frac{1}{z - \lambda_m(\xi')} \frac{1}{z - \lambda_l(\xi'')} dz \\
& \quad |e_n(\xi, \cdot)\rangle \langle e_n(\xi, \cdot)| V |e_m(\xi', \cdot)\rangle \langle e_m(\xi', \cdot)| V |e_l(\xi'', \cdot)\rangle \langle e_l(\xi'', \cdot)| d\xi d\xi' d\xi'', \\
&= \frac{1}{(2\pi)^9} \sum_{n,m,l=1}^Z \int_{\Gamma^*} \int_{\Gamma^*} \int_{\Gamma^*} \int_{\mathcal{C}} \left(\frac{\alpha}{z - \lambda_n(\xi)} + \frac{\beta}{z - \lambda_m(\xi')} + \frac{\gamma}{z - \lambda_l(\xi'')} \right) dz \\
& \quad |e_n(\xi, \cdot)\rangle \langle e_n(\xi, \cdot)| V |e_m(\xi', \cdot)\rangle \langle e_m(\xi', \cdot)| V |e_l(\xi'', \cdot)\rangle \langle e_l(\xi'', \cdot)| d\xi d\xi' d\xi'',
\end{aligned}$$

with

$$\alpha = \frac{1}{(\lambda_n - \lambda_m)(\lambda_n - \lambda_l)}, \beta = \frac{1}{(\lambda_m - \lambda_l)(\lambda_m - \lambda_n)}, \gamma = \frac{1}{(\lambda_l - \lambda_n)(\lambda_l - \lambda_m)}.$$

As $n, m, l \leq Z$, we have $\int_{\mathcal{C}} \left(\frac{\alpha}{z - \lambda_n(\xi)} + \frac{\beta}{z - \lambda_m(\xi')} + \frac{\gamma}{z - \lambda_l(\xi'')} \right) dz = \alpha + \beta + \gamma = 0$.

We now deal with $Q_2^{0,\perp}$

$$\begin{aligned}
Q_2^{0,\perp} &= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V (\gamma_{\text{per}}^0 + \gamma_{\text{per}}^{0,\perp}) (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz, \\
&= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} [V, \gamma_{\text{per}}^0] (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\
& \quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz, \\
&= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} [V, \gamma_{\text{per}}^0] (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\
& \quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V (\gamma_{\text{per}}^0 + \gamma_{\text{per}}^{0,\perp}) (z - H_{\text{per}}^0)^{-1} dz, \\
&= \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} [V, \gamma_{\text{per}}^0] (z - H_{\text{per}}^0)^{-1} V (z - H_{\text{per}}^0)^{-1} dz \\
& \quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} [V, \gamma_{\text{per}}^0] (z - H_{\text{per}}^0)^{-1} dz \\
& \quad + \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} \gamma_{\text{per}}^{0,\perp} V \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} dz,
\end{aligned}$$

and using the same arguments as for Q_2^0 , we obtain that $(1 - \Delta)Q_2^{0,\perp} \in \mathfrak{S}_2$. \square

Remark 3. Notice that it is natural to wonder whether $Q \in \mathfrak{S}_1$, which would in particular imply that $\rho_Q \in L^1(\mathbb{R}^3)$. This is known to be false for the Bogoliubov-Dirac-Fock model studied in [22–25]. We do not answer this question for our model here.

4.3.3 Existence of minimizers under a charge constraint

In the previous section, we stated the existence of minimizers for any chemical potential in the gap of the periodic operator H_{per}^0 , but of course the total charge $\text{Tr}_0(\bar{Q})$ of the obtained solution was unknown *a priori*. Here we tackle the more subtle problem of minimizing the energy while *imposing a charge constraint*. Mathematically this is more difficult because although the energy $\mathcal{E}^\nu(Q)$ is convex on \mathcal{K} and weakly lower semi-continuous (wlsc) for the weak-* topology of \mathcal{Q} (as was shown in the proof of Theorem 2), the γ_{per}^0 -trace functional $Q \in \mathcal{K} \mapsto \text{Tr}_0(Q)$ is continuous but *not* wlsc for the weak-* topology of \mathcal{Q} : in principle it is possible that a (positive or negative) part of the charge of a minimizing sequence for the charge-constrained minimization problem escapes to infinity, leaving at the limit a state of a different (lower or higher) charge. In fact, we can prove that a minimizer exists under a charge constraint, if and only if some binding conditions hold, the role of which being to prevent the lack of compactness.

As explained above, imposing $\text{Tr}_0(Q) = q$ should intuitively lead (for a sufficiently weak defect density ν) to a system of q electrons coupled to a polarized Fermi sea. Notice that we do *not* impose that $q = \int_{\mathbb{R}^3} \nu$, i.e. our model allows *a priori* to treat defects with non-zero total charge.

As usual in reduced Hartree-Fock theories, we consider the case of a real charge constraint $q \in \mathbb{R}$:

$$\boxed{E^\nu(q) := \inf\{\mathcal{E}^\nu(Q), Q \in \mathcal{K}, \text{Tr}_0(Q) = q\}}. \quad (4.41)$$

When no defect is present, $E^0(q)$ can be computed explicitly :

Proposition 3 (Defect-free charge-constrained energy). *Let $Z \in \mathbb{N} \setminus \{0\}$ and assume that **(A1)** holds. Then one has*

$$E^0(q) = \begin{cases} \Sigma_{Z+1}^- q & \text{when } q \geq 0 \\ \Sigma_Z^+ q & \text{when } q \leq 0. \end{cases}$$

The minimization problem (4.41) with $\nu \equiv 0$ has no solution except when $q = 0$.

Proof. Clearly

$$E^0(q) \geq \inf\{\text{Tr}_0(H_{\text{per}}^0 Q), Q \in \mathcal{K}, \text{Tr}_0(Q) = q\} := \tilde{E}^0(q) \quad (4.42)$$

since $D(\rho_Q, \rho_Q) \geq 0$. We begin by showing that

$$\tilde{E}^0(q) = \begin{cases} \Sigma_{Z+1}^- q & \text{when } q \geq 0 \\ \Sigma_Z^+ q & \text{when } q \leq 0, \end{cases}$$

Let us take first the case $q \geq 0$. Then,

$$\tilde{E}^0(q) - \Sigma_{Z+1}^- q = \inf\{\text{Tr}_0((H_{\text{per}}^0 - \Sigma_{Z+1}^-)Q), Q \in \mathcal{K}, \text{Tr}_0(Q) = q\},$$

and for every $Q \in \mathcal{K}$

$$\begin{aligned} \mathrm{Tr}_0((H_{\mathrm{per}}^0 - \Sigma_{Z+1}^-)Q) &= \mathrm{Tr}((H_{\mathrm{per}}^0 - \Sigma_{Z+1}^-)(Q^{++} + Q^{--})), \\ &= \mathrm{Tr}(|H_{\mathrm{per}}^0 - \Sigma_{Z+1}^-|(Q^{++} - Q^{--})), \\ &\geq 0, \end{aligned}$$

since $\gamma_{\mathrm{per}}^0(H_{\mathrm{per}}^0 - \mu)\gamma_{\mathrm{per}}^0 = -\gamma_{\mathrm{per}}^0|H_{\mathrm{per}}^0 - \mu|\gamma_{\mathrm{per}}^0$ and $(\gamma_{\mathrm{per}}^0)^\perp(H_{\mathrm{per}}^0 - \mu)(\gamma_{\mathrm{per}}^0)^\perp = (\gamma_{\mathrm{per}}^0)^\perp|H_{\mathrm{per}}^0 - \mu|(\gamma_{\mathrm{per}}^0)^\perp$ by definition, for every μ in the gap $[\Sigma_Z^+, \Sigma_{Z+1}^-]$.

It remains to show that $\tilde{E}^0(q) \leq \Sigma_{Z+1}^-q$. Let $\epsilon > 0$. As the spectrum of H_{per}^0 is only essential spectrum, $\chi_{[\Sigma_{Z+1}^-, \Sigma_{Z+1}^- + \epsilon]}(H_{\mathrm{per}}^0)$ is an infinite rank projector, and we can take an orthonormal basis of $N = [q] + 1$ functions φ_i in $L^2(\mathbb{R}^3)$ such that $\varphi_i \in \mathrm{Ran}(\chi_{[\Sigma_{Z+1}^-, \Sigma_{Z+1}^- + \epsilon]}(H_{\mathrm{per}}^0))$. We then define $Q = \sum_{i=1}^{N-1} |\varphi_i\rangle\langle\varphi_i| + (q - [q])|\varphi_N\rangle\langle\varphi_N|$. Q is finite rank, therefore trace-class and $\mathrm{Tr}(Q) = q$ so $Q \in \mathcal{K}$ and satisfies the constraint. Then,

$$\begin{aligned} \mathrm{Tr}_0(H_{\mathrm{per}}^0 Q) = \mathrm{Tr}(H_{\mathrm{per}}^0 Q) &= \sum_{i=1}^{N-1} \langle H_{\mathrm{per}}^0 \varphi_i, \varphi_i \rangle + (q - [q]) \langle H_{\mathrm{per}}^0 \varphi_N, \varphi_N \rangle, \\ &\leq (\Sigma_{Z+1}^- + \epsilon)q, \end{aligned}$$

since by hypothesis $\Sigma_{Z+1}^- \leq \langle H_{\mathrm{per}}^0 \varphi_i, \varphi_i \rangle \leq \Sigma_{Z+1}^- + \epsilon$. So we have $\Sigma_{Z+1}^-q \leq \tilde{E}^0(q) \leq (\Sigma_{Z+1}^- + \epsilon)q$ for every $\epsilon > 0$, thus $\tilde{E}^0(q) = \Sigma_{Z+1}^-q$ when $q \geq 0$.

Similarly, for $q \leq 0$, we first have that

$$\begin{aligned} \tilde{E}^0(q) - \Sigma_Z^+q &= \inf\{\mathrm{Tr}_0((H_{\mathrm{per}}^0 - \Sigma_Z^+)Q), Q \in \mathcal{K}, \mathrm{Tr}_0(Q) = q\} \\ &= \inf\{-\mathrm{Tr}_0(|H_{\mathrm{per}}^0 - \Sigma_Z^+|(Q^{++} - Q^{--})), Q \in \mathcal{K}, \mathrm{Tr}_0(Q) = q\} \leq 0. \end{aligned}$$

Then let $N = -[q]$, $(\varphi_i)_{i=1}^N$ be in $\mathrm{Ran}(\chi_{[\Sigma_Z^+ - \epsilon, \Sigma_Z^+]}(H_{\mathrm{per}}^0))$ and $Q = -\sum_{i=1}^{N-1} |\varphi_i\rangle\langle\varphi_i| - ([q] + 1 - q)|\varphi_N\rangle\langle\varphi_N|$. We have $\mathrm{Tr}(Q) = q$ and $\mathrm{Tr}_0(H_{\mathrm{per}}^0 Q) \geq (\Sigma_Z^+ - \epsilon)q$ so $\tilde{E}^0(q) = \Sigma_Z^+q$.

Thus it remains to prove that $E^0(q) \leq \tilde{E}^0(q)$ which we do by a kind of scaling argument. Let us deal with the case $q \geq 0$, the other case being similar. We can assume that $\Sigma_{Z+1}^- = \min_{\xi \in \Gamma^*} \lambda_{Z+1}(\xi) = \lambda_{Z+1}(\xi_0)$ since each $\lambda_n(\xi)$ is known to be continuous on Γ^* . For simplicity, we also assume that ξ_0 is in the interior of Γ^* (the proof can be easily adapted if this is not the case). Let us denote by $u_{Z+1}(\xi, \cdot) \in L_\xi^2$ an eigenvector associated with the eigenvalue $\lambda_{Z+1}(\xi)$ for any $\xi \in \Gamma^*$. Since $H_{\mathrm{per}}^0 u_{Z+1}(\xi, x) = \lambda_{Z+1}(\xi)u_{Z+1}(\xi, x)$ for any $\xi \in \Gamma^*$, it is clear that

$$\sup_{\xi \in \mathbb{R}^3} \|\Delta u_{Z+1}(\xi, \cdot)\|_{L_\xi^2(\Gamma)} < \infty,$$

i.e. $u_{Z+1} \in L^\infty(\mathbb{R}^3, H_{\mathrm{loc}}^2(\mathbb{R}^3))$ and $u_{Z+1} \in L^\infty(\mathbb{R}^3 \times \mathbb{R}^3)$ by the \mathbb{Z}^3 -periodicity (resp. the $2\pi\mathbb{Z}^3$ -periodicity) of $e^{-i\xi \cdot x}u_{Z+1}(\xi, x)$ with respect to x (resp. to ξ).

Consider now a fixed space V of dimension $[q] + 1$, consisting of C_0^∞ functions $f : \mathbb{R}^3 \rightarrow \mathbb{C}$ with support in the unit ball $B(0, 1)$ of \mathbb{R}^3 . For any λ large enough so

that the ball $B(\xi_0, \lambda^{-1})$ is contained in Γ^* , we introduce the following subspace of $L^2(\mathbb{R}^3)$:

$$W_\lambda := \left\{ g_\lambda := \lambda^{3/2} \int_{\mathbb{R}^3} f(\lambda(\xi - \xi_0)) u_{Z+1}(\xi, \cdot) d\xi, \quad f \in V \right\}$$

which has the same dimension as V . Noting that for any $g_\lambda \in W_\lambda$ arising from some $f \in V$

$$\begin{aligned} |g_\lambda(x)| &\leq \lambda^{3/2} \int_{\mathbb{R}^3} |f(\lambda(\xi - \xi_0)) u_{Z+1}(\xi, x)| d\xi, \\ &\leq \lambda^{-3/2} \int_{\mathbb{R}^3} |f(\xi') u_{Z+1}(\xi_0 + \lambda^{-1} \xi', x)| d\xi', \\ &\leq \lambda^{-3/2} \|u_{Z+1}\|_{L^\infty(\mathbb{R}^3 \times \mathbb{R}^3)} \int_{B(0,1)} |f(\xi)| d\xi, \\ &\leq C \lambda^{-3/2} \|u_{Z+1}\|_{L^\infty(\mathbb{R}^3 \times \mathbb{R}^3)} \|f\|_{L^2(\mathbb{R}^3)}, \\ &\rightarrow 0 \quad \text{when } \lambda \rightarrow \infty, \end{aligned}$$

we deduce thanks to the interpolation inequality that if g_λ is moreover bounded in $L^2(\mathbb{R}^3)$ (we take for instance $\|g\|_{L^2(\mathbb{R}^3)} = 1$),

$$\forall 2 < p \leq \infty, \quad \lim_{\lambda \rightarrow \infty} \sup_{\substack{g \in W_\lambda \\ \|g\|_{L^2(\mathbb{R}^3)} = 1}} \|g\|_{L^p(\mathbb{R}^3)} = 0. \quad (4.43)$$

By construction one also has for any fixed $f \in V$ with associated $g_\lambda \in W_\lambda$

$$\begin{aligned} \langle g_\lambda, (H_{\text{per}}^0 - \Sigma_{Z+1}^-) g_\lambda \rangle_{L_\xi^2} &= \lambda^3 \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} |f(\lambda(\xi - \xi_0))|^2 (\lambda_{Z+1}(\xi) - \Sigma_{Z+1}^-) |u_{Z+1}(\xi, x)|^2 d\xi dx, \\ &= \lambda^3 \int_{\mathbb{R}^3} |f(\lambda(\xi - \xi_0))|^2 (\lambda_{Z+1}(\xi) - \Sigma_{Z+1}^-) d\xi, \\ &= \int_{\mathbb{R}^3} |f(\xi')|^2 (\lambda_{Z+1}(\xi_0 + \lambda^{-1} \xi') - \Sigma_{Z+1}^-) d\xi', \\ &= \int_{B(0,1)} |f(\xi')|^2 (\lambda_{Z+1}(\xi_0 + \lambda^{-1} \xi') - \Sigma_{Z+1}^-) d\xi', \\ &\rightarrow_{\lambda \rightarrow \infty} 0, \end{aligned}$$

since f is bounded and $\lambda_{Z+1}(\xi_0 + \lambda^{-1} \xi') \rightarrow \lambda_{Z+1}(\xi_0) = \Sigma_{Z+1}^-$ when $\lambda \rightarrow \infty$.

Take now an orthonormal basis $(\varphi_1^\lambda, \dots, \varphi_{[q]+1}^\lambda)$ of W_λ and introduce the operator

$$Q^\lambda := \sum_{n=1}^{[q]} |\varphi_n^\lambda\rangle \langle \varphi_n^\lambda| + (q - [q]) |\varphi_{[q]+1}^\lambda\rangle \langle \varphi_{[q]+1}^\lambda|.$$

By construction $\gamma_{\text{per}}^0 \varphi_n^\lambda = 0$ for any $n = 1, \dots, [q]$ and, as before, Q is a finite rank operator and $\text{Tr}_0(Q^\lambda) = \text{Tr}(Q^\lambda) = q$, thus $Q^\lambda \in \mathcal{K}$ satisfies the charge constraint.

Then

$$\mathrm{Tr}_0(H_{\mathrm{per}}^0 Q^\lambda) = \mathrm{Tr}(H_{\mathrm{per}}^0 Q^\lambda) = \sum_{n=1}^{[q]} \langle \varphi_n^\lambda, H_{\mathrm{per}}^0 \varphi_n^\lambda \rangle + (q - [q]) \langle \varphi_{[q]+1}^\lambda, H_{\mathrm{per}}^0 \varphi_{[q]+1}^\lambda \rangle \rightarrow \Sigma_{Z+1}^- q$$

as $\lambda \rightarrow \infty$ since for every $n, 1 \leq n \leq [q] + 1$, $\langle \varphi_n^\lambda, (H_{\mathrm{per}}^0 - \Sigma_{Z+1}^-) \varphi_n^\lambda \rangle_{L_\xi^2} \rightarrow_{\lambda \rightarrow \infty} 0$. Besides, as $u_{Z+1}(\xi, \cdot)$ is in $H^1(\mathbb{R}^3)$, so are the (φ_n^λ) and therefore $(\rho_{Q^\lambda}^{1/2})_{\lambda \geq 1}$ is a bounded family in $H^1(\mathbb{R}^3)$ which converges to 0 in $L^p(\mathbb{R}^3)$ for any $p > 2$ thanks to (4.43). By the Hardy-Littlewood-Sobolev inequality, one has

$$D(\rho_{Q^\lambda}, \rho_{Q^\lambda}) \leq C \|\rho_{Q^\lambda}\|_{L^{6/5}}^2 \quad (4.44)$$

which implies $D(\rho_{Q^\lambda}, \rho_{Q^\lambda}) \rightarrow 0$ as $\lambda \rightarrow \infty$. Eventually $\mathcal{E}^0(Q^\lambda) \rightarrow \tilde{E}^0(q)$ and Proposition 3 is proved. \square

We now state the main result of this section, which is directly inspired from [24] :

Theorem 3 (Existence of minimizers under a charge constraint). *Let $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, $Z \in \mathbb{N} \setminus \{0\}$ and assume that **(A1)** holds. The following assertions are equivalent :*

(a) *Problem (4.41) admits a minimizer \bar{Q} ;*

(b) *Every minimizing sequence for (4.41) is precompact in \mathcal{Q} and converges towards a minimizer \bar{Q} of (4.41) ;*

(c) $\forall q' \in \mathbb{R} \setminus \{0\}$, $E^\nu(q) < E^\nu(q - q') + E^0(q')$.

Assume that the equivalent conditions (a), (b) and (c) above are fulfilled. In this case, the minimizer \bar{Q} is not necessarily unique, but all the minimizers share the same density $\bar{\rho} = \rho_{\bar{Q}}$. Besides, there exists $\epsilon_F \in [\Sigma_Z^+, \Sigma_{Z+1}^-]$ such that the obtained minimizer \bar{Q} is a global minimizer for $E_{\epsilon_F}^\nu$ defined in (4.31). It solves Equation (4.32) for some $0 \leq \delta \leq 1$ with $\mathrm{Ran}(\delta) \subseteq \ker(H_{\bar{Q}} - \epsilon_F)$. The operator δ is finite rank if $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$ and trace-class if $\epsilon_F \in \{\Sigma_Z^+, \Sigma_{Z+1}^-\}$.

Additionally the set of q 's in \mathbb{R} satisfying the above equivalent conditions is a non-empty closed interval $I \subseteq \mathbb{R}$. This is the largest interval on which $q \mapsto E^\nu(q)$ is strictly convex.

Remark 4. One has $I = \{\mathrm{Tr}_0(\bar{Q}), \bar{Q} \text{ min. of } E_{\epsilon_F}^\nu, \epsilon_F \in [\Sigma_Z^+, \Sigma_{Z+1}^-]\}$. Hence $I \neq \emptyset$ by Theorem 2.

Many of the above statements are very common in reduced Hartree-Fock theories and not all the details will be given in the proof (see, e.g. [46]). The difficult part is the proof that (b) is equivalent to (c), for which we use ideas from [24].

The proof that any minimizer solves (4.32) is the same as before and will be omitted.

Proof of Theorem 3.

Step 1 : Large HVZ-type inequalities

Let us start by the following result, which indeed shows that **(b)** \Rightarrow **(c)** :

Lemma 4 (Large HVZ-type inequalities). *Let $Z \in \mathbb{N} \setminus \{0\}$, $\nu \in \mathcal{C}$ and assume that **(A1)** holds. Then, for every $q, q' \in \mathbb{R}$, one has*

$$E^\nu(q) \leq E^\nu(q - q') + E^0(q').$$

If moreover there is a $q' \neq 0$ such that $E^\nu(q) = E^\nu(q - q') + E^0(q')$, then there is a minimizing sequence of $E^\nu(q)$ which is not precompact.

Proof. Here, thanks to Corollary 4, we can mimic the proof of [24, Prop. 6]. We first notice that $E^0(0) = 0$ so that there is nothing to prove in the case $q' = 0$. Let us now fix some $q' \neq 0$ and consider two states Q and $Q' \in \mathcal{K}$ which are almost minimizers of $E^\nu(q - q')$ and $E^0(q')$ respectively

$$\mathcal{E}^\nu(Q) \leq E^\nu(q - q') + \epsilon \quad \text{and} \quad \mathcal{E}^0(Q') \leq E^0(q') + \epsilon$$

for some $\epsilon > 0$. Corollary 4, together with Lemma 14 and Proposition 5, allows us to approximate any projector $Q = P - \gamma_{\text{per}}^0 \in \mathcal{K}$ by a finite rank projector $Q_k = P_k - \gamma_{\text{per}}^0$, also in \mathcal{K} so we can choose Q and Q' of the form

$$\begin{aligned} Q &= \lambda |\psi\rangle\langle\psi| + \sum_{n=1}^N |f_n\rangle\langle f_n| - \sum_{m=1}^M |g_m\rangle\langle g_m| + \sum_{i=1}^k \frac{\lambda_i^2}{1 + \lambda_i^2} (|v_i\rangle\langle v_i| - |u_i\rangle\langle u_i|) \\ &\quad + \sum_{i=1}^k \frac{\lambda_i}{1 + \lambda_i^2} (|u_i\rangle\langle v_i| - |v_i\rangle\langle u_i|) \end{aligned} \quad (4.45)$$

$$\begin{aligned} Q' &= \lambda' |\psi'\rangle\langle\psi'| + \sum_{n=1}^{N'} |f'_n\rangle\langle f'_n| - \sum_{m=1}^{M'} |g'_m\rangle\langle g'_m| + \sum_{i=1}^{k'} \frac{(\lambda'_i)^2}{1 + (\lambda'_i)^2} (|v'_i\rangle\langle v'_i| - |u'_i\rangle\langle u'_i|) \\ &\quad + \sum_{i=1}^{k'} \frac{\lambda'_i}{1 + (\lambda'_i)^2} (|u'_i\rangle\langle v'_i| - |v'_i\rangle\langle u'_i|) \end{aligned} \quad (4.46)$$

where $\lambda = q - q' - [q - q'] \in [0, 1)$, $\lambda' = q' - [q'] \in [0, 1)$, $N - M = [q - q']$ and $N' - M' = [q']$. In (4.45) and (4.46), $(f_n)_{n=1}^N \cup (v_i)_{i=1}^k$ and $(f'_n)_{n=1}^{N'} \cup (v'_i)_{i=1}^{k'}$ form two orthonormal sets of $\mathfrak{H}_+ = (1 - \gamma_{\text{per}}^0)L^2(\mathbb{R}^3)$, $(g_m)_{m=1}^M \cup (u_i)_{i=1}^k$ and $(g'_m)_{m=1}^{M'} \cup (u'_i)_{i=1}^{k'}$ form two orthonormal sets of $\mathfrak{H}_- = \gamma_{\text{per}}^0 L^2(\mathbb{R}^3)$.

The free minimization problem $E^0(q')$ being translation invariant, we have $\mathcal{E}^0(Q'_t) = \mathcal{E}^0(Q')$ for any $t \in \mathbb{R}$, where $Q'_t(x, y) := Q'(x - te, y - te)$, e being a fixed unitary vector in \mathbb{R}^3 . Easily,

$$\begin{aligned} Q'_t &= \lambda' |\psi^t\rangle\langle\psi^t| + \sum_{n=1}^{N'} |f_n^t\rangle\langle f_n^t| - \sum_{m=1}^{M'} |g_m^t\rangle\langle g_m^t| + \sum_{i=1}^{k'} \frac{(\lambda'_i)^2}{1 + (\lambda'_i)^2} (|v_i^t\rangle\langle v_i^t| - |u_i^t\rangle\langle u_i^t|) \\ &\quad + \sum_{i=1}^{k'} \frac{\lambda'_i}{1 + (\lambda'_i)^2} (|u_i^t\rangle\langle v_i^t| - |v_i^t\rangle\langle u_i^t|) \end{aligned}$$

where we have used the notation $\psi^t(x) = \psi'(x - te)$ and similarly for the other functions. Notice also that $(f_n^t)_{n=1}^{N'} \cup (v_i^t)_{i=1}^{k'} \subset \mathfrak{H}_+$ and $(g_m^t)_{m=1}^{M'} \cup (u_i^t)_{i=1}^{k'} \subset \mathfrak{H}_-$, since these two spaces are translation-invariant.

For every $f, g \in L^2(\mathbb{R}^3)$, we have $\lim_{t \rightarrow \infty} \langle f^t, g \rangle = 0$. So we can find for t large enough by the Gram-Schmidt orthonormalization procedure two orthonormal systems

$$(\tilde{f}_n^t)_{n=1}^{N'} \cup (\tilde{v}_i^t)_{i=1}^{k'} \subset \left(\text{span}_{\substack{1 \leq n \leq N' \\ 1 \leq i \leq k'}} \{f_n, v_i\} \right)^\perp \cap \mathfrak{H}_+$$

and

$$(\tilde{g}_m^t)_{m=1}^{M'} \cup (\tilde{u}_i^t)_{i=1}^{k'} \subset \left(\text{span}_{\substack{1 \leq m \leq M' \\ 1 \leq i \leq k'}} \{g_m, u_i\} \right)^\perp \cap \mathfrak{H}_-$$

which are such that

$$\lim_{t \rightarrow \infty} \left\| \tilde{f}_n^t - f_n^t \right\|_{L^2(\mathbb{R}^3)} = \lim_{t \rightarrow \infty} \left\| \tilde{u}_i^t - u_i^t \right\|_{L^2(\mathbb{R}^3)} = \lim_{t \rightarrow \infty} \left\| \tilde{g}_m^t - g_m^t \right\|_{L^2(\mathbb{R}^3)} = \lim_{t \rightarrow \infty} \left\| \tilde{v}_i^t - v_i^t \right\|_{L^2(\mathbb{R}^3)} = 0$$

for any $1 \leq n \leq N'$, $1 \leq m \leq M'$, $1 \leq i \leq k'$. We define $\tilde{\psi}^t$ and introduce

$$\begin{aligned} \tilde{Q}'_t &= \lambda' |\tilde{\psi}^t\rangle \langle \tilde{\psi}^t| + \sum_{n=1}^{N'} |\tilde{f}_n^t\rangle \langle \tilde{f}_n^t| - \sum_{m=1}^{M'} |\tilde{g}_m^t\rangle \langle \tilde{g}_m^t| + \sum_{i=1}^{k'} \frac{(\lambda'_i)^2}{1 + (\lambda'_i)^2} (|\tilde{v}_i^t\rangle \langle \tilde{v}_i^t| - |\tilde{u}_i^t\rangle \langle \tilde{u}_i^t|) \\ &\quad + \sum_{i=1}^{k'} \frac{\lambda'_i}{1 + (\lambda'_i)^2} (|\tilde{u}_i^t\rangle \langle \tilde{v}_i^t| - |\tilde{v}_i^t\rangle \langle \tilde{u}_i^t|) \end{aligned}$$

which is such that $\lim_{t \rightarrow \infty} \left\| \tilde{Q}'_t - Q'_t \right\|_{\mathfrak{S}_1(L^2(\mathbb{R}^3))} = 0$. We have by construction $Q + \tilde{Q}'_t \in \mathcal{K}$ and $\text{Tr}_0(Q + \tilde{Q}'_t) = q$, thus $Q + \tilde{Q}'_t$ satisfies the constraint. Now we shall compute the energy of $Q + \tilde{Q}'_t$

$$\begin{aligned} \mathcal{E}^\nu(Q + \tilde{Q}'_t) &= \text{Tr}_0(H_{\text{per}}^0 Q) + \text{Tr}_0(H_{\text{per}}^0 \tilde{Q}'_t) - D(\rho_Q, \nu) - D(\rho_{\tilde{Q}'_t}, \nu) + \frac{1}{2} D(\rho_{Q + \tilde{Q}'_t}, \rho_{Q + \tilde{Q}'_t}), \\ &= \mathcal{E}^\nu(Q) + \mathcal{E}^\nu(\tilde{Q}'_t) + D(\rho_Q, \rho_{\tilde{Q}'_t}), \\ &= \mathcal{E}^\nu(Q) + \mathcal{E}^0(\tilde{Q}'_t) + D(\rho_Q - \nu, \rho_{\tilde{Q}'_t}), \end{aligned}$$

and since Q and \tilde{Q}'_t are finite rank operators with their kernels in $L^2(\mathbb{R}^3)$ and ν is in $L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$,

$$\begin{aligned} D(\rho_Q - \nu, \rho_{\tilde{Q}'_t}) &= \int_{\mathbb{R}^3} 4\pi \frac{\overline{(\hat{\rho}_Q(k) - \hat{\nu}(k))} \hat{\rho}_{\tilde{Q}'_t}(k)}{|k|^2} dk \\ &= 4\pi \int_{\mathbb{R}^3} \frac{(\hat{\rho}_Q(k) - \hat{\nu}(k)) \hat{\rho}_{\tilde{Q}'_t}(k) e^{-ik \cdot te}}{|k|^2} dk \\ &= 2(2\pi)^{\frac{5}{2}} \hat{\varphi}(te), \end{aligned}$$

where $\varphi(k) = \frac{\overline{(\hat{\rho}_Q(k) - \hat{\nu}(k))} \hat{\rho}_{\tilde{Q}'_t}(k)}{|k|^2}$. $\varphi \in L^1(\mathbb{R}^3)$ since

$$\int_{\mathbb{R}^3} |\varphi(k)| \leq \frac{1}{4\pi} D(\rho_Q - \nu, \rho_Q - \nu)^{\frac{1}{2}} D(\rho_{\tilde{Q}'_t}, \rho_{\tilde{Q}'_t})^{\frac{1}{2}} \leq M,$$

with $0 < M < \infty$ as $\rho_Q, \nu, \rho_{\tilde{Q}'} \in \mathcal{C}$ by definition. Then, by the properties of Fourier transform for L^1 functions, $\hat{\varphi}(te) \rightarrow 0$ when $t \rightarrow \infty$, hence $D(\rho_Q - \nu, \rho_{\tilde{Q}'_t})$ converges to zero when $t \rightarrow \infty$.

So

$$\begin{aligned} E^\nu(q) &\leq \mathcal{E}^\nu(Q + \tilde{Q}'_t) = \mathcal{E}^\nu(Q) + \mathcal{E}^0(\tilde{Q}'_t) + o_{t \rightarrow \infty}(1) \\ &\leq E^\nu(q - q') + E^0(q') + 2\epsilon + o_{t \rightarrow \infty}(1), \end{aligned}$$

which, passing to the limit as $t \rightarrow \infty$ and $\epsilon \rightarrow 0^+$, ends the proof of this first step.

If $E^\nu(q) = E^\nu(q - q') + E^0(q')$ for some $q' \neq 0$, then one constructs a non-compact minimizing sequence by the same argument.

Step 2 : A necessary and sufficient condition for compactness

The following Proposition is the analogue of [24, Lemma 8] :

Proposition 4 (Conservation of charge implies compactness). *Let $Z \in \mathbb{N} \setminus \{0\}$, $\nu \in \mathcal{C}$, $q \in \mathbb{R}$ and assume that **(A1)** holds. Assume that $(Q_n)_{n \geq 1}$ is a minimizing sequence in \mathcal{K}_r for (4.31) such that $Q_n \rightharpoonup Q \in \mathcal{K}$ for the weak-* topology of \mathcal{Q} . Then $Q_n \rightarrow Q$ for the strong topology of \mathcal{Q} if and only if $\text{Tr}_0(Q) = q$.*

Proof of the Proposition. Let $(Q_n)_{n \geq 1} \subseteq \mathcal{K}_r$ be as stated and assume that $\text{Tr}_0(Q) = q$. We know from the proof of Theorem 2 that

$$\mathcal{E}^\nu(Q) \leq \lim_{n \rightarrow \infty} \mathcal{E}^\nu(Q_n) = E^\nu(q), \quad (4.47)$$

hence $Q \in \mathcal{K}$ is a minimizer of $E^\nu(q)$. Therefore Q satisfies the equation

$$Q = \chi_{(-\infty, \epsilon_F)}(H_Q) - \gamma_{\text{per}}^0 + \delta$$

for some $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$ and where δ is a finite rank operator satisfying $0 \leq \delta \leq 1$ and $\text{Ran}(\delta) \subseteq \ker(H_Q - \epsilon_F)$. In particular $Q \in \mathcal{K}_r$ by Proposition 2. We now introduce

$$P := \chi_{(-\infty, \epsilon_F)}(H_Q), \quad P' := \chi_{(\epsilon_F, \infty)}(H_Q) \quad \text{and} \quad \pi := \chi_{\{\epsilon_F\}}(H_Q).$$

Let us write

$$\mathcal{E}^\nu(Q_n) = \mathcal{E}^\nu(Q) + \text{Tr}_0(H_Q(Q_n - Q)) + \frac{1}{2}D(\rho_{Q_n} - \rho_Q, \rho_{Q_n} - \rho_Q).$$

Now using [22, Lemma 1], the fact that $P - \gamma_{\text{per}}^0 = Q - \delta \in \mathfrak{S}_2$ and the hypothesis $\text{Tr}_0(Q_n) = \text{Tr}_0(Q)$, we obtain

$$\begin{aligned} \text{Tr}_0(H_Q(Q_n - Q)) &= \text{Tr}_0((H_Q - \epsilon_F)(Q_n - Q)) = \text{Tr}_P((H_Q - \epsilon_F)(Q_n - Q)) \\ &= \text{Tr}(|H_Q - \epsilon_F|(P'(Q_n - Q)P' - P(Q_n - Q)P)), \end{aligned}$$

where we recall that by definition $\text{Tr}_P(A) = \text{Tr}(PAP + (1 - P)A(1 - P))$. Since $Q_n \in \mathcal{K}_r$, $-P - \delta \leq Q_n - Q \leq 1 - P - \delta$. Hence,

$$\begin{aligned} 0 &\leq Q_n - Q + P + \delta \leq 1, \\ (Q_n - Q + P + \delta)^2 &\leq Q_n - Q + P + \delta, \\ Q_n^2 - 2Q_nQ + 2Q_nP + 2Q_n\delta + Q^2 - 2QP - 2Q\delta + P + \delta^2 &\leq Q_n - Q + P + \delta, \\ (Q_n - Q)^2 &\leq Q_n - Q - 2(Q_n - Q)(P + \delta) + \delta - \delta^2, \end{aligned}$$

which implies $P(Q_n - Q)^2P \leq -P(Q_n - Q)P$ and $P'(Q_n - Q)^2P' \leq P'(Q_n - Q)P'$ since $P\delta = P'\delta = 0$ and $P^2 = P$ by construction. This yields

$$P(Q_n - Q)^2P + P'(Q_n - Q)^2P' \leq P'(Q_n - Q)P' - P(Q_n - Q)P \quad (4.48)$$

and in particular $P'(Q_n - Q)P' \geq 0$ and $P(Q_n - Q)P \leq 0$, i.e. $\text{Tr}_0(H_Q(Q_n - Q)) \geq 0$. Since we know that $\lim_{n \rightarrow \infty} \mathcal{E}^\nu(Q_n) = \mathcal{E}^\nu(Q)$, we have $\text{Tr}_0(H_Q(Q_n - Q)) + \frac{1}{2}D(\rho_{Q_n} - \rho_Q, \rho_{Q_n} - \rho_Q) \rightarrow 0$ when $n \rightarrow \infty$ and since both terms are nonnegative, we infer

$$\lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \epsilon_F|P'(Q_n - Q)P') = \lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \epsilon_F|P(Q_n - Q)P) = 0 \quad (4.49)$$

and from (4.48)

$$\lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \epsilon_F|P'(Q_n - Q)^2P') = \lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \epsilon_F|P(Q_n - Q)^2P) = 0.$$

On the one hand, we have that

$$P|H_Q - \epsilon_F|P \geq cP|H_Q - \kappa|P \quad \text{and} \quad P'|H_Q - \epsilon_F|P' \geq cP'|H_Q - \kappa|P'$$

for some small enough constant $c > 0$ and some $\kappa \notin \sigma(H_Q)$ close enough to ϵ_F . Indeed, there are two cases : if $\epsilon_F \notin \sigma(H_Q)$, this is obvious. Now let us assume that ϵ_F is an isolated eigenvalue of H_Q , then there exists α such that $(-\infty, \epsilon_F) \cap \sigma(H_Q) = (-\infty, \alpha]$. Let $\kappa = \frac{\alpha + \epsilon_F}{2}$, $\kappa \notin \sigma(H_Q)$. We show the inequality for P , the other case being similar. There exists $c > 0$ such that $\forall \lambda \in (-\infty, \alpha]$, $|\lambda - \epsilon_F| \geq c|\lambda - \kappa|$. Then,

$$P|H_Q - \epsilon_F|P = \int_{-\infty}^{\alpha} |\lambda - \epsilon_F| dP_\lambda \geq c \int_{-\infty}^{\alpha} |\lambda - \kappa| dP_\lambda = cP|H_Q - \kappa|P.$$

On the other hand, the fact that π is a ‘‘smooth’’ finite rank operator implies that $|H_Q - \kappa|\pi$ is finite rank and therefore Hilbert-Schmidt. This, added to the weak convergence of (Q_n) in \mathfrak{S}_2 , leads to

$$\lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \kappa|\pi(Q_n - Q)^2\pi) = \lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \kappa|\pi(Q_n - Q)\pi) = 0. \quad (4.50)$$

It is then clear that this yields

$$\lim_{n \rightarrow \infty} \text{Tr}(|H_Q - \kappa|(Q_n - Q)^2) = 0. \quad (4.51)$$

As we have chosen $\kappa \notin \sigma(H_Q)$, we can mimic the proof of Lemma 1 and obtain that

$$c_1(1 - \Delta) \leq |H_Q - \kappa| \leq c_2(1 - \Delta). \quad (4.52)$$

Hence (4.51) shows that $(1 - \Delta)^{1/2}(Q_n - Q) \rightarrow 0$ in \mathfrak{S}_2 .

Writing now

$$\begin{aligned}
(Q_n - Q)^{-} &= (\gamma_{\text{per}}^0 - P + P)(Q_n - Q)(\gamma_{\text{per}}^0 - P + P), \\
&= (\gamma_{\text{per}}^0 - P)(Q_n - Q)(\gamma_{\text{per}}^0 - P) + (\gamma_{\text{per}}^0 - P)(Q_n - Q)P \\
&\quad + P(Q_n - Q)(\gamma_{\text{per}}^0 - P) + P(Q_n - Q)P, \\
&= P(Q_n - Q)P + (\gamma_{\text{per}}^0 - P)(Q_n - Q)\gamma_{\text{per}}^0 \\
&\quad - (P - \gamma_{\text{per}}^0)(Q_n - Q)(P - \gamma_{\text{per}}^0) + \gamma_{\text{per}}^0(Q_n - Q)(\gamma_{\text{per}}^0 - P)
\end{aligned} \tag{4.53}$$

and using the results above, we are going to prove that $|H_Q - \kappa|^{1/2}(Q_n - Q)^{-}|H_Q - \kappa|^{1/2} \rightarrow 0$ and $(1 - \Delta)^{1/2}(Q_n - Q)^{-}(1 - \Delta)^{1/2} \rightarrow 0$ in \mathfrak{S}_1 .

We already have that $\| |H_Q - \kappa|^{1/2}P(Q_n - Q)P|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_1} \rightarrow_{n \rightarrow \infty} 0$. by (4.51).

Then, using that $\|AB_nA\|_{\mathfrak{S}_1} \leq \|A\|_{\mathfrak{S}_2} \|B_n\|_{\mathfrak{S}_2} \|A\|_{\mathfrak{S}_\infty} \leq \|A\|_{\mathfrak{S}_2} \|B_n\|_{\mathfrak{S}_2} \|A\|_{\mathfrak{S}_2}$,

$$\begin{aligned}
&\| |H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P)(Q_n - Q)\gamma_{\text{per}}^0|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_1} \\
&\leq \| (|H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P)(Q_n - Q)(\gamma_{\text{per}}^0 - P)|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_1} \\
&\quad + \| (|H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P)(Q_n - Q)|H_Q - \kappa|^{1/2}P \|_{\mathfrak{S}_1}, \\
&\leq \| |H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P) \|_{\mathfrak{S}_2} \| (Q_n - Q) \|_{\mathfrak{S}_2} \| (\gamma_{\text{per}}^0 - P)|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_2} \\
&\quad + \| (|H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P) \|_{\mathfrak{S}_2} \| (Q_n - Q)|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_2} \| P \|,
\end{aligned}$$

$|H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P)$, $(\gamma_{\text{per}}^0 - P)|H_Q - \kappa|^{1/2}$ being bounded in \mathfrak{S}_2 and P being a bounded operator, $(Q_n - Q) \rightarrow 0$ in \mathfrak{S}_2 and $(Q_n - Q)|H_Q - \kappa|^{1/2} \rightarrow 0$ in \mathfrak{S}_2 , we infer

$$\| |H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P)(Q_n - Q)\gamma_{\text{per}}^0|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_1} \rightarrow_{n \rightarrow \infty} 0.$$

In the same way, $\| |H_Q - \kappa|^{1/2}\gamma_{\text{per}}^0(Q_n - Q)(\gamma_{\text{per}}^0 - P)|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_1} \rightarrow_{n \rightarrow \infty} 0$ and $\| |H_Q - \kappa|^{1/2}(\gamma_{\text{per}}^0 - P)(Q_n - Q)(\gamma_{\text{per}}^0 - P)|H_Q - \kappa|^{1/2} \|_{\mathfrak{S}_1} \rightarrow_{n \rightarrow \infty} 0$. The proof is the same for $(Q_n - Q)^{++}$.

$(1 - \Delta)^{1/2}(Q_n - Q)^{-}(1 - \Delta)^{1/2} \rightarrow 0$ in \mathfrak{S}_1 comes from $|H_Q - \kappa|^{1/2}(Q_n - Q)^{-}|H_Q - \kappa|^{1/2} \rightarrow 0$ and (4.52).

Step 3 : Proof that (c) \Rightarrow (b)

We argue by contradiction. Let $(Q_n)_{n \geq 1} \subseteq \mathcal{K}$ be a minimizing sequence for $E^\nu(q)$ which is not precompact for the topology of \mathcal{Q} ¹. By the density of \mathcal{K}_r in \mathcal{K} , we can further assume that each $Q_n \in \mathcal{K}_r$. The bound (4.28) on the energy tells us that $(Q_n)_{n \geq 1}$ is bounded in \mathcal{Q} . Then, up to extraction and by Proposition 4, we can assume that $Q_n \rightharpoonup Q \in \mathcal{K}$ where $\text{Tr}_0(Q) \neq q$, and that $\rho_{Q_n} \rightharpoonup \rho_Q$ weakly in \mathcal{C} . We write $\text{Tr}_0(Q) = q - q'$ with $q' \neq 0$. We now prove that

$$E^\nu(q) \geq E^\nu(q - q') + E^0(q') \tag{4.54}$$

¹A sequence $(Q_n)_{n \in \mathbb{N}}$ is precompact for the topology of \mathcal{Q} if it admits at least a subsequence which converges strongly to some limit in \mathcal{Q} .

which will contradict (c). To this end, we argue like in the proof of [24, Thm. 3] : consider a smooth radial function χ with support in $B(0, 1)$ such that $0 \leq \chi \leq 1$ and $\chi \equiv 1$ in $B(0, 1/2)$; define $\chi_R(x) := \chi(x/R)$. Then let be $\eta_R := \sqrt{1 - \chi_R^2}$. Let us introduce the following localization operators

$$Y_R := \gamma_{\text{per}}^0 \eta_R \gamma_{\text{per}}^0 + (\gamma_{\text{per}}^0)^\perp \eta_R (\gamma_{\text{per}}^0)^\perp, \quad X_R = \sqrt{1 - Y_R^2}.$$

Lemma 5. *We have for all $3 < p \leq \infty$,*

$$\|[\eta_R, \gamma_{\text{per}}^0]\|_{\mathfrak{S}_p} = O(R^{-1+3/p}), \quad \|Y_R - \eta_R\|_{\mathfrak{S}_p} = O(R^{-1+3/p}). \quad (4.55)$$

Moreover $\|X_R^2 - \chi_R^2\|_{\mathfrak{S}_p} = O(R^{-1+3/p})$.

Proof of the Lemma. By (4.104) and (4.105),

$$\begin{aligned} [\gamma_{\text{per}}^0, \eta_R] &= \frac{1}{2i\pi} \int_{\mathcal{C}} [(z - H_{\text{per}}^0)^{-1}, \eta_R] dz, \\ &= \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} [H_{\text{per}}^0, \eta_R] (z - H_{\text{per}}^0)^{-1} dz, \\ &= -\frac{1}{4i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} [\Delta, \eta_R] (z - H_{\text{per}}^0)^{-1} dz, \\ &= -\frac{1}{4i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} (1 - \Delta) (1 - \Delta)^{-1} [\Delta, \eta_R] (1 - \Delta)^{-1} (1 - \Delta) (z - H_{\text{per}}^0)^{-1} dz, \end{aligned}$$

with $(z - H_{\text{per}}^0)^{-1} (1 - \Delta)$ uniformly bounded by Lemma 10 so it remains to prove that $(1 - \Delta)^{-1} [\Delta, \eta_R] (1 - \Delta)^{-1}$ is bounded in \mathfrak{S}_p . Moreover,

$$\begin{aligned} [\Delta, \eta_R] &= \Delta \eta_R + 2\nabla \eta_R \cdot \nabla, \\ &= \sum_i \partial_i^2 \eta_R + 2(\partial_i \eta_R) \partial_i, \\ &= \sum_i \partial_i (\partial_i \eta_R) + (\partial_i \eta_R) \partial_i, \end{aligned}$$

then,

$$\begin{aligned} \|(1 - \Delta)^{-1} [\Delta, \eta_R] (1 - \Delta)^{-1}\|_{\mathfrak{S}_p} &= \left\| (1 - \Delta)^{-1} \left(\sum_i \partial_i (\partial_i \eta_R) + (\partial_i \eta_R) \partial_i \right) (1 - \Delta)^{-1} \right\|_{\mathfrak{S}_p}, \\ &\leq \left\| (1 - \Delta)^{-1} \left(\sum_i \partial_i (\partial_i \eta_R) \right) (1 - \Delta)^{-1} \right\|_{\mathfrak{S}_p} \\ &\quad + \left\| (1 - \Delta)^{-1} \left(\sum_i (\partial_i \eta_R) \partial_i \right) (1 - \Delta)^{-1} \right\|_{\mathfrak{S}_p}, \\ &\leq 2 \left\| \left(\sum_i (\partial_i \eta_R) \right) (1 - \Delta)^{-1} \right\|_{\mathfrak{S}_p}, \\ &\leq C \left\| \sum_i (\partial_i \eta_R) \right\|_{L^q}, \end{aligned}$$

for $q > \frac{3}{2}$. We have first used that applying $(1 - \Delta)^{-1} \partial_i$ is equivalent to multiplying by $\frac{p_i}{1+|p|^2}$ in Fourier mode, which is bounded. Then with the Kato-Seiler-Simon inequality we got the last majoration. Finally,

$$\begin{aligned} \int_{\mathbb{R}^3} |\partial_i \eta_R(x)|^q dx &= \int_{\mathbb{R}^3} |R^{-1} \partial_i \eta_R(x/R)|^q dx, \\ &= \frac{R^3}{R^q} \int_{\mathbb{R}^3} |\partial_i \eta(x/R)|^q \frac{dx}{R^3}, \\ &= \frac{1}{R^{q-3}} \int_{\mathbb{R}^3} |\partial_i \eta(x)|^q dx, \end{aligned}$$

which yields $\|[\eta_R, \gamma_{\text{per}}^0]\|_{\mathfrak{S}_p} \leq C' \frac{1}{R^{1-3/q}} \|\nabla \eta\|_{L^q}$ for $q > 3$.

Eventually we notice

$$\begin{aligned} Y_R - \eta_R &= \gamma_{\text{per}}^0 \eta_R \gamma_{\text{per}}^0 + (\gamma_{\text{per}}^0)^\perp \eta_R (\gamma_{\text{per}}^0)^\perp - \eta_R, \\ &= -(1 - \gamma_{\text{per}}^0) \eta_R \gamma_{\text{per}}^0 - \gamma_{\text{per}}^0 \eta_R (1 - \gamma_{\text{per}}^0), \\ &= -(1 - \gamma_{\text{per}}^0) [\eta_R, \gamma_{\text{per}}^0] + \gamma_{\text{per}}^0 [\eta_R, \gamma_{\text{per}}^0], \\ &= (1 - 2\gamma_{\text{per}}^0) [\gamma_{\text{per}}^0, \eta_R] \end{aligned}$$

and thus (4.55) is proved since γ_{per}^0 is bounded. Finally, $X_R^2 - \chi_R^2 = 1 - Y_R^2 - \chi_R^2 = \eta_R^2 - Y_R^2$, then by (4.55) and since $\|Y_R\| \leq 1$, $\|\eta_R\| \leq 1$, $\|X_R^2 - \chi_R^2\|_{\mathfrak{S}_p} = O(R^{-1+3/p})$.

Lemma 6. *One has*

$$\| [Y_R, |H_{\text{per}}^0 - \Sigma|^{1/2}] |H_{\text{per}}^0 - \Sigma|^{-1/2} \| = O(R^{-1}) \quad (4.56)$$

where we recall that $\Sigma = (\Sigma_Z^+ + \Sigma_{Z+1}^-)/2$ is the middle of the gap.

Proof of the Lemma. We use the well-known integral representation of the square root [3]

$$|H_{\text{per}}^0 - \Sigma|^{1/2} = \frac{1}{\pi} \int_0^\infty \frac{|H_{\text{per}}^0 - \Sigma| ds}{s + |H_{\text{per}}^0 - \Sigma| \sqrt{s}}. \quad (4.57)$$

Recall that $Y_R = \gamma_{\text{per}}^0 \eta_R \gamma_{\text{per}}^0 + (\gamma_{\text{per}}^0)^\perp \eta_R (\gamma_{\text{per}}^0)^\perp$. For shortness, we only detail the estimate of the term involving $(\gamma_{\text{per}}^0)^\perp \eta_R (\gamma_{\text{per}}^0)^\perp$, the proof being the same with γ_{per}^0 as these operators are both bounded and commute with $|H_{\text{per}}^0 - \Sigma|$. Using that, and

then the fact that $[A, \frac{1}{B}] = \frac{1}{B}(BA - AB)\frac{1}{B}$, we see that it suffices to estimate

$$\begin{aligned}
& (\gamma_{\text{per}}^0)^\perp \int_0^\infty \left[\eta_R, \frac{|H_{\text{per}}^0 - \Sigma|}{s + |H_{\text{per}}^0 - \Sigma|} \right] |H_{\text{per}}^0 - \Sigma|^{-1/2} \frac{ds}{\sqrt{s}} (\gamma_{\text{per}}^0)^\perp \\
&= (\gamma_{\text{per}}^0)^\perp \int_0^\infty \left[\eta_R, 1 - \frac{s}{s + |H_{\text{per}}^0 - \Sigma|} \right] |H_{\text{per}}^0 - \Sigma|^{-1/2} \frac{ds}{\sqrt{s}} (\gamma_{\text{per}}^0)^\perp \\
&= (\gamma_{\text{per}}^0)^\perp \int_0^\infty -s \left[\eta_R, \frac{1}{s + |H_{\text{per}}^0 - \Sigma|} \right] |H_{\text{per}}^0 - \Sigma|^{-1/2} \frac{ds}{\sqrt{s}} (\gamma_{\text{per}}^0)^\perp \\
&= (\gamma_{\text{per}}^0)^\perp \int_0^\infty -\frac{s}{s + |H_{\text{per}}^0 - \Sigma|} \left[\eta_R, s + |H_{\text{per}}^0 - \Sigma| \right] \frac{1}{s + |H_{\text{per}}^0 - \Sigma|} |H_{\text{per}}^0 - \Sigma|^{-1/2} \frac{ds}{\sqrt{s}} (\gamma_{\text{per}}^0)^\perp \\
&= (\gamma_{\text{per}}^0)^\perp \int_0^\infty -\frac{1}{s + |H_{\text{per}}^0 - \Sigma|} \left[\eta_R, -\frac{\Delta}{2} \right] \frac{1}{s + |H_{\text{per}}^0 - \Sigma|} |H_{\text{per}}^0 - \Sigma|^{-1/2} \sqrt{s} ds (\gamma_{\text{per}}^0)^\perp \\
&= -\frac{1}{2} \int_0^\infty \frac{(\gamma_{\text{per}}^0)^\perp}{s + |H_{\text{per}}^0 - \Sigma|} \left[\eta_R, -\Delta \right] |H_{\text{per}}^0 - \Sigma|^{-1/2} \frac{(\gamma_{\text{per}}^0)^\perp}{s + |H_{\text{per}}^0 - \Sigma|} \sqrt{s} ds.
\end{aligned} \tag{4.58}$$

Then $[\eta_R, -\Delta] = (\Delta\eta_R) + 2(\nabla\eta_R) \cdot \nabla$, hence $\|[\eta_R, -\Delta] |H_{\text{per}}^0 - \Sigma|^{-1/2}\| = O(R^{-1})$ where we have used that $\nabla |H_{\text{per}}^0 - \Sigma|^{-1/2}$ is a bounded operator by Lemma 1. Then we use that $|H_{\text{per}}^0 - \Sigma| \geq \epsilon$ for some $\epsilon > 0$ to estimate the last right hand side of (4.58) in the operator norm by

$$\|[\eta_R, -\Delta] |H_{\text{per}}^0 - \Sigma|^{-1/2}\| \times \int_0^\infty \frac{\sqrt{s} ds}{(\epsilon + s)^2} = O(R^{-1}).$$

Notice now that $Y_R Q_n Y_R \in \mathcal{K}$ for all $R \geq 1$ (the same is true for $X_R Q_n X_R$ but we shall actually not need it). To see this, notice for instance that

$$\begin{aligned}
(1 + |\nabla|)(Y_R Q_n Y_R) &= -(1 + |\nabla|) |H_{\text{per}}^0 - \Sigma|^{-1/2} \left([Y_R, |H_{\text{per}}^0 - \Sigma|^{1/2}] \times \right. \\
&\quad \left. \times |H_{\text{per}}^0 - \Sigma|^{-1/2} |H_{\text{per}}^0 - \Sigma|^{1/2} Q_n Y_R - Y_R |H_{\text{per}}^0 - \Sigma|^{1/2} Q_n Y_R \right) \tag{4.59}
\end{aligned}$$

which belongs to \mathfrak{S}_2 since $|H_{\text{per}}^0 - \Sigma|^{1/2} Q_n \in \mathfrak{S}_2$ and $[Y_R, |H_{\text{per}}^0 - \Sigma|^{1/2}] |H_{\text{per}}^0 - \Sigma|^{-1/2}$ is bounded by Lemma 6. The proof that $(1 + |\nabla|)(Y_R Q_n Y_R)^{++}(1 + |\nabla|)$ and $(1 + |\nabla|)(Y_R Q_n Y_R)^{--}(1 + |\nabla|)$ are trace class is similar since γ_{per}^0 and $(\gamma_{\text{per}}^0)^\perp$ are bounded and commute with $|H_{\text{per}}^0 - \Sigma|$. Eventually, to prove that $-\gamma_{\text{per}}^0 \leq Y_R Q_n Y_R \leq 1 - \gamma_{\text{per}}^0$, we use the equivalent condition (4.17), the fact that γ_{per}^0 and $(\gamma_{\text{per}}^0)^\perp$ commute with Y_R and that $0 \leq Y_R \leq 1$

$$\begin{aligned}
(Q_n)^2 &\leq Q_n^{++} - Q_n^{--}, \\
Y_R(Q_n)^2 Y_R &\leq Y_R Q_n^{++} Y_R - Y_R Q_n^{--} Y_R, \\
Y_R Q_n (Y_R)^2 Q_n Y_R &\leq Y_R (Q_n)^2 Y_R \leq (\gamma_{\text{per}}^0)^\perp Y_R Q_n Y_R (\gamma_{\text{per}}^0)^\perp - \gamma_{\text{per}}^0 Y_R Q_n Y_R \gamma_{\text{per}}^0, \\
(Y_R Q_n Y_R)^2 &\leq (Y_R Q_n Y_R)^{++} - (Y_R Q_n Y_R)^{--}, \\
-\gamma_{\text{per}}^0 &\leq Y_R Q_n Y_R \leq 1 - \gamma_{\text{per}}^0.
\end{aligned}$$

We are now able to prove (4.54) as announced. We write, following [23],

$$\begin{aligned} \mathcal{E}^\nu(Q_n) &= \text{Tr}(X_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q_n^{++} - Q_n^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}X_R) \\ &\quad + \text{Tr}(Y_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q_n^{++} - Q_n^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}Y_R) + \Sigma\text{Tr}_0(X_RQ_nX_R) \\ &\quad + \Sigma\text{Tr}_0(Y_RQ_nY_R) + \frac{1}{2}D(\rho_{Q_n} - \nu, \rho_{Q_n} - \nu) - \frac{1}{2}D(\nu, \nu) \end{aligned} \quad (4.60)$$

where we have used that $[\gamma_{\text{per}}^0, X_R] = [\gamma_{\text{per}}^0, Y_R] = 0$ to infer $\text{Tr}_0(Q_n) = \text{Tr}_0(X_RQ_nX_R) + \text{Tr}_0(Y_RQ_nY_R)$. Then, by Lemma 6 and using the fact that $(Q_n)_{n \geq 1}$ is a bounded sequence in \mathcal{Q} , we deduce that

$$\begin{aligned} \text{Tr}(Y_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q_n^{++} - Q_n^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}Y_R) \\ \geq \text{Tr}(|H_{\text{per}}^0 - \Sigma|^{1/2}Y_R(Q_n^{++} - Q_n^{--})Y_R|H_{\text{per}}^0 - \Sigma|^{1/2}) - C/R \end{aligned} \quad (4.61)$$

for some constant $C > 0$. Arguing similarly for the other terms, we obtain

$$\begin{aligned} \mathcal{E}^\nu(Q_n) &\geq \tilde{\mathcal{E}}^0(Y_RQ_nY_R) + \text{Tr}(\chi_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q_n^{++} - Q_n^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}\chi_R) \\ &\quad + \Sigma\text{Tr}(\chi_R(Q_n^{++} + Q_n^{--})\chi_R) + \frac{1}{2}D(\rho_{Q_n} - \nu, \rho_{Q_n} - \nu) - \frac{1}{2}D(\nu, \nu) - C'/R \end{aligned} \quad (4.62)$$

for some constant C' , where

$$\tilde{\mathcal{E}}^0(Q) := \text{Tr}(|H_{\text{per}}^0 - \Sigma|^{1/2}(Q^{++} - Q^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}) + \Sigma\text{Tr}_0(Q).$$

Recall (Proposition 3)

$$E^0(q) = \inf\{\tilde{\mathcal{E}}^0(Q), \quad Q \in \mathcal{K}_r, \quad \text{Tr}_0(Q) = q\}.$$

Thus, using

$$q = \text{Tr}_0(Q_n) = \text{Tr}_0(Y_RQ_nY_R) + \text{Tr}_0(X_RQ_nX_R),$$

and the fact that $q \mapsto E^0(q)$ is Lipschitz by Proposition 3, (4.62) yields

$$\begin{aligned} \mathcal{E}^\nu(Q_n) &\geq \text{Tr}(\chi_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q_n^{++} - Q_n^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}\chi_R) \\ &\quad + \Sigma\text{Tr}(\chi_R(Q_n^{++} + Q_n^{--})\chi_R) + E^0(q - \text{Tr}(\chi_R(Q_n^{++} + Q_n^{--})\chi_R)) \\ &\quad + \frac{1}{2}D(\rho_{Q_n} - \nu, \rho_{Q_n} - \nu) - \frac{1}{2}D(\nu, \nu) - C'/R. \end{aligned} \quad (4.63)$$

Let us now pass to the limit $n \rightarrow \infty$. First we notice

$$\begin{aligned} \liminf_{n \rightarrow \infty} \text{Tr}(\chi_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q_n^{++} - Q_n^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}\chi_R) \\ \geq \text{Tr}(\chi_R|H_{\text{per}}^0 - \Sigma|^{1/2}(Q^{++} - Q^{--})|H_{\text{per}}^0 - \Sigma|^{1/2}\chi_R), \end{aligned} \quad (4.64)$$

$$\liminf_{n \rightarrow \infty} D(\rho_{Q_n} - \nu, \rho_{Q_n} - \nu) \geq D(\rho_Q - \nu, \rho_Q - \nu) \quad (4.65)$$

by Fatou's Lemma and the weak convergence $\rho_{Q_n} \rightharpoonup \rho_Q$ in \mathcal{C} . Then

$$\lim_{n \rightarrow \infty} \text{Tr}(\chi_RQ_n^{++}\chi_R) = \text{Tr}(\chi_RQ^{++}\chi_R), \quad \lim_{n \rightarrow \infty} \text{Tr}(\chi_RQ_n^{--}\chi_R) = \text{Tr}(\chi_RQ^{--}\chi_R)$$

which is obtained by writing for instance

$$\mathrm{Tr}(\chi_R Q_n^{++} \chi_R) = \mathrm{Tr}(\chi_R (1 + |\nabla|)^{-1} (1 + |\nabla|) Q_n^{++} (1 + |\nabla|) (1 + |\nabla|)^{-1} \chi_R)$$

and using that $\chi_R (1 + |\nabla|)^{-1}$ is compact (it belongs to \mathfrak{S}_p for $p > 3$ by the Kato-Seiler-Simon inequality) and that

$$(1 + |\nabla|) Q_n^{++} (1 + |\nabla|) \rightharpoonup (1 + |\nabla|) Q^{++} (1 + |\nabla|)$$

for the weak-* topology of \mathfrak{S}_1 . Thus,

$$\begin{aligned} E^\nu(q) &= \liminf_{n \rightarrow \infty} \mathcal{E}^\nu(Q_n) \geq \mathrm{Tr}(\chi_R |H_{\mathrm{per}}^0 - \Sigma|^{1/2} (Q^{++} - Q^{--}) |H_{\mathrm{per}}^0 - \Sigma|^{1/2} \chi_R) \\ &\quad + \Sigma \mathrm{Tr}(\chi_R (Q^{++} + Q^{--}) \chi_R) + E^0(q - \mathrm{Tr}(\chi_R (Q^{++} + Q^{--}) \chi_R)) \\ &\quad + \frac{1}{2} D(\rho_Q - \nu, \rho_Q - \nu) - \frac{1}{2} D(\nu, \nu) - C'/R. \end{aligned} \quad (4.66)$$

Passing now to the limit as $R \rightarrow \infty$, we have that $\chi_R \rightarrow \chi(0) = 1$ and therefore

$$\begin{aligned} E^\nu(q) &\geq \mathrm{Tr}(|H_{\mathrm{per}}^0 - \Sigma|^{1/2} (Q_n^{++} - Q_n^{--}) |H_{\mathrm{per}}^0 - \Sigma|^{1/2}) \\ &\quad + \Sigma \mathrm{Tr}(Q^{++} + Q^{--}) + E^0(q - \mathrm{Tr}(Q^{++} + Q^{--})) \\ &\quad + \frac{1}{2} D(\rho_Q - \nu, \rho_Q - \nu) - \frac{1}{2} D(\nu, \nu), \end{aligned}$$

which is equal to $E^\nu(q - q') + E^0(q')$ where we recall that $q' = q - \mathrm{Tr}(Q^{++} + Q^{--}) = q - \mathrm{Tr}_0(Q)$, so we obtain (4.54). This contradicts (3) and ends the proof that **(b)** \Leftrightarrow **(c)** in Theorem 3.

Step 4 : Characterization of the q 's such that **(c)** holds

Because $q \mapsto E^\nu(q)$ is a convex function, it is classical that the set $I = \{q \in \mathbb{R}, \text{ (c) holds}\}$ is a closed interval of \mathbb{R} , see e.g. [21, 46]. It is non-empty since it contains $\mathrm{Tr}_0(\bar{Q})$ for any minimizer \bar{Q} of $E_{\epsilon_F}^\nu$ obtained in Theorem 2, for any ϵ_F in the gap $(\Sigma_Z^+, \Sigma_{Z+1}^-)$. Additionally, $q \mapsto E^\nu(q)$ is linear on the connected components of $\mathbb{R} \setminus I$ and I is the largest interval on which this function is strictly convex. Let us now state and prove the

Lemma 7. *Let $Z \in \mathbb{N} \setminus \{0\}$, $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$, and assume that **(A1)** holds. Assume that Q_1 and Q_2 are respectively two minimizers of $E^\nu(q_1)$ and $E^\nu(q_2)$ with $q_1 \neq q_2$. Then $\rho_{Q_1} \neq \rho_{Q_2}$ and therefore*

$$E^\nu(tq_1 + (1-t)q_2) \leq \mathcal{E}^\nu(tQ_1 + (1-t)Q_2) < tE^\nu(q_1) + (1-t)E^\nu(q_2).$$

Proof of the Lemma. Assume by contradiction that $\rho_{Q_1} = \rho_{Q_2}$. As we saw before, the operators Q_1 and Q_2 satisfy the self-consistent equations

$$Q_1 = \chi_{(-\infty, \epsilon_1)}(H_{Q_1}) - \gamma_{\mathrm{per}}^0 + \delta_1, \quad Q_2 = \chi_{(-\infty, \epsilon_2)}(H_{Q_2}) - \gamma_{\mathrm{per}}^0 + \delta_2$$

where $0 \leq \delta_k \leq 1$ and $\text{Ran}(\delta_k) \subseteq \ker(H_{Q_k} - \epsilon_k)$ for $k = 1, 2$. Necessarily ϵ_1 and ϵ_2 are in $[\Sigma_Z^+, \Sigma_{Z+1}^-]$ otherwise Q_1 and Q_2 would not be compact, which is not possible since every operator of \mathcal{K} is compact. Since $H_{Q_1} = H_{\text{per}}^0 + (\rho_{Q_1} - \nu) * |\cdot|^{-1} = H_{\text{per}}^0 + (\rho_{Q_2} - \nu) * |\cdot|^{-1} = H_{Q_2}$ has only a point spectrum in the gap, we deduce that if $\epsilon_k \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, then necessarily $\ker(H_{Q_k} - \epsilon_k)$ is a finite dimension space and thus δ_k is finite rank. If $\epsilon_k \in \{\Sigma_Z^+, \Sigma_{Z+1}^-\}$, then it can easily be proved that at least $\delta_k \in \mathfrak{S}_1$. Hence Q_1 and Q_2 differ only by a trace-class operator : $Q_2 = Q_1 + \delta$, with $\delta = \delta_2 - \delta_1$, thus in \mathfrak{S}_1 : $\text{Tr}|\delta| < \infty$. Now $0 \neq q_2 - q_1 = \text{Tr}(\delta) = \int \rho_\delta$ which contradicts our assumption that $\rho_\delta = \rho_{Q_1} - \rho_{Q_2} = 0$. Then, for the energy, we have

$$\begin{aligned} E^\nu(tq_1 + (1-t)q_2) &= \inf\{\mathcal{E}^\nu(Q), Q \in \mathcal{K}, \text{Tr}_0(Q) = tq_1 + (1-t)q_2\}, \\ &\leq \mathcal{E}^\nu(tQ_1 + (1-t)Q_2), \\ &< t\mathcal{E}^\nu(Q_1) + (1-t)\mathcal{E}^\nu(Q_2), \\ &= tE^\nu(q_1) + (1-t)E^\nu(q_2), \end{aligned}$$

by the strict convexity of $\rho \mapsto D(\rho, \rho)$ and since $\rho_{Q_1} \neq \rho_{Q_2}$.

Corollary 3. *There is no minimizer for $E^\nu(q)$ if $q \notin I$, the interval on which (c) holds. Thus (a) implies (c).*

Proof of the Corollary. Assume that there is a minimizer Q_1 for some $q_1 \notin I$, for instance $q_1 > \max I := q_2$. Applying Lemma 7 to q_1 and q_2 shows that $E^\nu(\cdot)$ will be strictly convex (and cannot be linear) on $[q_2, q_1]$, which contradicts the definition of I being the biggest interval on which $E^\nu(\cdot)$ is strictly convex.

Conditions like (c) appear classically when analyzing the compactness properties of minimizing sequences, for instance by using the concentration-compactness principle of P.-L. Lions [39]. They are also very classical for linear models in which the bottom of the essential spectrum has the form of the minimum with respect to q' of the right hand side of (c), as expressed by the HVZ Theorem [28, 48, 49]. Assume for simplicity that $q > 0$ and that \bar{Q} can be written as in (4.36) and (4.37). When $0 < q' \leq q$, (c) means that it is not favorable to let q' electrons escape to infinity, while keeping $q - q'$ electrons near the defect. When $q' < 0$, it means that it is not favorable to let $|q'|$ holes escape to infinity, while keeping $q + |q'|$ electrons near the defect. When $q' > q$, it means that it is not favorable to let q' electrons escape to infinity, while keeping $q' - q$ holes near the defect.

We do not show here when (c) holds true. Proving (c) usually requires some decay property of the density of charge $\rho_{\bar{Q}}$ for a solution \bar{Q} of the nonlinear equation (4.32). In particular, knowing that $\rho_{\bar{Q}} \in L^1(\mathbb{R}^3)$ would be very useful (see Remark 3). We plan to investigate more closely the decay properties of $\rho_{\bar{Q}}$ and the validity of (c) in the near future. At present, the validity of a condition similar to (c) for the Bogoliubov-Dirac-Fock model was only proved in the nonrelativistic limit or in the weak coupling limit, see [24].

4.4 Thermodynamic limit of the supercell model

As mentioned before, we shall now justify the model of the previous section by proving that it is the thermodynamic limit of the supercell model.

Let us emphasize that there are several ways of performing thermodynamic limits. In [12], the authors consider a box of size L , $\Lambda_L := [-L/2, L/2]^3$, and assume that the nuclei are located on $\mathbb{Z}^3 \cap \Lambda_L$. Then they consider the rHF model of Section 4.1 for N electrons living *in the whole space*, with $N = ZL^3$ chosen to impose neutrality. Denoting by ρ_L the ground state electronic density of the latter problem, it is proved in [12, Thm 2.2] that the energy per unit cell converges to I_{per}^0 , and that the following holds :

$$\sqrt{L^{-3} \sum_{k \in \mathbb{Z}^3 \cap \Lambda_L} \rho_L(x - k)} \rightarrow \sqrt{\rho \gamma_{\text{per}}^0} \quad (4.67)$$

weakly in $H_{\text{loc}}^1(\mathbb{R}^3)$, strongly in $L_{\text{loc}}^p(\mathbb{R}^3)$ for all $2 \leq p < 6$ and almost everywhere on \mathbb{R}^3 when $L \rightarrow \infty$. Let us recall that I_{per}^0 and γ_{per}^0 are defined in Section 4.2.

Another way for performing thermodynamic limits is to confine the nuclei *and* the electrons in a domain Ω_L with $|\Omega_L| \rightarrow \infty$, by means of Dirichlet boundary conditions for the electrons. The latter approach was chosen for the Schrödinger model with quantum nuclei in the canonical and grand canonical ensembles [43] by Lieb and Lebowitz in the seminal paper [34] (see also [31]), where the existence of a limit for the energy per unit volume is proved. The crystal case in the Schrödinger model was tackled by Fefferman [18] in the same spirit. We do not know whether Fefferman's proof can be adapted to treat the Hartree-Fock case.

Another possibility, perhaps less satisfactory from a physical viewpoint but more directly related to practical calculations (see e.g. [16]), is to take $\Omega_L = \Lambda_L$ and to impose periodic boundary conditions on the box Λ_L . Usually the Coulomb interaction is also replaced by a $(L\mathbb{Z}^3)$ -periodic Coulomb potential, leading to the so-called *supercell model* which will be described in details below. This approach has the advantage of respecting the symmetry of the system in the crystal case. It was used by Hainzl, Lewin and Solovej in [25] to justify the Hartree-Fock approximation of no-photon Quantum Electrodynamics. The supercell limit of a linear model for photonic crystals is studied in [75].

Of course the conjecture is that the final results (the energy per unit cell and the ground state density of the crystal) should not depend on the chosen thermodynamic limit procedure. This is actually the case for the reduced Hartree-Fock model of the crystal. See [26] for a result in this direction for a model with quantum nuclei.

Let us now describe the supercell model. For $L \in \mathbb{N} \setminus \{0\}$, we introduce the supercell $\Lambda_L = [-L/2, L/2]^3$ and the Hilbert space

$$L_{\text{per}}^2(\Lambda_L) = \{ \varphi \in L_{\text{loc}}^2(\mathbb{R}^3) \mid \varphi \text{ (} LZ^3 \text{)-periodic} \}.$$

The lattice \mathcal{R} is now $L\mathbb{Z}^3$ and its reciprocal lattice $\mathcal{R}^* = \frac{2\pi}{L}\mathbb{Z}^3$, hence the Brillouin zone Γ^* becomes $[-\frac{2\pi\eta}{L}, \frac{2\pi(1-\eta)}{L}]^3$, with $\eta = 0$ if L is even and $\eta = 1/2$ if L is odd. We keep the notation $\Gamma = [-\frac{1}{2}, \frac{1}{2}]^3 = \Lambda_1$.

We also introduce the $L\mathbb{Z}^3$ -periodic Coulomb potential G_L defined as the unique solution to

$$\begin{cases} -\Delta G_L = 4\pi \left(\sum_{k \in \mathbb{Z}^3} \delta_k L - \frac{1}{L^3} \right) \\ G_L \text{ } L\mathbb{Z}^3\text{-periodic, } \min_{\mathbb{R}^3} G_L = 0. \end{cases}$$

It is easy to check (see Section 2.4) that $G_L(x) = L^{-1}G_1(x/L)$ and that

$$G_L(x) = \frac{c}{L} + \sum_{k \in \mathcal{R}^* \setminus \{0\}} \frac{4\pi}{|k|^2} \frac{1}{L^3} e^{ik \cdot x}.$$

Now we recall that for any $L\mathbb{Z}^3$ -periodic function g ,

$$(g \star_{\Lambda_L} G_L)(x) := \int_{\Lambda_L} G_L(x-y) g(y) dy,$$

$$D_{G_L}(f, g) := \int_{\Lambda_L} \int_{\Lambda_L} G_L(x-y) f(x) g(y) dx dy.$$

An admissible electronic state is then described by a one-body density matrix γ in

$$\mathcal{P}_{\text{sc},L} = \left\{ \gamma \in \mathfrak{S}_1(L^2_{\text{per}}(\Lambda_L)) \mid \gamma^* = \gamma, 0 \leq \gamma \leq 1, \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(-\Delta\gamma) < +\infty \right\}.$$

Any $\gamma \in \mathcal{P}_{\text{sc},L}$ has a well-defined density of charge $\rho_\gamma(x) = \gamma(x, x)$, where $\gamma(x, y)$ is the kernel of the operator γ . Notice that $\gamma(x + Lz, y + Lz') = \gamma(x, y)$ for any $z, z' \in \mathbb{Z}^3$, which implies that ρ_γ is $L\mathbb{Z}^3$ -periodic.

Throughout this section, we use the subscript ‘sc’ to indicate that we consider the thermodynamic limit of the supercell model.

4.4.1 Thermodynamic limit without defect

Because our model with defect uses the defect-free density matrix of the Fermi sea as a reference, we need to start with the study of the thermodynamic limit without defect. We are going to prove for the supercell model a result analogous to [12, Thm 2.2].

The reduced Hartree-Fock energy functional of the supercell model is defined for $\gamma \in \mathcal{P}_{\text{sc},L}$ as

$$\mathcal{E}_{\text{sc},L}^0(\gamma) = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(-\frac{1}{2} \Delta \gamma \right) + \frac{1}{2} D_{G_L}(\rho_\gamma - \mu_{\text{per}}, \rho_\gamma - \mu_{\text{per}})$$

where we recall that $\mu_{\text{per}}(x) = \sum_{R \in \mathbb{Z}^3} Zm(x-R)$ is a \mathbb{Z}^3 - (thus $L\mathbb{Z}^3$ -) periodic function. The reduced Hartree-Fock ground state energy for a neutral system in the box of size L is then given by

$$I_{\text{sc},L}^0 = \inf \left\{ \mathcal{E}_{\text{sc},L}^0(\gamma), \gamma \in \mathcal{P}_{\text{sc},L}, \int_{\Lambda_L} \rho_\gamma = \int_{\Lambda_L} \mu_{\text{per}} = ZL^3 \right\}. \quad (4.68)$$

Let us recall that I_{per}^0 , γ_{per}^0 and H_{per}^0 are defined in Section 4.2. We are going to prove the following

Theorem 4 (Thermodynamic limit of the defect-free supercell model). *Let $Z > 0$.*

i) For all $L \in \mathbb{N} \setminus \{0\}$, the minimizing problem $I_{\text{sc},L}^0$ has at least one minimizer, and all the minimizers share the same density. This density is \mathbb{Z}^3 -periodic. Besides, there is one minimizer $\gamma_{\text{sc},L}^0$ of (4.68) which commutes with the translations τ_k , $k \in \mathbb{Z}^3$.

ii) The following thermodynamic limit properties hold true :

- (Convergence of the energy per unit cell).

$$\lim_{L \rightarrow \infty} \frac{I_{\text{sc},L}^0}{L^3} = I_{\text{per}}^0;$$

- (Convergence of the density).

$$\sqrt{\rho_{\gamma_{\text{sc},L}^0}} \rightharpoonup \sqrt{\rho_{\gamma_{\text{per}}^0}} \quad \text{weakly in } H_{\text{loc}}^1(\mathbb{R}^3), \quad (4.69)$$

$$\rho_{\gamma_{\text{sc},L}^0} \rightarrow \rho_{\gamma_{\text{per}}^0} \quad \text{strongly in } L_{\text{loc}}^p(\mathbb{R}^3) \quad \text{for } 1 \leq p < 3 \text{ and a.e.};$$

- (Convergence of the mean-field Hamiltonian and its spectrum). *Let*

$$H_{\text{sc},L}^0 = -\frac{\Delta}{2} + (\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1$$

seen as an operator acting on $L^2(\mathbb{R}^3)$. Then, for all $L \in \mathbb{N} \setminus \{0\}$, $H_{\text{sc},L}^0 - H_{\text{per}}^0$ is a bounded operator and

$$\lim_{L \rightarrow \infty} \|H_{\text{sc},L}^0 - H_{\text{per}}^0\| = 0.$$

Denoting by $(\lambda_n^L(\xi))_{n \in \mathbb{N} \setminus \{0\}}$ the nondecreasing sequence of eigenvalues of $(H_{\text{sc},L}^0)_{\xi}$ for $\xi \in \Gamma^$, one has*

$$\lim_{L \rightarrow \infty} \sup_{n \in \mathbb{N} \setminus \{0\}} \sup_{\xi \in \Gamma^*} |\lambda_n^L(\xi) - \lambda_n(\xi)| = 0 \quad (4.70)$$

where $(\lambda_n(\xi))_{n \geq 1}$ are the eigenvalues of $(H_{\text{per}}^0)_{\xi}$ introduced in Theorem 1.

*iii) Assume in addition that **(A1)** holds. Fix some $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. Then for L large enough, the minimizer $\gamma_{\text{sc},L}^0$ of $I_{\text{sc},L}^0$ is unique. It is also the unique minimizer of the following problem*

$$I_{\text{sc},L,\epsilon_F}^0 := \inf \left\{ \mathcal{E}_{\text{sc},L}^0(\gamma) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(\gamma), \quad \gamma \in \mathcal{P}_{\text{sc},L} \right\}. \quad (4.71)$$

Proof.

Step 1

Let us first prove that $\limsup_{L \rightarrow +\infty} \frac{1}{L^3} I_{\text{sc},L}^0 \leq I_{\text{per}}^0$. Starting from the Bloch decomposition

$$\gamma_{\text{per}}^0 = \frac{1}{(2\pi)^3} \int_{\Gamma^*} (\gamma_{\text{per}}^0)_\xi d\xi$$

of γ_{per}^0 it is possible to construct a convenient test function $\tilde{\gamma}_{\text{sc},L}$ for (4.68) as follows :

$$\tilde{\gamma}_{\text{sc},L}(x, y) = \frac{1}{(2\pi)^3} \sum_{\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} e^{i\xi x} \left(\int_{\xi + [-\frac{2\pi\eta}{L}, \frac{2\pi(1-\eta)}{L}]^3} e^{-i\xi' x} (\gamma_{\text{per}}^0)_{\xi'}(x, y) e^{i\xi' y} d\xi' \right) e^{-i\xi y},$$

with $\eta = 0$ if L is even and $\eta = 1/2$ if L is odd.

It is easy to check that $\tilde{\gamma}_{\text{sc},L}$ is self-adjoint, nonnegative and satisfies $\rho_{\tilde{\gamma}_{\text{sc},L}} = \rho_{\gamma_{\text{per}}^0}$ (see Section 2.4). In particular, since $\frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_\xi^2(\Gamma)} (\gamma_{\text{per}}^0)_\xi d\xi = Z < \infty$, we can write

$$\int_{\Lambda_L} \rho_{\tilde{\gamma}_{\text{sc},L}} = \int_{\Lambda_L} \mu_{\text{per}} = ZL^3.$$

Then, when $f \in L_{\text{per}}^2(\Lambda_L)$, $\tilde{\gamma}_{\text{sc},L}f$ is also in $L_{\text{per}}^2(\Lambda_L)$ if

$$\forall z \in LZ^3, \tilde{\gamma}_{\text{sc},L}(x+z, y) = \tilde{\gamma}_{\text{sc},L}(x, y+z) = \tilde{\gamma}_{\text{sc},L}(x, y) \quad \forall x, y \in \mathbb{R}^3.$$

To ensure that, we use that $(\gamma_{\text{per}}^0)_\xi$ is in $L_\xi^2(\Gamma)$, which implies that $e^{-i\xi' x} (\gamma_{\text{per}}^0)_{\xi'}(x, y) e^{i\xi' y} \in L_{\text{per}}^2$: for $z \in LZ^3$,

$$e^{-i\xi'(x+z)} (\gamma_{\text{per}}^0)_{\xi'}(x+z, y) e^{i\xi' y} = e^{-i\xi' x} (\gamma_{\text{per}}^0)_{\xi'}(x, y) e^{i\xi' y} = e^{-i\xi' x} (\gamma_{\text{per}}^0)_{\xi'}(x, y+z) e^{i\xi'(y+z)}.$$

As $\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*$, $e^{i\xi(x+z)} = e^{i\xi x}$ and $e^{-i\xi(y+z)} = e^{-i\xi y}$ for every $z \in LZ^3$. So $\tilde{\gamma}_{\text{sc},L} \in \mathfrak{S}_1(L_{\text{per}}^2(\Lambda_L))$ and it was verified in Section 2.4 that finally $\tilde{\gamma}_{\text{sc},L}$ is in $\mathcal{P}_{\text{sc},L}$. Now, since both $\rho_{\gamma_{\text{per}}^0}$ and μ_{per} are \mathbb{Z}^3 -periodic,

$$D_{G_L} (\rho_{\tilde{\gamma}_{\text{sc},L}} - \mu_{\text{per}}, \rho_{\tilde{\gamma}_{\text{sc},L}} - \mu_{\text{per}}) = L^3 D_{G_1} (\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}, \rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}).$$

To establish this inequality, we use the fact that $W_L(x) = \int_{\Lambda_L} G_L(x-y)f(y) dy$ verifies

$$\begin{cases} -\Delta W_L = 4\pi(f - \int_{\Lambda_L} f), \\ \int_{\Lambda_L} W_L = 0, \end{cases} \quad (4.72)$$

with f \mathbb{Z}^3 -periodic and W_L $L\mathbb{Z}^3$ -periodic. Besides $W_1 = \int_{\Gamma} G_1(x-y)f(y) dy$ verifies

$$\begin{cases} -\Delta W_1 = 4\pi(f - \int_{\Gamma} f), \\ \int_{\Gamma} W_1 = 0, \end{cases}$$

with f \mathbb{Z}^3 -periodic and W_1 \mathbb{Z}^3 -periodic hence $L\mathbb{Z}^3$ -periodic. The potential W_1 therefore also verifies (4.72) and by uniqueness of the solution of this elliptic system, $W_L = W_1$ and therefore, taking $f = \rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}$, the result is proved.

Besides, the kinetic part of the energy reads

$$\begin{aligned} \frac{1}{L^3} \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{1}{2} \Delta \tilde{\gamma}_{\text{sc},L} \right) &= \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_{\xi}^2(\Gamma)} \left(-\frac{1}{2} \Delta (\gamma_{\text{per}}^0)_{\xi} \right) d\xi \\ &\quad - \frac{1}{2(2\pi)^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \int_{\xi + [-\frac{2\pi\eta}{L}, \frac{2\pi(1-\eta)}{L}]^3} |\xi - \xi'|^2 \text{Tr}_{L_{\xi'}^2(\Gamma)} ((\gamma_{\text{per}}^0)_{\xi'}) d\xi' \\ &\quad - \frac{i}{(2\pi)^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \int_{\xi + [-\frac{2\pi\eta}{L}, \frac{2\pi(1-\eta)}{L}]^3} (\xi - \xi') \cdot \text{Tr}_{L_{\xi'}^2(\Gamma)} (-i\nabla (\gamma_{\text{per}}^0)_{\xi'}) d\xi'. \end{aligned}$$

It follows from the boundedness of $\int_{\Gamma^*} \text{Tr}_{L_{\xi}^2(\Gamma)} ((1 - \Delta)(\gamma_{\text{per}}^0)_{\xi}) d\xi$ and from the inequality $|-2i\nabla| \leq (1 - \Delta)$ that the last two terms of the above expression are bounded by $C(\frac{2\pi}{L})^3$ and therefore go to zero when $L \rightarrow +\infty$, hence that

$$\lim_{L \rightarrow +\infty} \frac{1}{L^3} \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{1}{2} \Delta \tilde{\gamma}_{\text{sc},L} \right) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L_{\xi}^2(\Gamma)} \left(-\frac{1}{2} (\Delta \gamma_{\text{per}}^0)_{\xi} \right) d\xi.$$

Therefore $\lim_{L \rightarrow +\infty} L^{-3} \mathcal{E}_{\text{sc},L}^0(\tilde{\gamma}_{\text{sc},L}) = \mathcal{E}_{\text{per}}^0(\gamma_{\text{per}}^0) = I_{\text{per}}^0$ and consequently

$$\limsup_{L \rightarrow +\infty} \frac{1}{L^3} I_{\text{sc},L}^0 \leq \lim_{L \rightarrow +\infty} \frac{1}{L^3} \mathcal{E}_{\text{sc},L}^0(\tilde{\gamma}_{\text{sc},L}) = I_{\text{per}}^0.$$

Step 2

Let us now establish that $\liminf_{L \rightarrow +\infty} \frac{1}{L^3} I_{\text{sc},L}^0 \geq I_{\text{per}}^0$. First, the existence of a minimizer γ_L for (4.68) follows from the same arguments that we used before in Section 2.2 (see also the arguments in the proof of [12, Thm 2.1]). We now show the uniqueness of the corresponding density $\rho_{\text{sc},L}^0$. Let us assume that there exist two minimizers γ_1, γ_2 with corresponding densities ρ_1, ρ_2 , and $\rho_1 \neq \rho_2$. Then, by the strict convexity of D_{G_L} and the linearity of the trace, we have

$$\begin{aligned} \mathcal{E}_{\text{sc},L}^0(t\gamma_1 + (1-t)\gamma_2) &< t \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{\Delta}{2} \gamma_1 \right) + (1-t) \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{\Delta}{2} \gamma_2 \right) \\ &\quad + \frac{t}{2} D_{G_L}(\rho_1 - \mu_{\text{per}}, \rho_1 - \mu_{\text{per}}) + \frac{(1-t)}{2} D_{G_L}(\rho_2 - \mu_{\text{per}}, \rho_2 - \mu_{\text{per}}), \\ &= t \mathcal{E}_{\text{sc},L}^0(\gamma_1) + (1-t) \mathcal{E}_{\text{sc},L}^0(\gamma_2), \\ &= I_{\text{sc},L}^0, \end{aligned}$$

which contradicts the hypothesis of $I_{\text{sc},L}^0$ being the minimizer as $t\gamma_1 + (1-t)\gamma_2 \in \mathcal{P}_{\text{sc},L}$. Hence, $\rho_1 = \rho_2$.

Note that, by symmetry, $\rho_{\text{sc},L}^0$ is \mathbb{Z}^3 -periodic. We now define the operator $\gamma_{\text{sc},L}^0$ as

$$\gamma_{\text{sc},L}^0 = \frac{1}{L^3} \sum_{k \in \mathbb{Z}^3 \cap \Lambda_L} \tau_k^* \gamma_L \tau_k.$$

The Laplacian commuting with the translations of \mathbb{Z}^3 , it holds

$$\forall k \in \mathbb{Z}^3, \text{Tr} \left(-\frac{\Delta}{2} \tau_k^* \gamma_L \tau_k \right) = \text{Tr} \left(\tau_k^* \left(-\frac{\Delta}{2} \right) \gamma_L \tau_k \right) = \text{Tr} \left(-\frac{\Delta}{2} \gamma \right).$$

The second term of the energy is the same for $\tau_k^* \gamma_L \tau_k$ as for γ_L since $\rho_{\tau_k^* \gamma_L \tau_k}(x) = \rho_{\text{sc},L}^0(x - k)$ and $\rho_{\text{sc},L}^0$ is \mathbb{Z}^3 -periodic. Then it is clear that $\tau_k^* \gamma_L \tau_k$ is also a minimizer for (4.68) for all $k \in \mathbb{Z}^3$. By convexity, so is $\gamma_{\text{sc},L}^0$. Besides, $\gamma_{\text{sc},L}^0$ commutes with the translations τ_k for all $k \in \mathbb{Z}^3$ so that its kernel $\gamma_{\text{sc},L}^0(x, y)$ satisfies

$$\forall (x, y, z) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{Z}^3, \quad \gamma_{\text{sc},L}^0(x + z, y + z) = \gamma_{\text{sc},L}^0(x, y).$$

As we saw in Section 2.4, the optimality conditions imply that $\gamma_{\text{sc},L}^0$ can be expanded as follows

$$\gamma_{\text{sc},L}^0(x, y) = \frac{1}{L^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} n_{k,\xi}^L e^{i\xi \cdot x} v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)} e^{-i\xi \cdot y}$$

where for any $\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*$, $(v_{k,\xi}^L)_{k \geq 1}$ is an orthonormal basis of $L_{\text{per}}^2(\Gamma)$ consisting of eigenfunctions of the self-adjoint operator on $L_{\text{per}}^2(\Gamma)$ defined by

$$-\frac{1}{2} \Delta - i\xi \cdot \nabla + (\rho_{\text{sc},L}^0 - \mu_{\text{per}}) \star_{\Gamma} G_1 + \frac{1}{2} |\xi|^2$$

associated with eigenvalues $\lambda_1^L(\xi) \leq \lambda_2^L(\xi) \leq \dots$. The occupation numbers $n_{k,\xi}^L$ are in the range $[0, 1]$ and such that

$$\frac{1}{L^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} n_{k,\xi}^L = Z.$$

Lastly, there exists a Fermi level $\epsilon_F^L \in \mathbb{R}$ such that $n_{k,\xi}^L = 1$ whenever $\lambda_k^L(\xi) < \epsilon_F^L$ and $n_{k,\xi}^L = 0$ whenever $\lambda_k^L(\xi) > \epsilon_F^L$. To get the energy of $\gamma_{\text{sc},L}^0$, we first compute

$$\begin{aligned} \nabla_x \gamma_{\text{sc},L}^0 &= \frac{1}{L^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} n_{k,\xi}^L (i\xi e^{i\xi \cdot x} v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)} e^{-i\xi \cdot y} + e^{i\xi \cdot x} \nabla v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)} e^{-i\xi \cdot y}), \\ \Delta_x \gamma_{\text{sc},L}^0 &= \frac{1}{L^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} n_{k,\xi}^L (-|\xi|^2 e^{i\xi \cdot x} v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)} e^{-i\xi \cdot y} + 2i\xi e^{i\xi \cdot x} \nabla v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)} e^{-i\xi \cdot y} \\ &\quad + e^{i\xi \cdot x} \Delta v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)} e^{-i\xi \cdot y}), \\ \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{\Delta}{2} \gamma_{\text{sc},L}^0 \right) &= \int_{\Lambda_L} -\frac{\Delta}{2} \gamma_{\text{sc},L}^0(x, x) dx = L^3 \int_{\Gamma} -\frac{\Delta}{2} \gamma_{\text{sc},L}^0(x, x) dx \text{ by periodicity,} \\ &= \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \frac{n_{k,\xi}^L}{2} \int_{\Gamma} (|\xi|^2 - 2i\xi \cdot \nabla - \Delta) |v_{k,\xi}^L(x)|^2 dx, \\ &= \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \frac{n_{k,\xi}^L}{2} \|(\xi - i\nabla) v_{k,\xi}^L\|_{L_{\text{per}}^2(\Gamma)}^2. \end{aligned}$$

Hence one has

$$\begin{aligned} \frac{1}{L^3} I_{\text{sc},L}^0 &= \frac{1}{L^3} \mathcal{E}_{\text{sc},L}^0(\gamma_{\text{sc},L}^0) = \frac{1}{L^3} \sum_{\xi \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \frac{n_{k,\xi}^L}{2} \|(-i\nabla + \xi)v_{k,\xi}^L\|_{L_{\text{per}}^2(\Gamma)}^2 \\ &\quad + \frac{1}{2L^3} D_{G_L}(\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}, \rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}). \end{aligned} \quad (4.73)$$

We now introduce

$$\tilde{\gamma}_{\text{sc},L}^0(x, y) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{k \geq 1} n_{k,\beta_L(\xi)}^L e^{i\xi \cdot x} v_{k,\beta_L(\xi)}^L(x) \overline{v_{k,\beta_L(\xi)}^L(y)} e^{-i\xi \cdot y} d\xi$$

where $\beta_L(\xi)$ is the unique element of $\frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*$ such that $\xi \in \beta_L(\xi) + [-\frac{2\pi\eta}{L}, \frac{2\pi(1-\eta)}{L}]^3$. $\tilde{\gamma}_{\text{sc},L}^0$ can be rewritten

$$\tilde{\gamma}_{\text{sc},L}^0(x, y) = \frac{1}{(2\pi)^3} \int_{\Gamma^*} e^{i\xi \cdot x} (\gamma_{\text{sc},L}^0)_{\beta_L(\xi)}(x, y) e^{-i\xi \cdot y} d\xi,$$

where

$$(\gamma_{\text{sc},L}^0)_{\xi}(x, y) = \sum_{k \geq 1} n_{k,\xi}^L v_{k,\xi}^L(x) \overline{v_{k,\xi}^L(y)},$$

the above kernel being well-defined since $(\gamma_{\text{sc},L}^0)_{\xi}$ is trace-class and therefore Hilbert-Schmidt. Then, it is easy to check that $\tilde{\gamma}_{\text{sc},L}^0 \in \mathcal{P}_{\text{per}}^Z$ so that $\tilde{\gamma}_{\text{sc},L}^0$ can be used as a test function for (4.8). Therefore

$$\mathcal{E}_{\text{per}}^0(\tilde{\gamma}_{\text{sc},L}^0) \geq I_{\text{per}}^0. \quad (4.74)$$

Let us remark that

$$\begin{aligned} \rho_{\tilde{\gamma}_{\text{sc},L}^0}(x) &= \tilde{\gamma}_{\text{sc},L}^0(x, x) \\ &= \frac{1}{L^3} \int_{\Gamma^*} \sum_{k \geq 1} n_{k,\beta_L(\xi)}^L \left(\frac{L}{2\pi}\right)^3 |v_{k,\beta_L(\xi)}^L(x)|^2 d\xi, \\ &= \frac{1}{L^3} \sum_{\beta_L(\xi) \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \left(\frac{L}{2\pi}\right)^3 \int_{\beta_L(\xi) + [-\frac{2\pi\eta}{L}, \frac{2\pi(1-\eta)}{L}]^3} n_{k,\beta_L(\xi)}^L |v_{k,\beta_L(\xi)}^L(x)|^2 d\xi, \\ &= \frac{1}{L^3} \sum_{\beta_L(\xi) \in \frac{2\pi}{L} \mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} n_{k,\beta_L(\xi)}^L |v_{k,\beta_L(\xi)}^L(x)|^2, \\ &= \gamma_{\text{sc},L}^0(x, x) = \rho_{\gamma_{\text{sc},L}^0}(x). \end{aligned}$$

Consequently, $\rho_{\tilde{\gamma}_{\text{sc},L}^0}$ is \mathbb{Z}^3 -periodic, and one has

$$D_{G_1}(\rho_{\tilde{\gamma}_{\text{sc},L}^0} - \mu_{\text{per}}, \rho_{\tilde{\gamma}_{\text{sc},L}^0} - \mu_{\text{per}}) = \frac{1}{L^3} D_{G_L}(\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}, \rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}). \quad (4.75)$$

Besides,

$$\begin{aligned} \frac{1}{(2\pi)^3} \int_{\Gamma^*} \text{Tr}_{L^2(\Gamma)} \left(-\frac{1}{2} \Delta (\gamma_{\text{sc},L}^0)_\xi \right) d\xi &= \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \|(-i\nabla + \xi)v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)}^2 \\ &= \frac{1}{L^3} \sum_{\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \frac{n_{k,\xi}^L}{2} \|(-i\nabla + \xi)v_{k,\xi}^L\|_{L^2_{\text{per}}(\Gamma)}^2 + R_L \quad (4.76) \end{aligned}$$

with

$$R_L = \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \left(\|(-i\nabla + \xi)v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)}^2 - \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)}^2 \right).$$

Putting (4.73)–(4.76) together, we end up with $\frac{1}{L^3} I_{\text{sc},L}^0 + R_L \geq I_{\text{per}}^0$. As

$$\begin{aligned} |R_L| &\leq \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \left(\left(\|(-i\nabla + \xi)v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} - \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right) \right. \\ &\quad \left. \left(\|(-i\nabla + \xi)v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} + \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right) \right), \\ &\leq \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \left(\left(\|(\xi - \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right) \right. \\ &\quad \left. \left(\|(-i\nabla + \beta_L(\xi) - \beta_L(\xi) + \xi)v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} + \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right) \right), \\ &\leq \frac{1}{(2\pi)^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \left(\left(\frac{2\pi}{L} \right)^3 \left(\|(-\beta_L(\xi) + \xi)v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right) \right. \\ &\quad \left. + 2\|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right), \\ &\leq \frac{2}{L^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} + \frac{1}{L^3} \int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \left(\frac{2\pi}{L} \right)^3, \\ &\leq \frac{2}{L^3} \int_{\Gamma^*} \left(\sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right)^{\frac{1}{2}} \left(\sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \right)^{\frac{1}{2}} + \left(\frac{2\pi}{L} \right)^6 \frac{Z}{2}, \\ &\leq \frac{2}{L^3} \left(\int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right)^{\frac{1}{2}} \left(\int_{\Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \right)^{\frac{1}{2}} + \left(\frac{2\pi}{L} \right)^6 \frac{Z}{2}, \\ &\leq \frac{2}{L^3} \left(\left(\frac{2\pi}{L} \right)^3 \sum_{\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \|(-i\nabla + \beta_L(\xi))v_{k,\beta_L(\xi)}^L\|_{L^2_{\text{per}}(\Gamma)} \right)^{\frac{1}{2}} \times \\ &\quad \times \left(\left(\frac{2\pi}{L} \right)^3 \sum_{\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{k \geq 1} \frac{n_{k,\beta_L(\xi)}^L}{2} \right)^{\frac{1}{2}} + \left(\frac{2\pi}{L} \right)^6 \frac{Z}{2}, \\ &\leq \left(\frac{2\pi}{L} \right)^3 \left(\frac{2I_{\text{sc},L}^0 Z}{L^3} \right)^{\frac{1}{2}} + \left(\frac{2\pi}{L} \right)^6 \frac{Z}{2}, \end{aligned}$$

we finally obtain $\liminf_{L \rightarrow +\infty} \frac{1}{L^3} I_{sc,L}^0 \geq I_{per}^0$.

Step 3 : *Convergence of the density*

A byproduct of Steps 1 and 2 is that $(\tilde{\gamma}_{sc,L}^0)_{L \in \mathbb{N} \setminus \{0\}}$ is a minimizing sequence for I_{per}^0 . The convergence results on the density $\rho_{\gamma_{sc,L}^0} = \rho_{\tilde{\gamma}_{sc,L}^0}$ immediately follows from the proof of [12, Thm 2.1].

Step 4 : *Convergence of the mean-field Hamiltonian and its spectrum*

One has $H_{sc,L}^0 - H_{per}^0 = \Phi_L$ where Φ_L solves the Poisson equation $-\Delta \Phi_L = 4\pi(\rho_{\gamma_{sc,L}^0} - \rho_{\gamma_{per}^0})$ on Γ with periodic boundary conditions. As it follows from Step 3 that $(\rho_{\gamma_{sc,L}^0} - \rho_{\gamma_{per}^0})$ converges to zero in $L^2_{per}(\Gamma)$, we obtain that Φ_L converges to zero in $H^2_{per}(\Gamma)$, hence in $L^\infty(\mathbb{R}^3)$. Consequently,

$$\|H_{sc,L}^0 - H_{per}^0\| \leq \|\Phi_L\|_{L^\infty} \rightarrow 0 \quad (4.77)$$

as $L \rightarrow \infty$. By the min-max principle, we have

$$\begin{aligned} \lambda_n(\xi) &= \inf_{V_n \subset \mathcal{V}_n} \sup_{\varphi \in V_n \setminus \{0\}} \frac{\langle (H_{per}^0)_\xi \varphi, \varphi \rangle_{L^2_{per}(\Gamma)}}{\|\varphi\|_{L^2_{per}(\Gamma)}} \\ &= \inf_{V_n \subset \mathcal{V}_n} \sup_{\varphi \in V_n \setminus \{0\}} \frac{\|(-i\nabla + \xi)\varphi\|_{L^2_{per}(\Gamma)}^2 + \int_\Gamma V_{per}^0 |\varphi|^2}{\|\varphi\|_{L^2_{per}(\Gamma)}^2}, \end{aligned}$$

for $n \geq 1$, $\xi \in \Gamma^*$, $\lambda_n(\xi)$ being the n -th Bloch eigenvalue of H_{per}^0 in the increasing order and \mathcal{V}_n the set of all subspaces of dimension n included in the form domain of H_{per}^0 , which is in fact $H^1_{per}(\Gamma)$. Similarly,

$$\begin{aligned} \lambda_n^L(\xi) &= \inf_{V_n \subset \mathcal{V}_n} \sup_{\varphi \in V_n \setminus \{0\}} \frac{\langle (H_{sc,L}^0)_\xi \varphi, \varphi \rangle_{L^2_{per}(\Gamma)}}{\|\varphi\|_{L^2_{per}(\Gamma)}} \\ &= \inf_{V_n \subset \mathcal{V}_n} \sup_{\varphi \in V_n \setminus \{0\}} \frac{\|(-i\nabla + \xi)\varphi\|_{L^2_{per}(\Gamma)}^2 + \int_\Gamma V_{sc,L}^0 |\varphi|^2}{\|\varphi\|_{L^2_{per}(\Gamma)}^2}, \end{aligned}$$

where $\lambda_n^L(\xi)$ is the n -th Bloch eigenvalue of $H_{\text{sc},L}^0$ in the increasing order and $V_{\text{sc},L}^0 = (\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) * G_1$. Then,

$$\begin{aligned} |\lambda_n^L(\xi) - \lambda_n(\xi)| &= \left| \inf_{V_n \subset \mathcal{V}} \sup_{\varphi \in V_n \setminus \{0\}} \frac{\langle ((H_{\text{sc},L}^0)_\xi - (H_{\text{per}}^0)_\xi)\varphi, \varphi \rangle_{L^2_{\text{per}}(\Gamma)}}{\|\varphi\|_{L^2_{\text{per}}(\Gamma)}} \right|, \\ &\leq \inf_{V_n \subset \mathcal{V}} \sup_{\varphi \in V_n \setminus \{0\}} \left| \frac{\langle ((H_{\text{sc},L}^0)_\xi - (H_{\text{per}}^0)_\xi)\varphi, \varphi \rangle_{L^2_{\text{per}}(\Gamma)}}{\|\varphi\|_{L^2_{\text{per}}(\Gamma)}} \right|, \\ &\leq \inf_{V_n \subset \mathcal{V}} \sup_{\varphi \in V_n \setminus \{0\}} \|H_{\text{sc},L}^0 - H_{\text{per}}^0\| \frac{\|\varphi\|_{L^2_{\text{per}}(\Gamma)}}{\|\varphi\|_{L^2_{\text{per}}(\Gamma)}}, \\ &= \|H_{\text{sc},L}^0 - H_{\text{per}}^0\|, \end{aligned}$$

which clearly implies that

$$\sup_{n \geq 1} \sup_{\xi \in \Gamma^*} |\lambda_n^L(\xi) - \lambda_n(\xi)| \leq \|\Phi_L\|_{L^\infty} \xrightarrow{L \rightarrow \infty} 0.$$

Step 5 : Uniqueness of $\gamma_{\text{sc},L}^0$ for large values of L

In the remainder of the proof, we assume that **(A1)** holds, i.e. that H_{per}^0 has a gap.

The spectrum of $H_{\text{sc},L}^0$ considered as an operator on $L^2_{\text{per}}(\Lambda_L)$ is given by

$$\sigma_{L^2_{\text{per}}(\Lambda_L)}(H_{\text{sc},L}^0) = \bigcup_{n \in \mathbb{N} \setminus \{0\}} \bigcup_{\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \lambda_n^L(\xi).$$

It follows from Step 4 that there exists some $L_0 \in \mathbb{N} \setminus \{0\}$ such that for all $L \geq L_0$, the lowest ZL^3 eigenvalues of $H_{\text{sc},L}^0$ (including multiplicities) are

$$\bigcup_{1 \leq n \leq Z} \bigcup_{\xi \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \lambda_n^L(\xi)$$

and there is a gap between the (ZL^3) -th and the $(ZL^3 + 1)$ -st eigenvalues. As a consequence, $\gamma_{\text{sc},L}^0$ is uniquely defined : it is the spectral projector associated with the lowest ZL^3 eigenvalues of $H_{\text{sc},L}^0$, considered as an operator on $L^2_{\text{per}}(\Lambda_L)$.

Step 6

Let $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. For L large enough, $\gamma_{\text{sc},L}^0 = \chi_{(-\infty, \epsilon_F]}(H_{\text{sc},L}^0) = \chi_{(-\infty, 0]}(H_{\text{sc},L}^0 - \epsilon_F)$ as operators acting on $L^2_{\text{per}}(\Lambda_L)$. This means that $\gamma_{\text{sc},L}^0$ satisfies the Euler-Lagrange equation associated with $I_{\text{sc},L,\epsilon_F}^0$, see (4.71). As the functional $\gamma \mapsto \mathcal{E}_{\text{sc},L}^0(\gamma) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(\gamma)$ is convex on $\mathcal{P}_{\text{sc},L}$, $\gamma_{\text{sc},L}^0$ is a minimizer of this functional. Its uniqueness follows as usual from the uniqueness of the minimizing density, which implies the uniqueness of the operator $H_{\text{sc},L}^0$ and from the fact that 0 is not in the spectrum of $H_{\text{sc},L}^0 - \epsilon_F$. \square

Remark 5. Notice that some of the above assertions are more precise for the supercell model than for the thermodynamic limit procedure considered in [12, Thm 2.2] (compare for instance (4.69) with (4.67)). This is because the supercell model respects the symmetry of the system, allowing in particular to have a minimizer $\gamma_{\text{sc},L}^0$ in the box of size L^3 which is periodic for the lattice \mathbb{Z}^3 . For an insulator, the uniqueness of $\gamma_{\text{sc},L}^0$ for large L and the convergence properties of *iii*) are also very interesting for computational purposes.

4.4.2 Thermodynamic limit with defect

We end this section by considering the thermodynamic limit of the supercell model with a defect. Recall that $\nu \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3) \subset \mathcal{C}$ is the nuclear density of charge of the defect. First we recall that we need to periodize this function with respect to the supercell Λ_L ; for that reason we have defined (see Section 2.4)

$$\nu_L(x) = \sum_{z \in \mathbb{Z}^3} (\mathbf{1}_{\Lambda_L} \nu)(\cdot - Lz).$$

The reduced Hartree-Fock energy functional of the supercell model with defect is then defined for $\gamma \in \mathcal{P}_{\text{sc},L}$ as

$$\mathcal{E}_{\text{sc},L}^\nu(\gamma) = \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(-\frac{1}{2} \Delta \gamma \right) + \frac{1}{2} D_{G_L}(\rho_\gamma - \mu_{\text{per}} - \nu_L, \rho_\gamma - \mu_{\text{per}} - \nu_L).$$

For $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$, we consider the following minimization problem

$$I_{\text{sc},L,\epsilon_F}^\nu = \inf \left\{ \mathcal{E}_{\text{sc},L}^\nu(\gamma) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(\gamma), \gamma \in \mathcal{P}_{\text{sc},L} \right\}. \quad (4.78)$$

We recall that γ_{per}^0 is defined in Section 4.2, that E_{κ}^ν and \bar{Q} are defined in Section 4.3.2, and that $I_{\text{sc},L,\epsilon_F}^0$ is defined in Section 4.4.1. Now we prove the

Theorem 5 (Thermodynamic limit of the supercell model with defect). *Let $Z \in \mathbb{N} \setminus \{0\}$. Assume that **(A1)** holds and fix some $\epsilon_F \in (\Sigma_Z^+, \Sigma_{Z+1}^-)$. Then one has*

$$\lim_{L \rightarrow \infty} (I_{\text{sc},L,\epsilon_F}^\nu - I_{\text{sc},L,\epsilon_F}^0) = E_{\epsilon_F}^\nu - \int_{\mathbb{R}^3} \nu \left((\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1 \right) + \frac{1}{2} D(\nu, \nu). \quad (4.79)$$

Additionally, if $\gamma_{\text{sc},L}^\nu$ denotes a minimizer for (4.78), then one has, up to extraction of a subsequence,

$$(\gamma_{\text{sc},L}^\nu - \gamma_{\text{sc},L}^0)(x, y) \rightarrow \bar{Q}(x, y)$$

weakly in $H_{\text{loc}}^1(\mathbb{R}^3 \times \mathbb{R}^3)$ and strongly in $L_{\text{loc}}^2(\mathbb{R}^3 \times \mathbb{R}^3)$, where \bar{Q} is a minimizer of (4.31), as obtained in Theorem 2. Besides,

$$\rho_{\gamma_{\text{sc},L}^\nu} - \rho_{\gamma_{\text{sc},L}^0} \rightarrow \bar{\rho}$$

weakly in $L_{\text{loc}}^2(\mathbb{R}^3)$, where $\bar{\rho}$ is the common density of all the minimizers of (4.31).

Proof. We follow the method of [25]. As in the previous proof, we denote by $\gamma_{sc,L}^0$ the minimizer of (4.68), which is unique for L large enough and is also the unique minimizer of (4.71). Let \mathcal{K}_L be the set of operators Q_L on $L^2_{\text{per}}(\Lambda_L)$ such that $\gamma_{sc,L}^0 + Q_L \in \mathcal{P}_{sc,L}$. In fact

$$\mathcal{K}_L = \left\{ Q_L \in \mathfrak{S}_1(L^2_{\text{per}}(\Lambda_L)) \mid Q_L^* = Q_L, |\nabla|Q_L|\nabla| \in \mathfrak{S}_1(L^2_{\text{per}}(\Lambda_L)), \right. \\ \left. -\gamma_{sc,L}^0 \leq Q_L \leq 1 - \gamma_{sc,L}^0 \right\}.$$

We introduce $\mathcal{E}_{sc,L,\epsilon_F}^\nu := \mathcal{E}_{sc,L}^\nu - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}$. For $Q_L \in \mathcal{K}_L$, one has

$$\begin{aligned} & \mathcal{E}_{sc,L,\epsilon_F}^\nu(\gamma_{sc,L}^0 + Q_L) - \mathcal{E}_{sc,L,\epsilon_F}^\nu(\gamma_{sc,L}^0) = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(-\frac{1}{2}(\Delta\gamma_{sc,L}^0 + \Delta Q_L) \right) \\ & + \frac{1}{2}D_{G_L}(\rho_{\gamma_{sc,L}^0} + \rho_{Q_L} - \mu_{\text{per}} - \nu_L, \rho_{\gamma_{sc,L}^0} + \rho_{Q_L} - \mu_{\text{per}} - \nu_L) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(\gamma_{sc,L}^0) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L) \\ & - \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(-\frac{1}{2}\Delta\gamma_{sc,L}^0 \right) - \frac{1}{2}D_{G_L}(\rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}, \rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}) + \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(\gamma_{sc,L}^0), \\ & = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(-\frac{1}{2}\Delta Q_L \right) + D_{G_L}(\rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}, \rho_{Q_L} - \nu_L) + \frac{1}{2}D_{G_L}(\rho_{Q_L} - \nu_L, \rho_{Q_L} - \nu_L) \\ & \quad - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L), \\ & = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(H_{sc,L}^0 Q_L) - D_{G_L}(\rho_{Q_L}, \nu_L) + \frac{1}{2}D_{G_L}(\rho_{Q_L}, \rho_{Q_L}) \\ & \quad - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L) - D_{G_L}(\nu_L, \rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}) + \frac{1}{2}D_{G_L}(\nu_L, \nu_L). \end{aligned}$$

Note that in the above expression, $H_{sc,L}^0$ is considered as an operator on $L^2_{\text{per}}(\Lambda_L)$. Using Theorem 4, we can write that

$$\begin{aligned} & \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(H_{sc,L}^0 Q_L) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L) \\ & = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((H_{sc,L}^0 - \epsilon_F)(1 - \gamma_{sc,L}^0 + \gamma_{sc,L}^0)Q_L(1 - \gamma_{sc,L}^0 + \gamma_{sc,L}^0)), \\ & = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((H_{sc,L}^0 - \epsilon_F)(Q_L^{++L} + Q_L^{--L})), \\ & = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(|H_{sc,L}^0 - \epsilon_F|(Q_L^{++L} - Q_L^{--L})), \end{aligned}$$

since $\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((H_{sc,L}^0 - \epsilon_F)Q_L^{++L}) \geq 0$, $\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((H_{sc,L}^0 - \epsilon_F)Q_L^{--L}) \leq 0$ and $\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((H_{sc,L}^0 - \epsilon_F)(1 - \gamma_{sc,L}^0)Q_L\gamma_{sc,L}^0) = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((H_{sc,L}^0 - \epsilon_F)\gamma_{sc,L}^0 Q_L(1 - \gamma_{sc,L}^0)) = 0$, where we have set

$$Q_L^{++L} = (1 - \gamma_{sc,L}^0)Q_L(1 - \gamma_{sc,L}^0) \quad \text{and} \quad Q_L^{--L} = \gamma_{sc,L}^0 Q_L \gamma_{sc,L}^0.$$

Then the equality on the energies can be rewritten, for L large enough, as

$$\begin{aligned} & \mathcal{E}_{sc,L,\epsilon_F}^\nu(\gamma_{sc,L}^0 + Q_L) - \mathcal{E}_{sc,L,\epsilon_F}^\nu(\gamma_{sc,L}^0) = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(|H_{sc,L}^0 - \epsilon_F|(Q_L^{++L} - Q_L^{--L})) \\ & \quad + \frac{1}{2}D_{G_L}(\rho_{Q_L} - \nu_L, \rho_{Q_L} - \nu_L) - D_{G_L}(\nu_L, \rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}). \quad (4.80) \end{aligned}$$

It follows from (4.80) that

$$\begin{aligned} I_{sc,L,\epsilon_F}^\nu - I_{sc,L,\epsilon_F}^0 & = \inf \left\{ E_{sc,L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L), Q_L \in \mathcal{K}_L \right\} \\ & \quad - D_{G_L}(\nu_L, \rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}) + \frac{1}{2}D_{G_L}(\nu_L, \nu_L). \end{aligned}$$

where

$$\begin{aligned} E_{\text{sc},L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(Q_L) &:= -D_{G_L}(\rho_{Q_L}, \nu_L) + \frac{1}{2} D_{G_L}(\rho_{Q_L}, \rho_{Q_L}) \\ &\quad + \text{Tr}_{L_{\text{per}}^2(\Lambda_L)} \left(|H_{\text{sc},L}^0 - \epsilon_F|^{1/2} (Q_L^{++L} - Q_L^{--L}) |H_{\text{sc},L}^0 - \epsilon_F|^{1/2} \right). \end{aligned}$$

First, using ν being in $L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ and the convergence of $\Phi_L = (\rho_{\gamma_{\text{sc},L}^0} - \rho_{\gamma_{\text{per}}^0}) \star_\Gamma G_1$ to zero in L^∞ (see Step 4 of the proof of Theorem 4 in Section 4.4.1), we obtain

$$-D_{G_L}(\nu_L, \rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) + \frac{1}{2} D_{G_L}(\nu_L, \nu_L) \longrightarrow - \int_{\mathbb{R}^3} \nu \left((\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}) \star_\Gamma G_1 \right) + \frac{1}{2} D(\nu, \nu).$$

Indeed, we have

$$\begin{aligned} D_{G_L}(\nu_L, \rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) &= \int_{\Lambda_L} (G_L *_{\Lambda_L} \nu_L)(x) \nu_L(x) dx, \\ &= \sum_{k \in \frac{2\pi}{L} \mathbb{Z}^3 \setminus \{0\}} c_k(G_L) |c_k(\nu_L)|^2, \end{aligned}$$

where $c_k(f)$ denotes the k -th Fourier coefficient of f . We already know that $c_k(G_L) = \frac{4\pi}{|k|^2}$ for $k \neq 0$ and we compute

$$\begin{aligned} c_k(\nu_L) &= \frac{1}{L^{\frac{3}{2}}} \int_{\Lambda_L} \sum_{z \in \mathbb{Z}^3} \nu(x - Lz) e^{-ik \cdot x} dx, \\ &= \frac{1}{L^{\frac{3}{2}}} \sum_{z \in \mathbb{Z}^3} \int_{\Lambda_L} \nu(x - Lz) e^{-ik \cdot (x - Lz)} dx, \quad \text{by periodicity, for } k \in \frac{2\pi}{L} \mathbb{Z}^3, \\ &= \frac{1}{L^{\frac{3}{2}}} \sum_{z \in \mathbb{Z}^3} \int_{-Lz + \Lambda_L} \nu(y) e^{-ik \cdot y} dy, \\ &= \frac{1}{L^{\frac{3}{2}}} \int_{\mathbb{R}^3} \nu(y) e^{-ik \cdot y} dy, \\ &= \left(\frac{2\pi}{L} \right)^{\frac{3}{2}} \hat{\nu}(k), \end{aligned}$$

hence

$$D_{G_L}(\nu_L, \nu_L) = \sum_{k \in \frac{2\pi}{L} \mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} \left(\frac{2\pi}{L} \right) |\hat{\nu}(k)|^2,$$

which converges to $D(\nu, \nu) = 4\pi \int_{\mathbb{R}^3} \frac{|\hat{\nu}(k)|^2}{|k|^2}$ when $L \rightarrow \infty$ by convergence of the Riemann sums.

Thanks to the proof of Theorem 4 in Section 4.4.1, we have that $G_1 \star_\Gamma (\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) = G_L *_{\Lambda_L} (\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}})$ since $\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}$ is \mathbb{Z}^3 -periodic and its integral over

Γ is equal to zero. Then,

$$\begin{aligned} D_{G_L}(\nu_L, \rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}) &= \int_{\Lambda_L} (G_1 *_{\Gamma} (\rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}))(x) \nu_L(x) dx, \\ \left| D_{G_L}(\nu_L, \rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}) \right| &\leq \left| \int_{\Lambda_L} (G_1 *_{\Gamma} (\rho_{\gamma_{sc,L}^0} - \rho_{\gamma_{\text{per}}^0}))(x) \nu_L(x) dx \right| \\ &\quad + \left| \int_{\Lambda_L} (G_1 *_{\Gamma} (\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}}))(x) \nu_L(x) dx \right|, \end{aligned}$$

the first term converges to zero thanks to the convergence of $\Phi_L = (\rho_{\gamma_{sc,L}^0} - \rho_{\gamma_{\text{per}}^0}) *_{\Gamma} G_1$ to zero in L^∞ and the fact that ν is in $L^1(\mathbb{R}^3)$, yielding that the sequence $(\|\nu_L\|_{L^1(\Lambda_L)})$ is bounded. Then, similarly to what we did above, we have the convergence of the Fourier series of the second term to $4\pi \int_{\mathbb{R}^3} (\rho_{\gamma_{\text{per}}^0} - \mu_{\text{per}})(k) \hat{\nu}(k) / |k|^2 dk$.

Our goal is to prove that

$$\lim_{L \rightarrow \infty} E_{\epsilon_F, L}^\nu = E_{\epsilon_F}^\nu \quad (4.81)$$

where

$$E_{\epsilon_F, L}^\nu = \inf \left\{ E_{sc, L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L), Q_L \in \mathcal{K}_L \right\}. \quad (4.82)$$

Step 1 : Preliminaries

In the proof of (4.81), we shall need several times to compare states living in $L^2_{\text{per}}(\Lambda_L)$ with states living in $L^2(\mathbb{R}^3)$. To this end, we introduce the map

$$\begin{aligned} i_L : L^2(\mathbb{R}^3) &\rightarrow L^2_{\text{per}}(\Lambda_L) \\ \varphi &\mapsto \sum_{z \in \mathbb{Z}^3} (\mathbf{1}_{\Lambda_L} \varphi)(\cdot - Lz). \end{aligned}$$

Notice that $(i_L)^* : L^2_{\text{per}}(\Lambda_L) \rightarrow L^2(\mathbb{R}^3)$ is the operator which to any periodic function $\psi \in L^2_{\text{per}}(\Lambda_L)$ associates the function $\mathbf{1}_{\Lambda_L} \psi \in L^2(\mathbb{R}^3)$. Remark that $i_L(i_L)^* = Id_{L^2_{\text{per}}(\Lambda_L)}$ whereas $(i_L)^*i_L = \mathbf{1}_{\Lambda_L}$. Hence i_L defines an isometry from $L^2(\Lambda_L)$ to $L^2_{\text{per}}(\Lambda_L)$. The equality $(i_L)^*i_L\varphi = \varphi$ is only true when $\varphi \in L^2(\mathbb{R}^3)$ has its support in Λ_L . When $\varphi \in H^1(\mathbb{R}^3)$ satisfies $\text{Supp}(\varphi) \subset \Lambda_L$, then one also has $\partial_{x_i} i_L(\varphi) = i_L(\partial_{x_i} \varphi)$.

Notice in addition that if $A \in \mathfrak{S}_1(L^2_{\text{per}}(\Lambda_L))$, then $(i_L)^*Ai_L \in \mathfrak{S}_1(L^2(\Lambda_L)) \subseteq \mathfrak{S}_1(L^2(\mathbb{R}^3))$ and

$$\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(A) = \text{Tr}_{L^2(\mathbb{R}^3)}((i_L)^*Ai_L).$$

Similarly if $A \in \mathfrak{S}_p(L^2_{\text{per}}(\Lambda_L))$,

$$\|A\|_{\mathfrak{S}_p(L^2_{\text{per}}(\Lambda_L))} = \|(i_L)^*Ai_L\|_{\mathfrak{S}_p(L^2(\mathbb{R}^3))}. \quad (4.83)$$

Finally, we shall use that for any \mathbb{Z}^3 -periodic bounded function f , $i_L f = f i_L$, where we use the same notation f to denote the multiplication operator by the function f acting either on $L^2_{\text{per}}(\Lambda_L)$ or on $L^2(\mathbb{R}^3)$. Similarly, the operator $H_{sc, L}^0 =$

$-\Delta/2 + (\rho_{\gamma_{sc,L}^0} - \mu_{\text{per}}) \star_{\Gamma} G_1$ can be seen as acting on $L^2_{\text{per}}(\Lambda_L)$ or on $L^2(\mathbb{R}^3)$ and we use the same notation in the two cases. Then we have for any $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^3)$ satisfying $\text{Supp}(\varphi) \subseteq \Lambda_L$

$$H_{sc,L}^0 i_L \varphi = i_L H_{sc,L}^0 \varphi. \quad (4.84)$$

Notice that one can also define $-\Delta \nabla$ on $L^2(\mathbb{R}^3)$ or on $L^2_{\text{per}}(\Lambda_L)$ and we shall adopt the same notation for these two operators. We gather some useful limits in the following

Lemma 8. *Let be $\psi \in L^2(\mathbb{R}^3)$ and $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^3)$. Then we have as $L \rightarrow \infty$*

1. $(i_L)^* \gamma_{sc,L}^0 i_L \psi \rightarrow \gamma_{\text{per}}^0 \psi$ in $L^2(\mathbb{R}^3)$;
2. $(i_L)^* H_{sc,L}^0 \gamma_{sc,L}^0 i_L \varphi \rightarrow H_{\text{per}}^0 \gamma_{\text{per}}^0 \varphi$ in $L^2(\mathbb{R}^3)$;
3. $(i_L)^* \Delta \gamma_{sc,L}^0 i_L \varphi \rightarrow \Delta \gamma_{\text{per}}^0 \varphi$ in $L^2(\mathbb{R}^3)$;
4. $(i_L)^* (1 + |\nabla|) i_L \varphi \rightarrow (1 + |\nabla|) \varphi$ in $L^2(\mathbb{R}^3)$;
5. $(i_L)^* |H_{sc,L}^0 - \epsilon_F|^{1/2} i_L \varphi \rightarrow |H_{\text{per}}^0 - \epsilon_F|^{1/2} \varphi$ in $L^2(\mathbb{R}^3)$ for any fixed ϵ_F in the gap $(\Sigma_Z^+, \Sigma_{Z+1})$.

Proof. The operator $(i_L)^* \gamma_{sc,L}^0 i_L$ being uniformly bounded with respect to L , it suffices to prove the first assertion for a dense subset of $L^2(\mathbb{R}^3)$. Hence we may assume that $\psi = \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^3)$.

Let K be a compact set in the resolvent set of H_{per}^0 . We are going to prove that

$$\lim_{L \rightarrow \infty} (i_L)^* (z - H_{sc,L}^0)^{-1} i_L \varphi \rightarrow_{L \rightarrow \infty} (z - H_{\text{per}}^0)^{-1} \varphi \quad (4.85)$$

in $L^2(\mathbb{R}^3)$, uniformly for $z \in K$. To this end, we first notice that by Theorem 4, K is contained in the resolvent set of $H_{sc,L}^0$ for L large enough. More precisely, there exists $C(K) > 0$ such that for L large enough $\|(z - H_{sc,L}^0)^{-1}\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \leq C(K)$. We also know from Theorem 4 that for all $\epsilon > 0$, there exists $L_0 > 0$ such that for all $L \geq L_0$, $\|H_{\text{per}}^0 - H_{sc,L}^0\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \leq \frac{\epsilon}{C(K)}$. Then, using that $(z - A)^{-1} - (z - B)^{-1} = (z - A)^{-1} (1 - (1 - (z - A)^{-1} (B - A))^{-1})$, we obtain for L large enough

$$\begin{aligned} & \|(z - H_{sc,L}^0)^{-1} - (z - H_{\text{per}}^0)^{-1}\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \\ &= \|(z - H_{sc,L}^0)^{-1} (1 - (1 - (z - H_{sc,L}^0)^{-1} (H_{\text{per}}^0 - H_{sc,L}^0))^{-1})\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \\ &\leq \|(z - H_{sc,L}^0)^{-1}\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \|1 - (1 - (z - H_{sc,L}^0)^{-1} (H_{\text{per}}^0 - H_{sc,L}^0))^{-1}\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \\ &\leq C(K) \sum_{k=1}^{\infty} \|((z - H_{sc,L}^0)^{-1} (H_{\text{per}}^0 - H_{sc,L}^0))^k\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \\ &\leq C(K) \frac{\epsilon}{1 - \epsilon}, \end{aligned}$$

thus

$$\|(z - H_{sc,L}^0)^{-1} - (z - H_{\text{per}}^0)^{-1}\|_{\mathcal{B}(L^2_{\text{per}}(\Lambda_L))} \rightarrow 0.$$

Hence it suffices to prove (4.85) with $H_{\text{sc},L}^0$ replaced by H_{per}^0 (seen as an operator acting on $L^2_{\text{per}}(\Lambda_L)$). Assuming L large enough for $\text{Supp}(\varphi) \subset \Lambda_L$,

$$\begin{aligned}
& (i_L)^*(z - H_{\text{per}}^0)^{-1}i_L\varphi(x) \\
&= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \frac{L^{-3}}{z - \lambda_n(k)} \left(\int_{\mathbb{R}^3} \overline{e_n(k, y)} \sum_{z \in \mathbb{Z}^3} (\mathbf{1}_{\Lambda_L}\varphi)(y - Lz) dy \right) \mathbf{1}_{\Lambda_L}(x) e_n(k, x), \\
&= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \frac{L^{-3}}{z - \lambda_n(k)} \sum_{z \in \mathbb{Z}^3} \left(\int_{\Lambda_L} \overline{e_n(k, y - Lz)} \varphi(y - Lz) dy \right) \mathbf{1}_{\Lambda_L}(x) e_n(k, x), \\
&= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \frac{L^{-3}}{z - \lambda_n(k)} \sum_{z \in \mathbb{Z}^3} \left(\int_{-Lz + \Lambda_L} \overline{e_n(k, z')} \varphi(z') dz' \right) \mathbf{1}_{\Lambda_L}(x) e_n(k, x), \\
&= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \frac{L^{-3}}{z - \lambda_n(k)} \left(\int_{\mathbb{R}^3} \overline{e_n(k, \cdot)} \varphi \right) \mathbf{1}_{\Lambda_L}(x) e_n(k, x),
\end{aligned}$$

$$\|(i_L)^*(z - H_{\text{per}}^0)^{-1}i_L\varphi\|_{L^2(\mathbb{R}^3)}^2 = \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \frac{L^{-3}}{|z - \lambda_n(k)|^2} \left| \int_{\mathbb{R}^3} \overline{e_n(k, \cdot)} \varphi \right|^2.$$

Now, taking the scalar product against a function $\psi \in \mathcal{C}_0^\infty(\mathbb{R}^3)$, we check that $(i_L)^*(z - H_{\text{per}}^0)^{-1}i_L\varphi \rightarrow (z - H_{\text{per}}^0)^{-1}\varphi$ in $\mathcal{D}'(\mathbb{R}^3)$: for L large enough so that $\text{Supp}(\psi) \subset \Lambda_L$,

$$\begin{aligned}
& \langle (i_L)^*(z - H_{\text{per}}^0)^{-1}i_L\varphi, \psi \rangle_{L^2_{\text{per}}(\Lambda_L)} \\
&= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \frac{L^{-3}}{z - \lambda_n(k)} \left(\int_{\mathbb{R}^3} \overline{e_n(k, \cdot)} \varphi \right) \int_{\Lambda_L} e_n(k, x) \psi(x) dx, \\
&= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \cap \Gamma^*} \sum_{n \geq 1} \left(\frac{2\pi}{L} \right)^3 \frac{(2\pi)^{-3}}{z - \lambda_n(k)} \left(\int_{\mathbb{R}^3} \overline{e_n(k, \cdot)} \varphi \right) \int_{\mathbb{R}^3} e_n(k, x) \psi(x) dx, \\
&= \frac{1}{|\Gamma^*|} \int_{\Gamma^*} d\xi \sum_{n \geq 1} \frac{1}{z - \lambda_n(\xi)} \left(\int_{\mathbb{R}^3} \overline{e_n(\xi, \cdot)} \varphi \right) \int_{\mathbb{R}^3} e_n(\xi, x) \psi(x) dx, \\
&= \langle (z - H_{\text{per}}^0)^{-1}\varphi, \psi \rangle_{L^2_\xi(\Gamma)},
\end{aligned}$$

by convergence of the Riemann sums. Then, similarly, $\|(i_L)^*(z - H_{\text{per}}^0)^{-1}i_L\varphi\|_{L^2(\mathbb{R}^3)} \rightarrow \|(z - H_{\text{per}}^0)^{-1}\varphi\|_{L^2(\mathbb{R}^3)}$, also because it is a Riemann sum over k . This yields the strong convergence of $(i_L)^*(z - H_{\text{per}}^0)^{-1}i_L\varphi$ to $(z - H_{\text{per}}^0)^{-1}\varphi$ in $L^2(\mathbb{R}^3)$.

For the proof of (1), it then suffices to choose a curve \mathcal{C} around the first Z bands of H_{per}^0 and use the above convergence of the resolvent in the Cauchy formula. Assertion (2) is an easy consequence of (1) and (4.84). Indeed, by Theorem 4, we know that $\lim_{L \rightarrow \infty} \|H_{\text{sc},L}^0 - H_{\text{per}}^0\|_{\mathcal{B}(L^2(\mathbb{R}^3))} = 0$. Since $(i_L)^*\gamma_{\text{sc},L}^0 i_L$ is bounded, this implies that for L large enough such that $\text{Supp}(\varphi) \subseteq \Lambda_L$

$$\|(i_L)^*\gamma_{\text{sc},L}^0 (H_{\text{sc},L}^0 i_L - i_L H_{\text{per}}^0) \varphi\|_{L^2(\mathbb{R}^3)} = \|(i_L)^*\gamma_{\text{sc},L}^0 i_L (H_{\text{sc},L}^0 - H_{\text{per}}^0) \varphi\|_{L^2(\mathbb{R}^3)} \rightarrow 0$$

where we have used (4.84). Then we notice that $H_{\text{per}}^0 \varphi \in L^2(\mathbb{R}^3)$. Hence (1) implies that $\lim_{L \rightarrow \infty} \|((i_L)^* \gamma_{\text{sc},L}^0 i_L - \gamma_{\text{per}}^0) H_{\text{per}}^0 \varphi\|_{L^2(\mathbb{R}^3)} = 0$. The argument is exactly the same for the third assertion (3). Assertion (5) can be proved in the same way, using (4.84) and the integral representation of the square root (4.57).

Finally, it remains to prove that (4) is true, which is done by computing explicitly, for L large enough so that $\text{Supp}(\varphi) \subseteq \Lambda_L$,

$$(i_L)^*(1 + |\nabla|)i_L \varphi = \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{(2\pi)^{3/2}}{L^3} (1 + |k|) \widehat{\varphi}(k) e^{ik \cdot x} \mathbb{1}_{\Lambda_L}(x). \quad (4.86)$$

To prove (4.86), we begin by expanding the periodic function $i_L \varphi$ in Fourier series

$$(i_L \varphi)(x) = \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{1}{L^{3/2}} c_k(i_L \varphi) e^{ik \cdot x},$$

with

$$\begin{aligned} c_k(i_L \varphi) &= \frac{1}{L^{3/2}} \int_{\Lambda_L} \sum_{z \in \mathbb{Z}^3} (\mathbb{1}_{\Lambda_L} \varphi)(x - Lz) e^{-ik \cdot x} dx, \\ &= \frac{1}{L^{3/2}} \sum_{z \in \mathbb{Z}^3} \int_{\Lambda_L} \varphi(x - Lz) e^{-ik \cdot (x - Lz)} dx, \\ &= \frac{1}{L^{3/2}} \int_{\mathbb{R}^3} \varphi(x) e^{-ik \cdot x} dx, \\ &= \left(\frac{2\pi}{L} \right)^{3/2} \widehat{\varphi}(k), \end{aligned}$$

where $\widehat{\varphi}(k)$ is the Fourier transform at point k of the function $\varphi \in L^2(\mathbb{R}^3)$. Then, multiplying by $(1 + |\nabla|)$ corresponds to multiplying by $(1 + |k|)$ the k -th Fourier mode. Finally (i_L^*) consists of multiplying the $L^2_{\text{per}}(\Lambda_L)$ function by $\mathbb{1}_{\Lambda_L}$, hence (4.86). Consequently,

$$\|(i_L)^*(1 + |\nabla|)i_L \varphi\|_{L^2(\mathbb{R}^3)}^2 = \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{(2\pi)^3}{L^3} |(1 + |k|) \widehat{\varphi}(k)|^2 \xrightarrow{L \rightarrow \infty} \|(1 + |\nabla|)\varphi\|_{L^2(\mathbb{R}^3)}^2.$$

The strong convergence is obtained as above by the convergence of the Riemann sums when $L \rightarrow \infty$.

Lemma 9. *Let $V \in \mathcal{C}_0^\infty(\mathbb{R}^3)$. We have as $L \rightarrow \infty$*

$$(i_L)^*(1 - \Delta)^{-1} i_L(V) i_L \rightarrow (1 - \Delta)^{-1} V, \quad (4.87)$$

$$(i_L)^*(1 + |\nabla|)^{-1} i_L(V) (1 + |\nabla|)^{-1} i_L \rightarrow (1 + |\nabla|)^{-1} V (1 + |\nabla|)^{-1} \quad (4.88)$$

strongly in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$.

Proof. For L large enough so that $\text{Supp}(V) \subset \Lambda_L$, we have

$$\begin{aligned} \|(i_L)^*(1 - \Delta)^{-1}i_L(V)i_L\|_{\mathfrak{S}_2(L^2(\mathbb{R}^3))} &= \|(1 - \Delta)^{-1}i_L(V)\|_{\mathfrak{S}_2(L^2_{\text{per}}(\Lambda_L))} \\ &= \|i_L(V)\|_{L^2_{\text{per}}(\Lambda_L)} \|g_0\|_{L^2_{\text{per}}(\Lambda_L)} \\ &= \|V\|_{L^2(\mathbb{R}^3)} \|g_0\|_{L^2_{\text{per}}(\Lambda_L)}, \end{aligned}$$

where $g_0(x - y)$ denotes the kernel of the operator $(1 - \Delta)^{-1}$, considered as an operator on $L^2_{\text{per}}(\Lambda_L)$. As the Fourier space representation of this operator reads

$$\sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{1}{1 + |k|^2} |e_k\rangle \langle e_k|,$$

with $e_k(x) = \frac{1}{L^{3/2}} e^{ik \cdot x}$, we obtain

$$g_0(x - y) = \frac{1}{(2\pi)^3} \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \left(\frac{2\pi}{L}\right)^3 \frac{1}{1 + |k|^2} e^{ik \cdot (x-y)},$$

and therefore $\|g_0\|_{L^2_{\text{per}}(\Lambda_L)} = (2\pi)^{-3/2} \left(\sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{(2\pi/L)^3}{(1 + |k|^2)^2}\right)^{1/2}$. Now, since

$$\lim_{L \rightarrow \infty} \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{(2\pi/L)^3}{(1 + |k|^2)^2} = \int_{\mathbb{R}^3} |g(p)|^2 dp, \quad g(p) = (1 + |p|^2)^{-1}, \quad (4.89)$$

$$\begin{aligned} \lim_{L \rightarrow \infty} \|(i_L)^*(1 - \Delta)^{-1}i_L(V)i_L\|_{\mathfrak{S}_2(L^2(\mathbb{R}^3))} &= \frac{\|V\|_{L^2(\mathbb{R}^3)} \|g\|_{L^2(\mathbb{R}^3)}}{(2\pi)^{3/2}} \\ &= \|(1 - \Delta)^{-1}V\|_{\mathfrak{S}_2(L^2(\mathbb{R}^3))}. \end{aligned} \quad (4.90)$$

Arguing as in the proof of the fourth assertion of Lemma 8, we can prove that the convergence (4.87) holds in the strong operator sense. It then follows from (4.90) that the convergence holds strongly in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$.

The argument is the same for (4.88), noticing that

$$\begin{aligned} &\|(i_L)^*(1 + |\nabla|)^{-1}i_L(V)(1 + |\nabla|)^{-1}i_L\|_{\mathfrak{S}_2(L^2(\mathbb{R}^3))}^2 \\ &= \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left((1 + |\nabla|)^{-2}i_L(V)(1 + |\nabla|)^{-2}i_L(V) \right) \\ &= (2\pi)^{-3/2} \iint_{(\Lambda_L)^2} |h_L(x - y)|^2 V(x)V(y) dx dy \\ &\rightarrow_{L \rightarrow \infty} \|(1 + |\nabla|)^{-1}V(1 + |\nabla|)^{-1}\|_{\mathfrak{S}_2(L^2(\mathbb{R}^3))}^2 \end{aligned}$$

where we have used that

$$h_L(x) := \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \frac{(2\pi)^{3/2}}{L^3(1 + |k|^2)} e^{ik \cdot x}$$

converges to the Fourier inverse $\mathcal{F}^{-1}(h)$ of $h(p) = (1 + |p|)^{-2}$, strongly in $L^2_{\text{loc}}(\mathbb{R}^3)$.

Step 2 : Upper bound

We prove here that $\limsup_{L \rightarrow \infty} E_{\epsilon_F, L}^\nu \leq E_{\epsilon_F}^\nu$. Let $\epsilon > 0$. Using Lemma 14, Proposition 5, Corollary 4, and the notation therein, one can find, like in the proof of Theorem 3, a finite rank operator $Q \in \mathcal{K}_r$ such that

$$E_{\epsilon_F}^\nu \leq \mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}_0(Q) \leq E_{\epsilon_F}^\nu + \epsilon, \quad (4.91)$$

of the form

$$\begin{aligned} Q = & \sum_{m=-M}^{-1} |v_m\rangle\langle v_m| - \sum_{n=-N}^{-1} |u_n\rangle\langle u_n| + \sum_{i=0}^k \frac{\lambda_i^2}{1 + \lambda_i^2} (|v_i\rangle\langle v_i| - |u_i\rangle\langle u_i|) \\ & + \sum_{i=0}^k \frac{\lambda_i}{1 + \lambda_i^2} (|u_i\rangle\langle v_i| + |v_i\rangle\langle u_i|) + \delta' \quad \text{with} \quad \delta' = \sum_{j=1}^J n_j |w_j\rangle\langle w_j|. \end{aligned} \quad (4.92)$$

Let $0 < \eta \ll 1$. It is possible to choose a family of orthonormal functions $u_n^\eta, v_m^\eta, w_j^\eta$ in $C_0^\infty(\mathbb{R}^3)$ such that

$$\|u_n^\eta - u_n\|_{H^2} \leq \eta, \quad \|v_m^\eta - v_m\|_{H^2} \leq \eta, \quad \|w_j^\eta - w_j\|_{H^2} \leq \eta. \quad (4.93)$$

for all $n = -N \dots k, m = -M \dots k$ and $j = 1 \dots J$. Let us define the Gram matrices

$$(S_-^\eta)_{i,j} := \langle \gamma_{\text{per}}^0 u_i^\eta, u_j^\eta \rangle, \quad (S_+^\eta)_{i,j} := \langle (1 - \gamma_{\text{per}}^0) v_i^\eta, v_j^\eta \rangle$$

which, by (4.93) satisfy $S_+^\eta = Id_{M+k+1} + o(1)_{\eta \rightarrow 0}$ and $S_-^\eta = Id_{N+k+1} + o(1)_{\eta \rightarrow 0}$. We also introduce the orthogonal projector Π^η on $\text{Span}\{\gamma_{\text{per}}^0 u_n^\eta, (1 - \gamma_{\text{per}}^0) v_m^\eta\}$ and define $(S_w^\eta)_{i,j} := \langle (1 - \Pi^\eta) w_i^\eta, w_j^\eta \rangle$. Clearly $S_w^\eta = Id_J + o(1)_{\eta \rightarrow 0}$.

Now we introduce a new orthonormal system in $L_{\text{per}}^2(\Lambda_L)$

$$u_{i,L}^\eta := \sum_{n=-N}^k (S_{-,L})_{i,n}^{-1/2} \gamma_{\text{sc},L}^0 i_L u_n^\eta, \quad v_{i,L}^\eta := \sum_{m=-M}^k (S_{+,L})_{i,m}^{-1/2} (1 - \gamma_{\text{sc},L}^0) i_L v_m^\eta,$$

$$(S_{-,L})_{i,j} = \langle \gamma_{\text{sc},L}^0 i_L u_i^\eta, i_L u_j^\eta \rangle_{L_{\text{per}}^2(\Lambda_L)}, \quad (S_{+,L})_{i,j} = \langle (1 - \gamma_{\text{sc},L}^0) i_L v_i^\eta, i_L v_j^\eta \rangle_{L_{\text{per}}^2(\Lambda_L)}.$$

Notice that by the first assertion of Lemma 8, $\lim_{L \rightarrow \infty} S_{\pm,L} = S_{\pm}^\eta$. Finally, we introduce the projector Π_L on $\text{Span}(u_{n,L}^\eta, v_{m,L}^\eta)$ and define

$$w_{j,L}^\eta := \sum_{\ell=1}^J (S_{w,L}^{-1/2})_{j,\ell} (1 - \Pi_L) i_L w_\ell^\eta, \quad (S_{w,L})_{i,j} = \langle (1 - \Pi_L) i_L w_i^\eta, i_L w_j^\eta \rangle.$$

We now define a state in \mathcal{K}_L by

$$\begin{aligned} Q_L^\eta = & \sum_{m=-M}^{-1} |v_{m,L}^\eta\rangle\langle v_{m,L}^\eta| - \sum_{n=-N}^{-1} |u_{n,L}^\eta\rangle\langle u_{n,L}^\eta| + \sum_{i=0}^k \frac{\lambda_i^2}{1 + \lambda_i^2} (|v_{i,L}^\eta\rangle\langle v_{i,L}^\eta| - |u_{i,L}^\eta\rangle\langle u_{i,L}^\eta|) \\ & + \sum_{i=0}^k \frac{\lambda_i}{1 + \lambda_i^2} (|u_{i,L}^\eta\rangle\langle v_{i,L}^\eta| + |v_{i,L}^\eta\rangle\langle u_{i,L}^\eta|) + \sum_{j=1}^J n_j |w_{j,L}^\eta\rangle\langle w_{j,L}^\eta|. \end{aligned}$$

By Lemma 8, we have

$$(i_L)^* u_{i,L}^\eta \rightarrow_{L \rightarrow \infty} \tilde{u}_i^\eta, \quad (i_L)^* v_{i,L}^\eta \rightarrow_{L \rightarrow \infty} \tilde{v}_i^\eta \quad \text{and} \quad (i_L)^* w_{j,L}^\eta \rightarrow_{L \rightarrow \infty} \tilde{w}_j^\eta \quad (4.94)$$

in $L^2(\mathbb{R}^3) \cap L^\infty(\mathbb{R}^3)$, where the limits are defined by

$$\begin{aligned} \tilde{u}_i^\eta &:= \sum_{n=-N}^k (S_-^\eta)^{-1/2} \gamma_{\text{per}}^0 u_n^\eta, & \tilde{v}_i^\eta &:= \sum_{m=-M}^k (S_+^\eta)^{-1/2} (1 - \gamma_{\text{per}}^0) v_m^\eta, \\ \tilde{w}_j^\eta &:= \sum_{\ell=1}^J (S_w^\eta)^{-1/2} (1 - \Pi^\eta) w_\ell^\eta. \end{aligned}$$

By Lemma 8, we know that for any fixed $\varphi, \psi \in \mathcal{C}_0^\infty(\mathbb{R}^3)$,

$$\lim_{L \rightarrow \infty} \langle i_L^* (H_{\text{sc},L}^0 - \epsilon_F) \gamma_{\text{sc},L}^0 i_L \varphi, (i_L)^* \gamma_{\text{sc},L}^0 i_L \psi \rangle = \langle (H_{\text{per}}^0 - \epsilon_F) \gamma_{\text{per}}^0 \varphi, \gamma_{\text{per}}^0 \psi \rangle.$$

Hence, inserting the definition of Q_L^η in the kinetic energy and using the convergence of the Gram matrices, we obtain

$$\lim_{L \rightarrow \infty} \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} ((H_{\text{sc},L}^0 - \epsilon_F) Q_L^\eta) = \text{Tr}_{L^2(\mathbb{R}^3)} ((H_{\text{per}}^0 - \epsilon_F) \tilde{Q}^\eta)$$

where \tilde{Q}^η is defined similarly as Q_L^η but with the functions $(\tilde{u}_n^\eta, \tilde{v}_m^\eta, \tilde{w}_j^\eta)$ instead of $(u_{n,L}^\eta, v_{m,L}^\eta, w_{j,L}^\eta)$. Indeed, let us show it for $v_{m,L}^\eta$; the proof being very similar for the other basis functions and the sum being finite, the limit is then straightforward.

$$\begin{aligned} \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} ((H_{\text{sc},L}^0 - \epsilon_F) |v_{m,L}^\eta\rangle \langle v_{m,L}^\eta|) &= \langle (H_{\text{sc},L}^0 - \epsilon_F) v_{m,L}^\eta, v_{m,L}^\eta \rangle_{L^2_{\text{per}}(\Lambda_L)}, \\ &= \langle (H_{\text{sc},L}^0 - \epsilon_F) v_{m,L}^\eta, i_L i_L^* v_{m,L}^\eta \rangle_{L^2_{\text{per}}(\Lambda_L)}, \\ &= \left\langle \sum_{i=-M}^k i_L^* (H_{\text{sc},L}^0 - \epsilon_F) (S_{+,L}^\eta)^{-\frac{1}{2}} (1 - \gamma_{\text{sc},L}^0) i_L v_i^\eta, \sum_{j=-M}^k i_L^* (S_{+,L}^\eta)^{-\frac{1}{2}} (1 - \gamma_{\text{sc},L}^0) i_L v_j^\eta \right\rangle_{L^2_{\text{per}}(\Lambda_L)}, \\ &\rightarrow_{L \rightarrow \infty} \left\langle \sum_{i=-M}^k (H_{\text{per}}^0 - \epsilon_F) (S_+^\eta)^{-\frac{1}{2}} (1 - \gamma_{\text{per}}^0) v_i^\eta, \sum_{j=-M}^k (S_+^\eta)^{-\frac{1}{2}} (1 - \gamma_{\text{per}}^0) v_j^\eta \right\rangle_{L^2(\mathbb{R}^3)}, \\ &= \langle (H_{\text{per}}^0 - \epsilon_F) \tilde{v}_m^\eta, \tilde{v}_m^\eta \rangle_{L^2(\mathbb{R}^3)} \\ &= \text{Tr}_{L^2(\mathbb{R}^3)} ((H_{\text{per}}^0 - \epsilon_F) |\tilde{v}_m^\eta\rangle \langle \tilde{v}_m^\eta|). \end{aligned}$$

Let us now prove that

$$\lim_{L \rightarrow \infty} D_{G_L}(\rho_{Q_L^\eta}, \rho_{Q_L^\eta}) = D(\rho_{\tilde{Q}^\eta}, \rho_{\tilde{Q}^\eta}).$$

The convergence (4.94) implies that $\mathbf{1}_{\Lambda_L} \rho_{Q_L^\eta}$ converges to $\rho_{\tilde{Q}^\eta}$ in particular in $L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$.

Notice the definition of $D_{G_L}(\cdot, \cdot)$ implies that

$$\forall \rho \in L^1_{\text{per}}(\Lambda_L) \cap L^2_{\text{per}}(\Lambda_L), \quad D_{G_L}(\rho, \rho) \leq C \left(\|\rho\|_{L^1_{\text{per}}(\Lambda_L)}^2 + \|\rho\|_{L^2_{\text{per}}(\Lambda_L)}^2 \right) \quad (4.95)$$

for a constant C independent of L . Indeed,

$$\begin{aligned}
D_{G_L}(\rho, \rho) &= \int_{\Lambda_L} (\rho *_{\Lambda_L} G_L)(x) \rho(x) dx \\
&= \sum_{k \in \frac{2\pi}{L} \mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} |c_k(\rho)|^2 \\
&= \sum_{k \in \frac{2\pi}{L} \mathbb{Z}^3 \setminus \{0\}, |k| \leq 1} \frac{4\pi}{|k|^2} |c_k(\rho)|^2 + \sum_{k \in \frac{2\pi}{L} \mathbb{Z}^3 \setminus \{0\}, |k| > 1} \frac{4\pi}{|k|^2} |c_k(\rho)|^2 \\
&\leq c_1 \|\rho\|_{L^1_{\text{per}}(\Lambda_L)}^2 + 4\pi \|\rho\|_{L^2_{\text{per}}(\Lambda_L)}^2
\end{aligned}$$

where

$$c_1 = \sup_L \frac{1}{(2\pi)^3} \sum_{k \in \frac{2\pi}{L} \mathbb{Z}^3 \setminus \{0\}, |k| \leq 1} \left(\frac{2\pi}{L} \right)^3 \frac{4\pi}{|k|^2} < \infty.$$

Let us now write $\rho_{Q_L^\eta} = \rho_{1,L} + \rho_{2,L}$ where $\rho_{1,L}$ is the periodic function which equals $\mathbf{1}_{B(0,L/4)} \rho_{Q_L^\eta}$ on Λ_L . The convergence of $\mathbf{1}_{\Lambda_L} \rho_{Q_L^\eta}$ towards $\rho_{\tilde{Q}^\eta}$ in $L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ implies by [6, Theorem IV.9] that, up to an extraction,

$$\mathbf{1}_{\Lambda_L} \rho_{Q_L^\eta}(x) \rightarrow \rho_{\tilde{Q}^\eta}(x) \quad \text{a.e. on } \mathbb{R}^3,$$

$$|\mathbf{1}_{\Lambda_L} \rho_{Q_L^\eta}(x)| \leq \tilde{\rho}(x) \quad \forall L, \quad \text{a.e. on } \mathbb{R}^3 \quad \text{with } \tilde{\rho}(x) \in L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3).$$

Thus, applying the Lebesgue dominated convergence theorem to $\rho_{2,L} = (\mathbf{1}_{\Lambda_L} - \mathbf{1}_{B(0,L/4)}) \rho_{Q_L^\eta}$ which converges to 0 almost everywhere on Λ_L , we get

$$\lim_{L \rightarrow \infty} \|\rho_{2,L}\|_{L^1_{\text{per}}(\Lambda_L)} = \lim_{L \rightarrow \infty} \|\rho_{2,L}\|_{L^2_{\text{per}}(\Lambda_L)} = 0.$$

Hence, by (4.95),

$$\lim_{L \rightarrow \infty} D_{G_L}(\rho_{2,L}, \rho_{2,L}) = 0.$$

Then it remains to show that

$$\lim_{L \rightarrow \infty} D_{G_L}(\rho_{1,L}, \rho_{1,L}) = D(\rho_{\tilde{Q}^\eta}, \rho_{\tilde{Q}^\eta}). \quad (4.96)$$

To this end we use the estimate [37]

$$\sup_{x \in \Lambda_L} \left| G_L(x) - \frac{1}{|x|} \right| = O(L^{-1}),$$

to obtain

$$\begin{aligned}
D_{G_L}(\rho_{1,L}, \rho_{1,L}) &= \iint_{(\Lambda_L)^2} G_L(x-y) \rho_{1,L}(x) \rho_{1,L}(y) dx dy, \\
&= \iint_{(\Lambda_L)^2} \frac{\rho_{1,L}(x) \rho_{1,L}(y)}{|x-y|} dx dy + O(L^{-1}), \\
&= D(\mathbf{1}_{\Lambda_L} \rho_{1,L}, \mathbf{1}_{\Lambda_L} \rho_{1,L}) + O(L^{-1}),
\end{aligned}$$

where we have used that $\|\rho_{1,L}\|_{L^1_{\text{per}}(\Lambda_L)}$ is uniformly bounded and that $x - y \in \Lambda_L$ for any $x, y \in B(0, L/4)$, the support of $\rho_{1,L}$. The convergence of $\mathbb{1}_{\Lambda_L}\rho_{1,L}$ towards $\rho_{\tilde{Q}^\eta}$ in $L^1(\mathbb{R}^3) \cap L^2(\mathbb{R}^3)$ then proves (4.96). Using the same argument for the term $D_{G_L}(\rho_{Q_L^\eta}, \nu_L)$ we obtain

$$\lim_{L \rightarrow \infty} \left(-D_{G_L}(\rho_{Q_L^\eta}, \nu_L) + \frac{1}{2}D_{G_L}(\rho_{Q_L^\eta}, \rho_{Q_L^\eta}) \right) = -D(\rho_{\tilde{Q}^\eta}, \nu) + \frac{1}{2}D(\rho_{\tilde{Q}^\eta}, \rho_{\tilde{Q}^\eta}).$$

Finally

$$\lim_{L \rightarrow \infty} E_{\text{sc},L}^\nu(Q_L^\eta) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L^\eta) = \mathcal{E}^\nu(\tilde{Q}^\eta) - \epsilon_F \text{Tr}(\tilde{Q}^\eta). \quad (4.97)$$

Passing to the limit as $\eta \rightarrow 0$ using (4.93) and the convergence of the Gram matrices S_\pm^η and S_w^η , we eventually obtain that $\tilde{Q}^\eta \rightarrow Q$ and therefore

$$\limsup_{L \rightarrow \infty} E_{\epsilon_F, L}^\nu \leq \mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}(Q) \leq E_{\epsilon_F}^\nu + \epsilon.$$

Step 3. Lower bound

We end the proof by showing that $\liminf_{L \rightarrow \infty} E_{\epsilon_F, L}^\nu \geq E_{\epsilon_F}^\nu$. As Λ_L is bounded for any fixed L , the existence of a minimizer Q_L of

$$\inf \left\{ E_{\text{sc},L}^\nu(Q_L) - \epsilon_F \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L), Q_L \in \mathcal{K}_L \right\}$$

is straightforward. In addition, the spectrum of $H_{\text{sc},L}^0$, considered as an operator on $L^2_{\text{per}}(\Lambda_L)$, is purely discrete and bounded below (see Section 2.4). Then, as Q_L is a projector over the first eigenstates, it is finite rank.

Reasoning as in the proof of Lemma 1 (see Section 4.3.1), we note first that $V_{\text{sc},L}^0 = (\rho_{\gamma_{\text{sc},L}^0} - \mu_{\text{per}}) *_{\Gamma} G_1$ is bounded and therefore, using (4.70) which assures the existence of a gap for L large enough, we know that there exists a constant $c > 0$ (independent of L) such that $|H_{\text{sc},L}^0 - \epsilon_F| \geq c(1 - \Delta)$ on $L^2_{\text{per}}(\Lambda_L)$. We have the following uniform bounds :

$$\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(|H_{\text{sc},L}^0 - \epsilon_F|^{1/2}(Q_L^{+,L} - Q_L^{-,L})|H_{\text{sc},L}^0 - \epsilon_F|^{1/2}) \leq C, \quad (4.98)$$

$$\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((1 + |\nabla|)(Q_L^{+,L} - Q_L^{-,L})(1 + |\nabla|)) \leq C, \quad (4.99)$$

$$\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((1 + |\nabla|)Q_L^2(1 + |\nabla|)) \leq C, \quad (4.100)$$

$$D_{G_L}(\rho_{Q_L} - \nu_L, \rho_{Q_L} - \nu_L) \leq C, \quad (4.101)$$

with C independent of L . (4.98) and (4.101) follow from the boundedness from above of the energy, showed in Step 2. Then, (4.99) comes from (4.98) and $|H_{\text{sc},L}^0 - \epsilon_F| \geq c(1 - \Delta)$, which is equivalent to $(1 + |\nabla|)^2$. Finally, (4.100) comes directly from (4.99) and $Q_L^2 \leq Q_L^{+,L} - Q_L^{-,L}$ being true in \mathcal{K}_L .

Consider now the sequence of operators $\tilde{Q}_L := (i_L)^* Q_L i_L$ acting on $L^2(\mathbb{R}^3)$. It is bounded in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$ by (4.100) and since $\text{Tr}_{L^2(\mathbb{R}^3)}(\tilde{Q}_L^2) = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L^2)$ by (4.83). Hence \tilde{Q}_L weakly converges, up to extraction, to some $Q \in \mathfrak{S}_2(L^2(\mathbb{R}^3))$.

Similarly, the Hilbert-Schmidt operator $R_L := (i_L)^* Q_L (1 + |\nabla|) i_L$ weakly converges up to extraction to some R in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$. Let φ and ψ be in $C_0^\infty(\mathbb{R}^3)$ and assume that $\text{Supp}(\varphi) \cup \text{Supp}(\psi) \subset \Lambda_L$. Then

$$\begin{aligned} \langle (i_L)^* Q_L (1 + |\nabla|) i_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)} &= \left\langle \tilde{Q}_L (i_L)^* (1 + |\nabla|) i_L \varphi, \psi \right\rangle_{L^2(\mathbb{R}^3)} \\ &\rightarrow_{L \rightarrow \infty} \langle Q (1 + |\nabla|) \varphi, \psi \rangle_{L^2(\mathbb{R}^3)}, \end{aligned}$$

where we have used that $\tilde{Q}_L \rightharpoonup Q$ weakly in \mathfrak{S}_2 and that $(i_L)^* (1 + |\nabla|) i_L \varphi \rightarrow (1 + |\nabla|) \varphi$ strongly in $L^2(\mathbb{R}^3)$ by the third assertion of Lemma 8. Hence $Q (1 + |\nabla|) = R \in \mathfrak{S}_2(L^2(\mathbb{R}^3))$.

Similarly, define the operator $S_L := (i_L)^* Q_L^{-,L} i_L$ which is nonpositive and yields a bounded sequence in $\mathfrak{S}_1(L^2(\mathbb{R}^3))$ by (4.99). Up to extraction, we may assume that (S_L) converges for the weak-* topology to some $S \in \mathfrak{S}_1(L^2(\mathbb{R}^3))$. To identify the limit S , we compute as above for $\varphi, \psi \in L^2(\mathbb{R}^3)$,

$$\begin{aligned} \langle S_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)} &= \langle (i_L)^* \gamma_{\text{sc},L}^0 Q_L \gamma_{\text{sc},L}^0 i_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)} \\ &= \left\langle \tilde{Q}_L (i_L)^* \gamma_{\text{sc},L}^0 i_L \varphi, (i_L)^* \gamma_{\text{sc},L}^0 i_L \psi \right\rangle_{L^2(\mathbb{R}^3)}. \end{aligned}$$

Using now the first assertion of Lemma 8 we obtain $\lim_{L \rightarrow \infty} \langle S_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)} = \langle Q^{-} \varphi, \psi \rangle_{L^2(\mathbb{R}^3)}$.

Hence $Q^{-} = S \in \mathfrak{S}_1$. The same arguments work for $(i_L)^* Q_L^{+,L} i_L$ which converges to $Q^{+} \in \mathfrak{S}_1$. Then, since $(\tilde{Q}_L)^* = \tilde{Q}_L$, it is easy to see that $Q = Q^*$. So it remains to show that $|\nabla| Q^{-} |\nabla| \in \mathfrak{S}_1$ and $-\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0$.

By (4.100), $(i_L)^* (1 + |\nabla|) Q_L^{-,L} (1 + |\nabla|) i_L$ is bounded in $\mathfrak{S}_1(L^2(\mathbb{R}^3))$, hence converges up to extraction to some $P \in \mathfrak{S}_1(L^2(\mathbb{R}^3))$ for the weak-* topology of $\mathfrak{S}_1(L^2(\mathbb{R}^3))$. Besides, for $\varphi, \psi \in L^2(\mathbb{R}^3)$,

$$\begin{aligned} &\left\langle (i_L)^* (1 + |\nabla|) Q_L^{-,L} (1 + |\nabla|) i_L \varphi, \psi \right\rangle_{L^2(\mathbb{R}^3)} \\ &= \left\langle (i_L)^* Q_L i_L (i_L)^* \gamma_{\text{sc},L}^0 i_L (i_L)^* (1 + |\nabla|) i_L \varphi, (i_L)^* \gamma_{\text{sc},L}^0 i_L (i_L)^* (1 + |\nabla|) i_L \psi \right\rangle_{L^2(\mathbb{R}^3)} \\ &= \left\langle \tilde{Q}_L (i_L)^* \gamma_{\text{sc},L}^0 i_L (i_L)^* (1 + |\nabla|) i_L \varphi, (i_L)^* \gamma_{\text{sc},L}^0 i_L (i_L)^* (1 + |\nabla|) i_L \psi \right\rangle_{L^2(\mathbb{R}^3)}, \\ &\longrightarrow_{L \rightarrow \infty} \langle Q \gamma_{\text{per}}^0 (1 + |\nabla|) \varphi, \gamma_{\text{per}}^0 (1 + |\nabla|) \psi \rangle_{L^2(\mathbb{R}^3)}, \end{aligned}$$

which means that $(i_L)^* (1 + |\nabla|) Q_L^{-,L} (1 + |\nabla|) i_L$ converges to $(1 + |\nabla|) Q^{-} (1 + |\nabla|) \in \mathfrak{S}_1(L^2(\mathbb{R}^3))$ in the weak operator sense. Hence, $(1 + |\nabla|) Q^{-} (1 + |\nabla|) = P \in \mathfrak{S}_1(L^2(\mathbb{R}^3))$.

Lastly, as $Q_L \in \mathcal{K}_L$, we have for every $\varphi \in \mathcal{D}(\mathbb{R}^3)$ and all L large enough so that $\text{Supp}(\varphi) \subset \Lambda_L$,

$$0 \leq \langle (Q_L + \gamma_{\text{sc},L}^0) i_L \varphi, i_L \varphi \rangle_{L^2_{\text{per}}(\Lambda_L)} \leq \langle i_L \varphi, i_L \varphi \rangle_{L^2_{\text{per}}(\Lambda_L)} = \langle \varphi, \varphi \rangle_{L^2(\mathbb{R}^3)}.$$

Besides, using the first assertion of Lemma 8,

$$\langle (Q_L + \gamma_{\text{sc},L}^0) i_L \varphi, i_L \varphi \rangle_{L^2_{\text{per}}(\Lambda_L)} = \langle (i_L)^* (Q_L + \gamma_{\text{sc},L}^0) i_L \varphi, \varphi \rangle_{L^2(\mathbb{R}^3)} \rightarrow \langle (Q + \gamma_{\text{per}}^0) \varphi, \varphi \rangle_{L^2(\mathbb{R}^3)}.$$

Hence, $0 \leq Q + \gamma_{\text{per}}^0 \leq 1$, which allows us to conclude that $Q \in \mathcal{K}$.

Now, let $T_L := (i_L)^* |H_{\text{sc},L}^0 - \epsilon_F|^{1/2} Q_L^{-,L} |H_{\text{sc},L}^0 - \epsilon_F|^{1/2} i_L$ which also defines a bounded sequence in $\mathfrak{S}_1(L^2(\mathbb{R}^3))$. Up to extraction, we may assume that $T_L \rightharpoonup T$ for the weak-* topology of \mathfrak{S}_1 . Arguing as above and using assertions 1 and 5 of Lemma 8, we deduce that $T = |H_{\text{per}}^0 - \epsilon_F|^{1/2} Q^{-} |H_{\text{per}}^0 - \epsilon_F|^{1/2}$. Now, Fatou's Lemma yields

$$\begin{aligned} \liminf_{L \rightarrow \infty} \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(|H_{\text{sc},L}^0 - \epsilon_F|^{1/2} (-Q_L^{-,L}) |H_{\text{sc},L}^0 - \epsilon_F|^{1/2} \right) \\ = \liminf_{L \rightarrow \infty} \text{Tr}_{L^2(\mathbb{R}^3)}(-T_L) \geq \text{Tr}_{L^2(\mathbb{R}^3)} \left(|H_{\text{per}}^0 - \epsilon_F|^{1/2} (-Q^{-}) |H_{\text{per}}^0 - \epsilon_F|^{1/2} \right) \end{aligned}$$

This proves that

$$\liminf_{L \rightarrow \infty} \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left((H_{\text{sc},L}^0 - \epsilon_F) Q_L \right) \geq \text{Tr}_0(H_{\text{per}}^0 Q) - \epsilon_F \text{Tr}_0(Q).$$

We now study the term involving the density ρ_{Q_L} . First, following the proof of Proposition 1 and using the bounds (4.98)–(4.100), we are going to prove that there exists a constant C such that for all L large enough $\|\rho_{Q_L}\|_{L^2_{\text{per}}(\Lambda_L)} \leq C$.

Let V be in $L^2_{\text{per}}(\Lambda_L) \cap L^\infty_{\text{per}}(\Lambda_L)$. As in the proof of Proposition 1, we can notice that

$$(Q_L V)^{+,L} = (\gamma_{\text{sc},L}^0)^\perp (Q_L V) (\gamma_{\text{sc},L}^0)^\perp = Q_L^{+,L} V (\gamma_{\text{sc},L}^0)^\perp + Q_L^{+,L} [\gamma_{\text{sc},L}^0, V] (\gamma_{\text{sc},L}^0)^\perp.$$

Similarly,

$$(Q_L V)^{-,L} = \gamma_{\text{sc},L}^0 (Q_L V) \gamma_{\text{sc},L}^0 = Q_L^{-,L} V \gamma_{\text{sc},L}^0 - Q_L^{-,L} [\gamma_{\text{sc},L}^0, V] \gamma_{\text{sc},L}^0, \quad (4.102)$$

where $(\gamma_{\text{sc},L}^0)^\perp = 1 - \gamma_{\text{sc},L}^0$. We only treat the $(Q_L V)^{-,L}$ term, the argument being the same for $(Q_L V)^{+,L}$.

First we write $Q_L^{-,L} V \gamma_{\text{sc},L}^0 = (1 + |\nabla|)^{-1} (1 + |\nabla|) Q_L^{-,L} (1 + |\nabla|) (1 + |\nabla|)^{-1} V \gamma_{\text{sc},L}^0$, and notice that $(1 + |\nabla|)^{-1} V$ is bounded since $V \in L^\infty_{\text{per}}(\Lambda_L)$ by assumption. As $(1 + |\nabla|)^{-1}$ and $\gamma_{\text{sc},L}^0$ are also bounded, and as $(1 + |\nabla|) Q_L^{-,L} (1 + |\nabla|)$ is trace-class thanks to (4.99), this proves that $Q_L^{-,L} V \gamma_{\text{sc},L}^0$ is a trace-class operator. Thus the following is true :

$$\begin{aligned} |\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L^{-,L} V \gamma_{\text{sc},L}^0)| &= |\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L^{-,L} V)| \\ &= |\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}((1 + |\nabla|) Q_L^{-,L} (1 + |\nabla|) (1 + |\nabla|)^{-1} V (1 + |\nabla|)^{-1})| \\ &\leq \|Q_L\|_{\mathcal{Q}_L} \|(1 + |\nabla|)^{-1} V (1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty(L^2_{\text{per}}(\Lambda_L))}. \end{aligned}$$

Then, reasoning as in the proof of Proposition 1, we obtain

$$\|(1 + |\nabla|)^{-1} V (1 + |\nabla|)^{-1}\|_{\mathfrak{S}_\infty(L^2_{\text{per}}(\Lambda_L))} \leq C \|V\|_{L^2_{\text{per}}(\Lambda_L)},$$

Hence,

$$|\text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L^{-,L} V \gamma_{\text{sc},L}^0)| \leq C \|Q_L\|_{\mathcal{Q}_L} \|V\|_{L^2_{\text{per}}(\Lambda_L)},$$

where the constant C does not depend on L (nor of course on V).

For the second term of (4.102), we use the fact that $Q_L^{-+,L} \in \mathfrak{S}_2$, $[\gamma_{sc,L}^0, V] \in \mathfrak{S}_2$, and that $\gamma_{sc,L}^0$ is a bounded operator, to claim that

$$Q_L^{-+,L}[\gamma_{sc,L}^0, V]\gamma_{sc,L}^0 \in \mathfrak{S}_1.$$

Additionally, by similar arguments as those used in the proof of Lemma 12, $\|[\gamma_{sc,L}^0, V]\|_{\mathfrak{S}_2(L_{\text{per}}^2(\Lambda_L))} \leq C \|V\|_{L_{\text{per}}^2(\Lambda_L)}$, with a constant C independent on L . Thus, we get

$$|\text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(Q_L^{-+,L}[\gamma_{sc,L}^0, V]\gamma_{sc,L}^0)| \leq C \|Q_L^{-+,L}\|_{\mathfrak{S}_2} \|V\|_{L_{\text{per}}^2(\Lambda_L)}.$$

Hence,

$$\begin{aligned} \|\rho_{Q_L}\|_{L_{\text{per}}^2(\Lambda_L)} &= \sup_{V \in L_{\text{per}}^2(\Lambda_L)} \frac{|\text{Tr}_{L_{\text{per}}^2(\Lambda_L)}(Q_L V)|}{\|V\|_{L_{\text{per}}^2(\Lambda_L)}}, \\ &\leq C \|Q_L\|_{Q_L}, \\ &\leq C'. \end{aligned}$$

Therefore, up to extraction, we have $\mathbf{1}_{\Lambda_L} \rho_{Q_L} \rightharpoonup \rho$ weakly in $L^2(\mathbb{R}^3)$ for some function $\rho \in L^2(\mathbb{R}^3)$. We now introduce an auxiliary function $\rho_L \in L^2(\mathbb{R}^3)$ defined in Fourier space as follows :

$$\widehat{\rho}_L := \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} \frac{c_{k,L}(\rho_{Q_L})}{|P_k|^{1/2}} \mathbf{1}_{P_k}$$

where for any $k \in (2\pi/L)\mathbb{Z}^3$, $P_k := k + [-\frac{\pi}{L}, \frac{\pi}{L}]^3$.

Notice that ρ_L is bounded in $L^2(\mathbb{R}^3)$ as we have

$$\int_{\mathbb{R}^3} \rho_L^2 = \int_{\mathbb{R}^3} |\widehat{\rho}_L|^2 = \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} |c_{k,L}(\rho_{Q_L})|^2 \leq \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} |c_{k,L}(\rho_{Q_L})|^2 = \int_{\Lambda_L} \rho_{Q_L}^2.$$

On the other hand (up to extraction) $\rho_L \rightharpoonup \rho$ weakly in $L^2(\mathbb{R}^3)$, the same weak limit as $\mathbf{1}_{\Lambda_L} \rho_{Q_L}$. To see this, we compute first the scalar product of $\mathbf{1}_{\Lambda_L} \rho_{Q_L}$ against a fixed function $\varphi \in C_0^\infty(\mathbb{R}^3)$: for L large enough so that $\text{Supp}(\varphi) \subset \Lambda_L$,

$$\begin{aligned} \int_{\mathbb{R}^3} \mathbf{1}_{\Lambda_L} \rho_{Q_L}(x) \varphi(x) dx &= \int_{\Lambda_L} \rho_{Q_L}(x) \varphi(x) dx \\ &= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} c_{k,L}(\rho_{Q_L}) \overline{c_{k,L}(i_L(\varphi))}, \\ &= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} c_{k,L}(\rho_{Q_L}) \frac{1}{L^{\frac{3}{2}}} \overline{\int_{\Lambda_L} \varphi(x) e^{-ik \cdot x} dx}, \\ &= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3} \left(\frac{2\pi}{L}\right)^{\frac{3}{2}} c_{k,L}(\rho_{Q_L}) \overline{\widehat{\varphi}(k)}, \end{aligned}$$

where $c_{k,L}(g)$ denotes the k -th Fourier coefficient of $g \in L^2_{\text{per}}(\Lambda_L)$ and $\hat{f}(k)$ the Fourier transform of $f \in L^2(\mathbb{R}^3)$ at point k .

Then,

$$\begin{aligned} \int_{\mathbb{R}^3} \rho_L(x) \varphi(x) dx &= \int_{\mathbb{R}^3} \hat{\rho}_L(\xi) \overline{\hat{\varphi}(\xi)} d\xi, \\ &= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} \frac{c_{k,L}(\rho_{Q_L})}{|P_k|^{\frac{1}{2}}} \int_{P_k} \overline{\hat{\varphi}(\xi)} d\xi, \\ &= \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} \left(\frac{2\pi}{L}\right)^{\frac{3}{2}} c_{k,L}(\rho_{Q_L}) \frac{1}{|P_k|} \int_{P_k} \overline{\hat{\varphi}(\xi)} d\xi. \end{aligned}$$

Thus,

$$\begin{aligned} &\left| \int_{\mathbb{R}^3} \mathbf{1}_{\Lambda_L} \rho_{Q_L}(x) \varphi(x) dx - \int_{\mathbb{R}^3} \rho_L(x) \varphi(x) dx \right| \\ &\leq \left(\left(\frac{2\pi}{L}\right)^3 |\hat{\varphi}(0)| + \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} \left(\frac{2\pi}{L}\right)^3 \left| \hat{\varphi}(k) - \frac{1}{|P_k|} \int_{P_k} \overline{\hat{\varphi}} \right| \right) \|\rho_{Q_L}\|_{L^2}^2 \end{aligned}$$

$\rightarrow 0$ when L goes to infinity.

Now by the choice of the cubes P_k , we also have for L large enough

$$\begin{aligned} D(\rho_L, \rho_L) &= 4\pi \int_{\mathbb{R}^3} \frac{|\widehat{\rho}_L(k')|^2}{|k'|^2} dk', \\ &= 4\pi \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} \frac{|c_{k,L}(\rho_{Q_L})|^2}{|P_k|} \int_{P_k} \frac{1}{|k'|^2} dk', \\ &\leq 4 \sum_{k \in \frac{2\pi}{L}\mathbb{Z}^3 \setminus \{0\}} \frac{4\pi}{|k|^2} |c_{k,L}(\rho_{Q_L})|^2, \\ &= D_{G_L}(\rho_{Q_L}, \rho_{Q_L}) \leq C. \end{aligned}$$

Hence, up to extraction we may assume that $\rho_L \rightharpoonup \rho$ weakly in \mathcal{C} . Using the regularity of $\widehat{\nu}$, we also deduce that

$$\liminf_{L \rightarrow \infty} \left(-D_{G_L}(\rho_{Q_L}, \nu_L) + \frac{1}{2} D_{G_L}(\rho_{Q_L}, \rho_{Q_L}) \right) \geq -D(\rho, \nu) + \frac{1}{2} D(\rho, \rho).$$

What remains to be proved is that $\rho = \rho_Q$ where Q is the weak limit of $(i_L)^* Q_L i_L$ obtained above. This will clearly show

$$\liminf_{L \rightarrow \infty} E_{\epsilon_F, L}^\nu \geq \mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}_0(Q) \geq E_{\epsilon_F}^\nu$$

and end the proof of Theorem 5. We identify the limit of $\mathbf{1}_{\Lambda_L} \rho_{Q_L}$ using its weak convergence to ρ in $L^2(\mathbb{R}^3)$.

We start with $\rho_{Q_L^{++},L}$ and write, fixing some $V \in C_0^\infty(\mathbb{R}^3)$ and assuming L large enough for $\text{Supp}(V) \subset \Lambda_L$,

$$\int_{\Lambda_L} \rho_{Q_L^{++},L} V = \text{Tr}_{L^2_{\text{per}}(\Lambda_L)}(Q_L^{++},L i_L(V)) = \text{Tr}_{L^2(\mathbb{R}^3)}(A_L B_L) \quad \text{with}$$

$$A_L := (i_L)^*(1 + |\nabla|)Q_L^{++},L(1 + |\nabla|)i_L, \quad B_L := (i_L)^*(1 + |\nabla|)^{-1}i_L(V)(1 + |\nabla|)^{-1}i_L.$$

The sequence (A_L) is bounded in $\mathfrak{S}_1(L^2(\mathbb{R}^3))$, hence in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$, by (4.99) and converges (up to extraction) weakly in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$. We proceed as above to identify the weak limit, using the fourth assertion of Lemma 8, and some $\varphi, \psi \in C_0^\infty(\mathbb{R}^3)$ such that $\text{Supp}(\varphi) \cup \text{Supp}(\psi) \subset \Lambda_L$

$$\begin{aligned} \langle A_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)} &= \left\langle (i_L)^*(1 + |\nabla|)Q_L^{++},L(1 + |\nabla|)i_L \varphi, \psi \right\rangle_{L^2(\mathbb{R}^3)}, \\ &= \left\langle Q_L^{++},L(1 + |\nabla|)i_L \varphi, (1 + |\nabla|)i_L \psi \right\rangle_{L^2(\mathbb{R}^3)}, \\ &\longrightarrow_{L \rightarrow \infty} \langle Q^{++}(1 + |\nabla|)\varphi, (1 + |\nabla|)\psi \rangle_{L^2(\mathbb{R}^3)}, \\ &= \langle (1 + |\nabla|)Q^{++}(1 + |\nabla|)\varphi, \psi \rangle_{L^2(\mathbb{R}^3)}, \end{aligned}$$

which means A_L converges weakly towards $(1 + |\nabla|)Q^{++}(1 + |\nabla|)$. By Lemma 9, B_L converges towards $(1 + |\nabla|)^{-1}V(1 + |\nabla|)^{-1}$ strongly in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$. We thus obtain

$$\lim_{L \rightarrow \infty} \int_{\Lambda_L} \rho_{Q_L^{++},L} V = \text{Tr}_{L^2(\mathbb{R}^3)}(Q^{++}V) = \int_{\mathbb{R}^3} \rho_{Q^{++}} V.$$

Likewise, it can be proved that the weak limit of $\rho_{Q_L^{--},L}$ is $\rho_{Q^{--}}$.

Let us now treat $\rho_{Q_L^{+-},L}$ (the other case $\rho_{Q_L^{-+},L}$ being similar). Using Cauchy's formula and (4.105) we write

$$\begin{aligned} [\gamma_{\text{sc},L}^0, i_L(V)] &= \frac{1}{2i\pi} \int_{\mathcal{C}} dz [(z - H_{\text{sc},L}^0)^{-1}, i_L(V)], \\ &= \frac{1}{2i\pi} \int_{\mathcal{C}} dz (z - H_{\text{sc},L}^0)^{-1} [H_{\text{sc},L}^0, i_L(V)] (z - H_{\text{sc},L}^0)^{-1}, \\ &= -\frac{1}{4i\pi} \int_{\mathcal{C}} dz (z - H_{\text{sc},L}^0)^{-1} [\Delta, i_L(V)] (z - H_{\text{sc},L}^0)^{-1}, \end{aligned}$$

Thus,

$$\begin{aligned} \int_{\Lambda_L} \rho_{Q_L^{+-},L} V &= \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(Q_L^{+-},L [\gamma_{\text{sc},L}^0, i_L(V)] \right) \\ &= -\frac{1}{4i\pi} \int_{\mathcal{C}} dz \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(Q_L^{+-},L (z - H_{\text{sc},L}^0)^{-1} (\Delta i_L(V) - i_L(V)\Delta) (z - H_{\text{sc},L}^0)^{-1} \right). \end{aligned}$$

The two terms being similar, we only detail the argument to pass to the limit in

$$\begin{aligned} &\text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(Q_L^{+-},L (z - H_{\text{sc},L}^0)^{-1} i_L(V)\Delta (z - H_{\text{sc},L}^0)^{-1} \right) \\ &= \text{Tr}_{L^2_{\text{per}}(\Lambda_L)} \left(\Delta (z - H_{\text{sc},L}^0)^{-1} Q_L^{+-},L (z - H_{\text{sc},L}^0)^{-1} (1 - \Delta)(1 - \Delta)^{-1} i_L(V) \right) \\ &= \text{Tr}_{L^2(\mathbb{R}^3)} \left(C_L (i_L)^*(1 - \Delta)^{-1} i_L(V) i_L \right) \end{aligned}$$

with $C_L := (i_L)^* \Delta (z - H_{\text{sc},L}^0)^{-1} Q_L^{+-,L} (z - H_{\text{sc},L}^0)^{-1} (1 - \Delta) i_L$. Let us first remark that C_L is bounded in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$ as $(Q_L^{+-,L})$ is bounded in \mathfrak{S}_2 and as $(\Delta (z - H_{\text{sc},L}^0)^{-1})$ and $((z - H_{\text{sc},L}^0)^{-1} (1 - \Delta))$ are uniformly bounded, by Lemma 10. Then we can identify the weak limit by passing to the limit in $\langle C_L \varphi, \psi \rangle$ for some fixed $\varphi, \psi \in \mathcal{C}_0^\infty(\mathbb{R}^3)$, using the uniform convergence of the resolvent for $z \in \mathcal{C}$, as shown in the proof of Lemma 8

$$\begin{aligned} \langle C_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)} &= \langle (i_L)^* \Delta (z - H_{\text{sc},L}^0)^{-1} (\gamma_{\text{sc},L}^0)^\perp Q_L \gamma_{\text{sc},L}^0 (z - H_{\text{sc},L}^0)^{-1} (1 - \Delta) i_L \varphi, \psi \rangle_{L^2(\mathbb{R}^3)}, \\ &= \langle Q_L \gamma_{\text{sc},L}^0 (z - H_{\text{sc},L}^0)^{-1} i_L (i_L)^* (1 - \Delta) i_L \varphi, (\gamma_{\text{sc},L}^0)^\perp (z - H_{\text{sc},L}^0)^{-1} i_L (i_L)^* \Delta i_L \psi \rangle_{L^2(\mathbb{R}^3)}, \\ &\longrightarrow_{L \rightarrow \infty} \left\langle Q \gamma_{\text{per}}^0 (z - H_{\text{per}}^0)^{-1} (1 - \Delta) \varphi, \gamma_{\text{per}}^0{}^\perp (z - H_{\text{per}}^0)^{-1} \Delta \psi \right\rangle_{L^2(\mathbb{R}^3)}, \\ &= \langle \Delta (z - H_{\text{per}}^0)^{-1} Q^{+-} (z - H_{\text{per}}^0)^{-1} (1 - \Delta) \varphi, \psi \rangle_{L^2(\mathbb{R}^3)}. \end{aligned} \quad (4.103)$$

Therefore, $C_L \rightharpoonup \Delta (z - H_{\text{per}}^0)^{-1} Q^{+-} (z - H_{\text{per}}^0)^{-1} (1 - \Delta)$ weakly in $\mathfrak{S}_2(L^2(\mathbb{R}^3))$. Then by Lemma 9 we know that $(i_L)^* (1 - \Delta)^{-1} i_L (V) i_L$ converges towards $(1 - \Delta)^{-1} V$ strongly in $\mathfrak{S}_2(L^2(\Lambda_L))$. Hence we can pass to the limit in (4.103), uniformly in $z \in \mathcal{C}$. We conclude that

$$\lim_{L \rightarrow \infty} \int_{\mathbb{R}^3} \mathbb{1}_{\Lambda_L} \rho_{Q_L} V = \int_{\mathbb{R}^3} \rho_Q V$$

for any $V \in \mathcal{C}_0^\infty(\mathbb{R}^3)$, thus $\rho = \rho_Q$. \square

Remark 6. In numerical simulations, the right-hand side of (4.79) is approximated by $I_{\text{sc},L,\epsilon_F}^\nu - I_{\text{sc},L,\epsilon_F}^0$ for a given value of L . This approach has several drawbacks. First, the values of L that lead to tractable numerical simulations are in many cases much too small to provide a correct estimation of the limit $L \rightarrow \infty$. Second, it is not easy to extend this method for computing E_κ^ν , to the direct evaluation of $E^\nu(q)$ for a given q (i.e. the energy of a defect with a prescribed total charge). The formalism introduced in the present article (problems (4.31) and (4.41)) suggests an alternative way for computing energies of defects in crystalline materials. Work in this direction is in progress [58].

4.5 Some useful lemmas

We gather in this section some results which we needed throughout the proofs. We start with some

4.5.1 Commutator estimates

Throughout this chapter, we have used Cauchy's formula to express the projector γ_{per}^0 :

$$\gamma_{\text{per}}^0 = \frac{1}{2i\pi} \int_{\mathcal{C}} (z - H_{\text{per}}^0)^{-1} dz, \quad (4.104)$$

where \mathcal{C} is a fixed regular bounded closed contour enclosing the lowest Z bands of the spectrum of H_{per}^0 .

The following result is a useful tool to replace the resolvent $(z - H_{\text{per}}^0)^{-1}$ with the operator $(-\Delta + 1)^{-1}$ which is easier to manipulate. Its proof is the same as the one of Lemma 1.

Lemma 10. *The operator $B(z) = (z - H_{\text{per}}^0)^{-1}(-\Delta + 1)$ defined in Lemma 2 and its inverse are bounded uniformly in $z \in \mathcal{C}$.*

The next result provides some useful properties of commutators :

Lemma 11. *The operators $[\gamma_{\text{per}}^0, \Delta]$ and $(1 - \Delta)[\gamma_{\text{per}}^0, |\nabla|](1 - \Delta)$ are bounded.*

Proof. The boundedness of $[\gamma_{\text{per}}^0, \Delta]$ follows from (4.104) and from the fact that $[(z - H_{\text{per}}^0)^{-1}, \Delta]$ is bounded uniformly in $z \in \mathcal{C}$ by Lemma 10.

Using again (4.104), it suffices to prove that $(1 - \Delta)[(z - H_{\text{per}}^0)^{-1}, |\nabla|](1 - \Delta)$ is bounded uniformly in $z \in \mathcal{C}$ to infer that $(1 - \Delta)[\gamma_{\text{per}}^0, |\nabla|](1 - \Delta)$ is bounded. In order to prove the uniform boundedness of $(1 - \Delta)[(z - H_{\text{per}}^0)^{-1}, |\nabla|](1 - \Delta)$, we use the formal equality

$$[(z - A)^{-1}, B] = (z - A)^{-1}[A, B](z - A)^{-1}. \quad (4.105)$$

We thus obtain

$$(1 - \Delta)[(z - H_{\text{per}}^0)^{-1}, |\nabla|](1 - \Delta) = B(z)^* [V_{\text{per}}, |\nabla|] B(z). \quad (4.106)$$

Using (4.104) and lemma 10, we obtain

$$\|(1 - \Delta)[\gamma_{\text{per}}^0, |\nabla|](1 - \Delta)\| \leq C \|[V_{\text{per}}, |\nabla|]\| \leq C \|\nabla V_{\text{per}}\|_{L^\infty(\mathbb{R}^3)}.$$

Lemma 12. *Let $V = V_1 + V_2$ with $V_1 \in \mathcal{C}'$ and $V_2 \in L^2(\mathbb{R}^3)$. Then $[\gamma_{\text{per}}^0, V] \in \mathfrak{S}_2$ and there exists a positive real constant C such that*

$$\|[\gamma_{\text{per}}^0, V]\|_{\mathfrak{S}_2} \leq C(\|V_1\|_{\mathcal{C}'} + \|V_2\|_{L^2(\mathbb{R}^3)}).$$

Proof. Formulas (4.104) and (4.105) lead to

$$\begin{aligned} [\gamma_{\text{per}}^0, V_2] &= -\frac{1}{4i\pi} \int_{\mathcal{C}} B(z)(-\Delta + 1)^{-1}[\Delta, V_2](-\Delta + 1)^{-1}B(z)^* dz \\ &= -\frac{1}{4i\pi} \int_{\mathcal{C}} B(z)((-\Delta + 1)^{-1}\Delta)V_2(-\Delta + 1)^{-1}B(z)^* dz \\ &\quad + \frac{1}{4i\pi} \int_{\mathcal{C}} B(z)(-\Delta + 1)^{-1}V_2(\Delta(-\Delta + 1)^{-1})B(z)^* dz. \end{aligned}$$

As $(-\Delta + 1)^{-1}$ and $(-\Delta + 1)^{-1}\Delta$ are bounded operators, we obtain, using Lemma 10,

$$\|[\gamma_{\text{per}}^0, V_2]\|_{\mathfrak{S}_2} \leq C \|(-\Delta + 1)^{-1}V_2\|_{\mathfrak{S}_2},$$

for some constant C independent of V_2 . Likewise,

$$\begin{aligned} [\gamma_{\text{per}}^0, V_1] &= -\frac{1}{4i\pi} \int_{\mathcal{C}} B(z)(-\Delta + 1)^{-1}[\Delta, V_1](-\Delta + 1)^{-1}B(z)^* dz \\ &= -\sum_{i=1}^3 \frac{1}{4i\pi} \int_{\mathcal{C}} B(z) ((-\Delta + 1)^{-1}\partial_{x_i}) \frac{\partial V_1}{\partial x_i} (-\Delta + 1)^{-1}B(z)^* dz \\ &\quad + \sum_{i=1}^3 \frac{1}{4i\pi} \int_{\mathcal{C}} B(z)(-\Delta + 1)^{-1} \frac{\partial V_1}{\partial x_i} (\partial_{x_i}(-\Delta + 1)^{-1})B(z)^* dz, \end{aligned}$$

which implies

$$\|[\gamma_{\text{per}}^0, V_1]\|_{\mathfrak{S}_2} \leq C \|(-\Delta + 1)^{-1}\nabla V_1\|_{\mathfrak{S}_2}$$

for some constant C independent of V_1 . We then use the Kato-Seiler-Simon inequality which we recalled in Section 6.1 to infer

$$\|[\gamma_{\text{per}}^0, V_2]\|_{\mathfrak{S}_2} \leq C' \|V_2\|_{L^2(\mathbb{R}^3)}, \quad (4.107)$$

$$\|[\gamma_{\text{per}}^0, V_1]\|_{\mathfrak{S}_2} \leq C' \|\nabla V_1\|_{L^2(\mathbb{R}^3)} = C' \|V_1\|_{C'}. \quad (4.108)$$

Lemma 13. *The operator $[|H_{\text{per}}^0 - \kappa|, |\nabla|]$ is bounded for any κ in the gap $(\Sigma_Z^+, \Sigma_{Z+1}^-)$.*

Proof. We have $|H_{\text{per}}^0 - \kappa| = -(H_{\text{per}}^0 - \kappa)\gamma_{\text{per}}^0 + (H_{\text{per}}^0 - \kappa)(1 - \gamma_{\text{per}}^0)$ and thus

$$\begin{aligned} [|H_{\text{per}}^0 - \kappa|, |\nabla|] &= -2(H_{\text{per}}^0 - \kappa)[\gamma_{\text{per}}^0, |\nabla|] + [|\nabla|, V_{\text{per}}](2\gamma_{\text{per}}^0 - 1) \\ &= 2(B(\kappa)^*)^{-1}(1 - \Delta)[\gamma_{\text{per}}^0, |\nabla|] + [|\nabla|, V_{\text{per}}](2\gamma_{\text{per}}^0 - 1) \end{aligned}$$

which gives the result since $\|[|\nabla|, V_{\text{per}}]\| \leq \|\nabla V_{\text{per}}\|_{L^\infty(\mathbb{R}^3)}$ and $(1 - \Delta)[\gamma_{\text{per}}^0, |\nabla|]$ is bounded by Lemma 11.

4.5.2 Density of finite rank operators in \mathcal{K}

This section is devoted to the generalization of results in [24, Appendix B] concerning the density of finite rank operators, which is useful for proving Theorem 3.

Lemma 14. *For any $Q \in \mathcal{K}$ there exist an orthogonal projector P such that $P - \gamma_{\text{per}}^0 \in \mathcal{K}$ and a trace-class operator $\delta \in \mathcal{Q}$ such that $0 \leq \delta \leq 1$, $[P, \delta] = 0$ and*

$$Q = P - \gamma_{\text{per}}^0 + \delta.$$

Proof. This is an easy adaptation of [24, Lemma 19].

We denote for simplicity $\mathfrak{H}_+ = (1 - \gamma_{\text{per}}^0)L^2(\mathbb{R}^3)$ and $\mathfrak{H}_- = \gamma_{\text{per}}^0 L^2(\mathbb{R}^3)$.

Proposition 5. *Let P be an orthogonal projector on $L^2(\mathbb{R}^3)$ such that $Q = P - \gamma_{\text{per}}^0 \in \mathcal{K}$. Denote by $(f_1, \dots, f_N) \in (\mathfrak{H}_+ \cap H^1(\mathbb{R}^3))^N$ an orthonormal basis of $E_1 = \ker(P - \gamma_{\text{per}}^0 - 1) = \ker(\gamma_{\text{per}}^0) \cap \ker(1 - P)$ and by $(g_1, \dots, g_M) \in (\mathfrak{H}_- \cap H^1(\mathbb{R}^3))^M$ an orthonormal basis of $E_{-1} = \ker(P - \gamma_{\text{per}}^0 + 1) = \ker(1 - \gamma_{\text{per}}^0) \cap \ker(P)$. Then there*

exist an orthonormal basis $(v_i)_{i \geq 1} \subset \mathfrak{H}_+ \cap H^1(\mathbb{R}^3)$ of $(E_1)^\perp$ in \mathfrak{H}_+ , an orthonormal basis $(u_i)_{i \geq 1} \subset \mathfrak{H}_- \cap H^1(\mathbb{R}^3)$ of $(E_{-1})^\perp$ in \mathfrak{H}_- , and a sequence $(\lambda_i)_{i \geq 1} \in \ell_2(\mathbb{R}^+)$ such that

$$P = \sum_{n=1}^N |f_n\rangle\langle f_n| + \sum_{i=1}^{\infty} \frac{|u_i + \lambda_i v_i\rangle\langle u_i + \lambda_i v_i|}{1 + \lambda_i^2}, \quad (4.109)$$

$$1 - P = \sum_{m=1}^M |g_m\rangle\langle g_m| + \sum_{i=1}^{\infty} \frac{|v_i - \lambda_i u_i\rangle\langle v_i - \lambda_i u_i|}{1 + \lambda_i^2}. \quad (4.110)$$

Additionally $\sum_{i \geq 1} \lambda_i^2 (\|\nabla u_i\|_{L^2}^2 + \|\nabla v_i\|_{L^2}^2) < \infty$.

Proof. This is an obvious corollary of [24, Theorem 7].

Corollary 4. *Let $Q \in \mathcal{K}$. Then there exists a sequence $\{Q_k\}_{k \geq 1}$ of finite rank operators belonging to \mathcal{K}_r such that $\|Q_k - Q\|_{\mathcal{Q}} \rightarrow 0$ as $k \rightarrow \infty$ and for any $k \geq 1$, $\text{Tr}_0(Q_k) = \text{Tr}_0(Q)$.*

Proof. Taking $\lambda_i = 0$ for $i > k$ in the decomposition of Proposition 5, one can approximate P by another projector P_k such that $P_k - \gamma_{\text{per}}^0 \rightarrow P - \gamma_{\text{per}}^0$ in \mathcal{Q} as $k \rightarrow \infty$ and $P_k - \gamma_{\text{per}}^0$ is finite rank :

$$\begin{aligned} P_k - \gamma_{\text{per}}^0 &= \sum_{n=1}^N |f_n\rangle\langle f_n| - \sum_{m=1}^M |g_m\rangle\langle g_m| + \sum_{i=1}^k \frac{\lambda_i^2}{1 + \lambda_i^2} (|v_i\rangle\langle v_i| - |u_i\rangle\langle u_i|) \\ &\quad + \sum_{i=1}^k \frac{\lambda_i}{1 + \lambda_i^2} (|u_i\rangle\langle v_i| + |v_i\rangle\langle u_i|). \end{aligned} \quad (4.111)$$

It then suffices to approximate each function in (4.111) by a smoother one, for instance by defining for $\epsilon \ll 1$, $\tilde{u}_i := \|R_\epsilon u_i\|_{L^2}^{-1} R_\epsilon u_i$ and orthonormalizing these new functions, where R_ϵ was previously defined by (4.24).

Then for any $Q = P - \gamma_{\text{per}}^0 + \delta$ of the form given by Lemma 14, it remains to approximate δ by a finite rank operator δ_k such that $[P_k, \delta_k] = 0$, which can be done in the same way.

Chapitre 5

Numerical simulation of crystals models

In this chapter we have put the article which we have published [58] upon our numerical work. We present a new variational model for computing the electronic first-order density matrix of a crystalline material in presence of a local defect. A natural way to obtain variational discretizations of this model is to expand the difference Q between the density matrix of the defective crystal and the density matrix of the perfect crystal, in a basis of precomputed maximally localized Wannier functions of the reference perfect crystal. This approach can be used within any semi-empirical or Density Functional Theory framework.

Describing the electronic state of crystals with local defects is a major issue in solid-state physics, materials science and nano-electronics [61, 68, 76]. The first self-consistent electronic structure calculations for defective crystals were performed in the late 70', by means of nonlinear Green functions methods [51, 53, 77]. In the 90', it became possible to solve the Kohn-Sham equations [62] for systems with several hundreds of electrons, and Green function methods were superseded by supercell methods [66, 69]. However, supercell methods have several drawbacks. First, the defect interacts with its periodic images. Second, the supercell must have a neutral total charge, so that in the simulation of charged defects, an artificial charge distribution (a jellium for instance) needs to be introduced to counterbalance the charge of the defect. These two drawbacks may lead to large, uncontrolled errors in the estimation of the energy of the defect. In practice, *ad hoc* correction terms are introduced to account for these errors [63]. A refinement of the supercell approach, based on a more careful treatment of the Coulomb interaction, has also been proposed in [73].

In a recent article [7], we have used rigorous thermodynamic limit arguments to derive a variational model allowing to directly compute the modification of the electronic first order density matrix generated by a (neutral or charged) local defect, when the host crystal is an insulator (or a semi-conductor). This model has a structure similar to the Chaix-Iracane model in quantum electrodynamics [13, 60]. This similarity originates from formal analogies between the Fermi sea of a defective

crystal and the Dirac sea in presence of atomic nuclei. For technical reasons, the reference model considered in [7] was the reduced Hartree-Fock model, or in other words, a Kohn-Sham model with fractional occupancies and exchange-correlation energy set to zero.

The purpose of the present article is twofold. First, the extension of our model to a generic exchange-correlation functional is discussed. Second, a rigorous justification of the numerical method consisting in expanding the difference between the density matrix of the defective crystal and the density matrix of the perfect crystal, in a basis of well-chosen Wannier functions of the reference perfect crystal, is provided : this method can be seen as a variational approximation of our model.

5.1 Derivation of the model

We consider a generic Kohn-Sham model (or rather a generic *extended* Kohn-Sham model in which fractional occupancies are allowed) with exchange correlation energy functional $E^{\text{xc}}(\rho)$. For the sake of simplicity, we omit the spin variable. The ground state of a molecular system with nuclear charge density ρ^{nuc} and \mathcal{N} electrons is obtained by solving

$$\inf \{ E_{\rho^{\text{nuc}}}^{\text{KS}}(\gamma), 0 \leq \gamma \leq 1, \text{Tr}(\gamma) = \mathcal{N} \}, \quad (5.1)$$

$$E_{\rho^{\text{nuc}}}^{\text{KS}}(\gamma) = \text{Tr} \left(-\frac{1}{2} \Delta \gamma \right) - D(\rho^{\text{nuc}}, \rho_\gamma) + \frac{1}{2} D(\rho_\gamma, \rho_\gamma) + E^{\text{xc}}(\rho_\gamma), \quad (5.2)$$

where $\rho_\gamma(\mathbf{r}) = \gamma(\mathbf{r}, \mathbf{r})$ and where

$$D(f, g) = \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{f(\mathbf{r}) g(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}'$$

is the Coulomb interaction. Still for simplicity, we detail the case of the $X\alpha$ exchange-correlation functional

$$E^{\text{xc}}(\rho) = -C_{X\alpha} \int_{\mathbb{R}^3} \rho^{4/3},$$

the extension to more accurate LDA functionals being straightforward. Likewise, replacing the all electron model considered here with a valence electron model with pseudopotentials does not bring any additional difficulty.

The above model describes a *finite* system of \mathcal{N} electrons in the electrostatic field created by the density ρ^{nuc} . Our goal is to describe an *infinite* crystalline material obtained in the thermodynamic limit $\mathcal{N} \rightarrow \infty$. In fact we shall consider two such systems. The first one is the periodic crystal obtained when, in the thermodynamic limit, the nuclear density approaches the periodic nuclear distribution of the perfect crystal :

$$\rho^{\text{nuc}} \rightarrow \rho_{\text{per}}^{\text{nuc}}, \quad (5.3)$$

$\rho_{\text{per}}^{\text{nuc}}$ being a periodic distribution. The second system is the previous crystal in presence of a local defect :

$$\rho^{\text{nuc}} \rightarrow \rho_{\text{per}}^{\text{nuc}} + \nu. \quad (5.4)$$

Typically, ν describes nuclear vacancies, interstitial nuclei, or impurities together with possible local rearrangement of the nuclei of the host crystal in the vicinity of the defect. In the simple case of a reference perfect crystal with a single atom per unit cell

$$\rho_{\text{per}}^{\text{nuc}} = \sum_{\mathbf{R} \in \mathcal{R}} z \delta_{\mathbf{R}}$$

where \mathcal{R} is the Bravais lattice of the host crystal and $\delta_{\mathbf{R}}$ is the Dirac delta measure at \mathbf{R} . If the defect consists in a impurity (the nucleus of charge z at $\mathbf{R} = \mathbf{0}$ being replaced with a nucleus of charge z'), the charge distribution ν reads

$$\nu = z' \delta_{\mathbf{U}(\mathbf{0})} - z \delta_{\mathbf{0}} + \sum_{\mathbf{R} \in \mathcal{R} \setminus \{\mathbf{0}\}} z (\delta_{\mathbf{R} + \mathbf{U}(\mathbf{R})} - \delta_{\mathbf{R}}),$$

where \mathbf{U} is the displacement field of the nuclei generated by the relaxation of the crystal. It is therefore composed of nuclei of positive charges and of “ghost nuclei” of negative charges. In this article, we assume that ν is given, and we focus on the calculation of the electronic density matrix.

The form of the density matrix γ_{per}^0 of the *perfect crystal* obtained in the thermodynamic limit (5.3) is well-known. The matrix γ_{per}^0 is a solution to the self-consistent equation

$$\gamma_{\text{per}}^0 = \chi_{(-\infty; \epsilon_F]}(H_{\text{per}}^0) \quad (5.5)$$

$$H_{\text{per}}^0 = -\frac{1}{2}\Delta + \Phi_{\text{per}} - \frac{4}{3}C_{X\alpha} \rho_{\text{per}}^{0/3}, \quad (5.6)$$

$$-\Delta\Phi_{\text{per}} = 4\pi (\rho_{\text{per}}^0 - \rho_{\text{per}}^{\text{nuc}}), \quad \Phi_{\text{per}} \text{ } \mathcal{R}\text{-periodic.}$$

The notation $P = \chi_{(-\infty; \epsilon_F]}(A)$ means that P is the spectral orthogonal projector of the self-adjoint operator A corresponding to filling all the energies up to the Fermi level ϵ_F (see for instance [40]). In our case, (5.5) means that γ_{per}^0 is the spectral projector which fills all the energies of H_{per}^0 up to the Fermi level ϵ_F , see Figure 5.1.

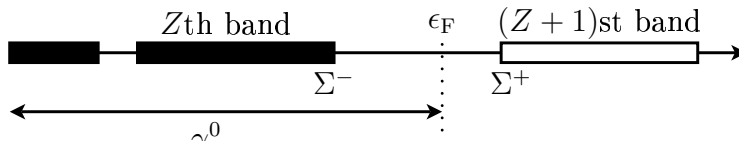


FIG. 5.1 – Spectrum of H_{per}^0 .

The density of the periodic Fermi sea is $\rho_{\text{per}}^0(\mathbf{r}) = \gamma_{\text{per}}^0(\mathbf{r}, \mathbf{r})$. Note that the system is locally neutral :

$$\int_{\Omega} \rho_{\text{per}}^0 = \int_{\Omega} \rho_{\text{per}}^{\text{nuc}},$$

where Ω is a reference unit cell, the Fermi level ϵ_F being chosen to ensure this equality. For the rest of the article, we assume that the host crystal is an insulator (or a semi-conductor), i.e. that there is a gap $g = \Sigma^+ - \Sigma^- > 0$ between the highest occupied and the lowest virtual bands. Then the Fermi level can be any number $\Sigma^- \leq \epsilon_F < \Sigma^+$.

Now we consider the system obtained in the thermodynamic limit (5.4) when there is a defect ν and derive a nonlinear variational model for it. We shall describe the variations of the Fermi sea with respect to the periodic state γ_{per}^0 . The relevant variable therefore is

$$Q = \gamma - \gamma_{\text{per}}^0$$

where γ is the density matrix of the defective Fermi sea. Notice that the constraint that γ is a density matrix ($0 \leq \gamma \leq 1$) translates into $-\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0$ for the new variable Q .

The energy of Q is by definition the difference of two infinite quantities : the energy of the state γ and the energy of the periodic Fermi sea γ_{per}^0 . Using (5.2), one obtains :

$$\mathcal{E}^\nu(Q) = \text{Tr}(H_{\text{per}}^0 Q) - D(\nu, \rho_Q) + \frac{1}{2}D(\rho_Q, \rho_Q) + \epsilon^{\text{xc}}(\rho_Q) \quad (5.7)$$

where

$$\epsilon^{\text{xc}}(\rho_Q) = -C_{\text{X}\alpha} \int_{\mathbb{R}^3} (\rho_{\text{per}}^0 + \rho_Q)^{4/3} - \rho_{\text{per}}^{0 4/3} - \frac{4}{3} \rho_{\text{per}}^{0 1/3} \rho_Q.$$

If we want to describe a defective crystal of electronic charge q (q electrons in excess with respect to the perfect crystal if $q > 0$, or $-q$ holes if $q < 0$) interacting with the self-consistent Fermi sea in the presence of the defect, we have to consider the minimization principle

$$E^\nu(q) = \inf \{ \mathcal{E}^\nu(Q), -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0, \text{Tr}(Q) = q \}. \quad (5.8)$$

We obtain in this way a model which apparently renders possible the direct calculation of the defective Fermi sea in presence of the nuclear charge defect ν , when q electrons (or $-q$ holes) are trapped by the defect. A globally neutral system would correspond to $q = \int_{\mathbb{R}^3} \nu$ but there is no obstacle in applying (5.8) to charged defects.

Alternatively, one can, instead of imposing *a priori* the total charge q of the system (microcanonical viewpoint), rather fix the Fermi level $\epsilon_{\text{F}} \in (\Sigma^-, \Sigma^+)$ (grand-canonical viewpoint). This amounts to considering the Legendre transform of (5.8) :

$$E_{\epsilon_{\text{F}}}^\nu = \inf \{ \mathcal{E}^\nu(Q) - \epsilon_{\text{F}} \text{Tr}(Q), -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0 \}. \quad (5.9)$$

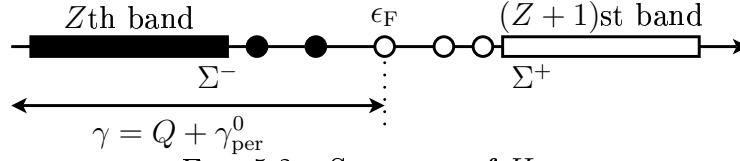
Any solution of (5.8) or (5.9) satisfies the SCF equation

$$Q = \chi_{(-\infty, \epsilon_{\text{F}})}(H_Q) - \gamma_{\text{per}}^0 + \delta, \quad (5.10)$$

where

$$H_Q = -\frac{\Delta}{2} + \Phi_{\text{per}} + (\rho_Q - \nu) * \frac{1}{|x|} - \frac{4}{3} C_{\text{X}\alpha} (\rho_{\text{per}}^0 + \rho_Q)^{1/3}$$

and where $0 \leq \delta \leq 1$ is a finite-rank self-adjoint operator on $L^2(\mathbb{R}^3)$ such that $\text{Ran}(\delta) \subset \text{Ker}(H_Q - \epsilon_{\text{F}})$. In the case of (5.8), the Fermi level ϵ_{F} is the Lagrange multiplier associated with the constraint $\text{Tr}(Q) = q$. The essential spectrum of H_Q is the same as the one of H_{per}^0 and is therefore composed of bands. On the other hand, the discrete spectrum of H_{per}^0 is empty, while the discrete spectrum of H_Q

FIG. 5.2 – Spectrum of H_Q .

may contain isolated eigenvalues of finite multiplicities located below the essential spectrum and between the bands. Each filled (or unfilled) eigenvalue may correspond to electrons (or holes) which are trapped by the defect.

The SCF equation (5.10) is equivalent to the usual Dyson equation, which is at the basis of Green function methods.

5.2 Proper definition of the variational set

The variational models (5.8) and (5.9) may look similar to the usual Kohn-Sham models for molecules and perfect crystals. Their mathematical structure is however dramatically more complex. To design consistent numerical methods for solving (5.8) and (5.9), a deeper understanding of the mathematical setting is needed.

The biggest issue with problems (5.8) and (5.9) is to properly define the variational set, that is the set of all Q 's on which one has to minimize the energy functional $\mathcal{E}^\nu(Q)$ or the free energy functional $\mathcal{E}^\nu(Q) - \epsilon_F \text{Tr}(Q)$. For usual Kohn-Sham models, the variational set is very simple : it is the largest set of density matrices for which each term of the energy functional is a well-defined number and the constraints are satisfied. This is the reason why it is not a problem to omit the precise definition of the variational set when dealing with usual Kohn-Sham models. For instance, the variational set for (5.1) is

$$\{\gamma \mid 0 \leq \gamma \leq 1, \text{Tr}(\gamma) = \mathcal{N}, \text{Tr}(|\nabla|\gamma|\nabla|) < \infty\}. \quad (5.11)$$

Let us recall (see [40] for instance) that if B is a non-negative self-adjoint operator on $L^2(\mathbb{R}^3)$ and if $(\psi_i)_{i \in \mathbb{N}}$ is an orthonormal basis of $L^2(\mathbb{R}^3)$, the series of non-negative numbers $\sum_{i=0}^{+\infty} \langle \psi_i | B | \psi_i \rangle$ converges in $\mathbb{R}_+ \cup \{+\infty\}$ towards a limit denoted by $\text{Tr}(B)$, which does not depend on the chosen basis. The operator B is said to be trace-class if $\text{Tr}(B) < \infty$. A bounded operator A on $L^2(\mathbb{R}^3)$ is trace-class if $\sqrt{A^*A}$ is trace-class. In this case, the scalar $\text{Tr}(A) = \sum_{i=0}^{+\infty} \langle \psi_i | A | \psi_i \rangle$ is well-defined and does not depend on the chosen basis. On the other hand, if A is not trace-class, the series $\sum_{i=0}^{+\infty} \langle \psi_i | A | \psi_i \rangle$ may converge for one specific basis and diverge (or converge to a different limit) in another basis.

The condition $\text{Tr}(|\nabla|\gamma|\nabla|) < \infty$ in (5.11) is a necessary and sufficient condition for each term of (5.2) being well-defined. In terms of Kohn-Sham orbitals, this conditions means that each orbital φ_i is in the Sobolev space $H^1(\mathbb{R}^3) = \{\varphi \in L^2(\mathbb{R}^3) \mid \nabla\varphi \in (L^2(\mathbb{R}^3))^3\}$.

The difficulty with the variational models (5.8) and (5.9) is that the variational set has not so simple a structure. It was shown in [7] that an appropriate variational

set is the convex set

$$\mathcal{K} = \left\{ Q \mid -\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0, \text{Tr}(1 + |\nabla|)Q^2(1 + |\nabla|) + \text{Tr}(1 + |\nabla|)(Q^{++} - Q^{--})(1 + |\nabla|) < \infty \right\}.$$

In the above expression, we have used the notation

$$Q = \left(\begin{array}{c|c} Q^{--} & Q^{-+} \\ \hline Q^{+-} & Q^{++} \end{array} \right)$$

with

$$\begin{aligned} Q^{--} &= \gamma_{\text{per}}^0 Q \gamma_{\text{per}}^0, & Q^{-+} &= \gamma_{\text{per}}^0 Q (1 - \gamma_{\text{per}}^0), \\ Q^{+-} &= (1 - \gamma_{\text{per}}^0) Q \gamma_{\text{per}}^0, & Q^{++} &= (1 - \gamma_{\text{per}}^0) Q (1 - \gamma_{\text{per}}^0), \end{aligned}$$

corresponding to the decomposition

$$L^2(\mathbb{R}^3) = \mathcal{H}_- \oplus \mathcal{H}_+, \quad (5.12)$$

where $\mathcal{H}_- = \gamma_{\text{per}}^0 L^2(\mathbb{R}^3)$ and $\mathcal{H}_+ = (1 - \gamma_{\text{per}}^0) L^2(\mathbb{R}^3)$ are respectively the occupied and virtual spaces of the reference perfect crystal.

Notice that when Q satisfies the constraint $-\gamma_{\text{per}}^0 \leq Q \leq 1 - \gamma_{\text{per}}^0$, one has $Q^{++} \geq 0$ and $Q^{--} \leq 0$. A remarkable point, proved in [7], is that the density ρ_Q of any operator $Q \in \mathcal{K}$ is a well-defined function which satisfies

$$\int_{\mathbb{R}^3} \rho_Q^2 + D(\rho_Q, \rho_Q) < \infty.$$

This shows that the electrostatic components of the energy $\mathcal{E}^\nu(\gamma)$ are well-defined and that so is the exchange-correlation contribution : as ρ_{per}^0 is periodic, continuous and positive on \mathbb{R}^3 and as $\rho_Q \in L^2(\mathbb{R}^3)$, the fifth term of (5.7) which was not considered in [7] is also well-defined. Finally, following [22], the generalized trace of an operator $Q \in \mathcal{K}$ is defined by

$$\text{Tr}(Q) = \text{Tr}(Q^{++}) + \text{Tr}(Q^{--}), \quad (5.13)$$

and for any $Q \in \mathcal{K}$, one sets

$$\text{Tr}(H_{\text{per}}^0 Q) = \text{Tr}([H_{\text{per}}^0]^{++} Q^{++}) + \text{Tr}([H_{\text{per}}^0]^{--} Q^{--}),$$

where $[H_{\text{per}}^0]^{--}$ and $[H_{\text{per}}^0]^{++}$ are respectively the restrictions to the occupied and virtual spaces of the periodic Kohn-Sham hamiltonian of the perfect crystal. Note that H_{per}^0 is block diagonal in the decomposition (5.12) :

$$H_{\text{per}}^0 = \left(\begin{array}{c|c} [H_{\text{per}}^0]^{--} & 0 \\ \hline 0 & [H_{\text{per}}^0]^{++} \end{array} \right).$$

The definition (5.13) of the trace function is an extension of the standard trace function defined on the set of trace-class operators. Note that this extension depends of

γ_{per}^0 through the decomposition (5.12) of the L^2 space. In the Quantum Electrodynamical model studied in [22–25, 60], minimizers are never trace-class (this property being related to renormalization). Whether or not the minimizers of (5.8) and (5.9) are trace-class still is an open question.

To our knowledge, the variational interpretation of the ground state solutions of the self-consistent equation (5.10) as minimizers of the energy (5.7) on the set \mathcal{K} with a constraint on the generalized trace (5.13), is new. This interpretation allows to rigorously justify the numerical method described in Section 5.4.

5.3 Interpretation in terms of Bogoliubov states

The density matrix formalism used in the previous section can be reinterpreted in terms of Bogoliubov states, following [13].

Let γ be an orthogonal projector acting on $L^2(\mathbb{R}^3)$ such that $Q = \gamma - \gamma_{\text{per}}^0 \in \mathcal{K}$. It can be proved [24] that there exists an orthonormal basis $(\varphi_i^-)_{i \geq -N_-}$ of \mathcal{H}_- and an orthonormal basis $(\varphi_i^+)_{i \geq -N_+}$ of \mathcal{H}_+ such that in this basis

$$Q = \left(\begin{array}{c|c|c|c} -I_{N_-} & 0 & 0 & 0 \\ \hline 0 & \text{diag}(-p_i) & 0 & \text{diag}(p'_i) \\ \hline 0 & 0 & I_{N_+} & 0 \\ \hline 0 & \text{diag}(p'_i) & 0 & \text{diag}(p_i) \end{array} \right) \quad (5.14)$$

with $0 \leq p_i < 1$, $\sum_{i=0}^{+\infty} p_i < \infty$, $p'_i = \sqrt{p_i(1-p_i)}$. Notice that Q is a trace-class operator if and only if $\sum_{i=0}^{+\infty} \sqrt{p_i} < \infty$. Let us assume for simplicity that in equation (5.10), the Fermi level ϵ_F is either empty or fully occupied. In this case, $\chi_{(-\infty, \epsilon_F)}(H_Q) + \delta$ is an orthogonal projector, which implies that Q can be decomposed as in (5.14). It is important to mention that in this case, the generalized trace of Q is the integer $N_+ - N_-$.

Formula (5.14) can be interpreted in terms of Bogoliubov states. The orbitals $\varphi_{-N_+}^+, \dots, \varphi_{-1}^+$ describe bound electrons in the virtual bands of the reference perfect crystal, while the orbitals $\varphi_{-N_-}^-, \dots, \varphi_{-1}^-$ represent bound holes in the occupied bands. Likewise, each pair $(\varphi_i^+, \varphi_i^-)$ with $i \geq 0$ and $0 < p_i < 1$ is a virtual electron-hole pair, and φ_i^+ and φ_i^- are the states of the corresponding Bogoliubov quasiparticles. The angle $\theta_i = \text{asin}(p_i)$ is then called the Bogoliubov angle of the virtual pair.

Formula (5.14) can itself be rewritten in a second quantized form, using the Fock space built upon the decomposition (5.12). Let us introduce the N -electron sector $\mathcal{F}_+^N := \bigwedge_1^N \mathcal{H}_+$ and the M -hole sector $\mathcal{F}_-^M := \bigwedge_1^M \mathcal{H}_-$. The electron-hole Fock space is defined as

$$\mathcal{F} := \bigoplus_{N, M \geq 0} \mathcal{F}_+^N \otimes \mathcal{F}_-^M.$$

We denote by a_i^\dagger the creation operator of an electron in the state φ_i^+ and by b_i^\dagger the creation operator of a hole in the state φ_i^- . In this formalism, the vacuum state

$\Omega_0 = 1 \otimes 1 \in \mathcal{F}_+^0 \otimes \mathcal{F}_-^0$ corresponds to the periodic Fermi sea of the perfect crystal, represented by the density matrix γ_{per}^0 in the usual Kohn-Sham description. We may also define the charge operator acting on the Fock space \mathcal{F} by

$$Q = \sum_{i \geq -N_+} a_i^\dagger a_i - \sum_{i \geq -N_-} b_i^\dagger b_i.$$

There is a special subclass of states in \mathcal{F} called Bogoliubov states [13, 22, 50, 52]. Each Bogoliubov state $\Omega_\gamma \in \mathcal{F}$ is completely characterized by its one-body density matrix γ , an orthogonal projector acting on $L^2(\mathbb{R}^3)$. Conversely, any projector γ gives rise to a Bogoliubov state under the Shale-Stinespring [71, 74] condition that $Q = \gamma - \gamma_{\text{per}}^0$ is a Hilbert-Schmidt operator (which means $\text{Tr}(Q^2) < \infty$). The role of the Shale-Stinespring condition is to ensure that Ω_γ is a well-defined state in the same Fock space as the vacuum state Ω_0 . Saying differently, this ensures that the Fock space representation associated with the splitting $L^2(\mathbb{R}^3) = \gamma L^2(\mathbb{R}^3) \oplus (1 - \gamma)L^2(\mathbb{R}^3)$ is equivalent to the one induced by (5.12) (i.e. $L^2(\mathbb{R}^3) = \gamma_{\text{per}}^0 L^2(\mathbb{R}^3) \oplus (1 - \gamma_{\text{per}}^0)L^2(\mathbb{R}^3)$). Notice the Hilbert-Schmidt condition $\text{Tr}(Q^2) < \infty$ is satisfied for any $Q = \gamma - \gamma_{\text{per}}^0$ in \mathcal{K} . Hence the variational set \mathcal{K} can be identified with a variational set of Bogoliubov states $\{\Omega_\gamma\}_{\gamma \in \mathcal{K}}$ in the Fock space \mathcal{F} .

The expression of the Bogoliubov state Ω_γ in the Fock space \mathcal{F} is given by [50, 71, 72]

$$\Omega_\gamma = c a_{-N_+}^\dagger \cdots a_{-1}^\dagger b_{-N_-}^\dagger \cdots b_{-1}^\dagger \exp \left(\sum_{i \geq 0} \lambda_i a_i^\dagger b_i^\dagger \right) \Omega_0$$

where $\lambda_i = \tan(\theta_i)$, and where c is a normalization constant. The above expression can be considered as the second-quantized formulation of (5.14). It can then easily be checked [22] that the charge of each Bogoliubov state Ω_γ (counted relatively to that of the vacuum Ω_0) is actually given by (5.13) :

$$\langle \Omega_\gamma | Q | \Omega_\gamma \rangle = \text{Tr}(Q^{++}) + \text{Tr}(Q^{--}) = N_+ - N_-$$

where $Q = \gamma - \gamma_{\text{per}}^0$.

5.4 Variational approximation

Let us now come to the discretization of problem (5.8).

If one discretizes (5.8) in a local basis without taking care of the constraint $Q \in \mathcal{K}$, there is a risk to obtain meaningless numerical results. On the other hand, selecting a basis set which respects the decomposition (5.12), will lead to a well-behaved variational approximation of (5.8) (the constraint $Q \in \mathcal{K}$ will be implicitly taken into account). Let V_\pm^h be finite-dimensional subspaces of the occupied and virtual spaces \mathcal{H}_\pm of the reference perfect crystal. Consider the finite-dimensional subspace $V^h = V_-^h \oplus V_+^h$ of $L^2(\mathbb{R}^3)$, the latter decomposition being the finite-dimensional

counterpart of (5.12). Let $(\varphi_1, \dots, \varphi_{m_-})$ (resp. $(\varphi_{m_-+1}, \dots, \varphi_{N_b})$) be an orthonormal basis of V_-^h (resp. of V_+^h). We denote for simplicity $m_+ := N_b - m_-$. The approximation set for Q consists of the finite-rank operators

$$Q = \sum_{i,j=1}^{N_b} Q_{ij}^h |\varphi_i\rangle\langle\varphi_j| \quad (5.15)$$

with $Q^h \in \mathcal{K}^h = \{Q^h = [Q^h]^T, 0 \leq \mathcal{I} + Q^h \leq 1\}$, where \mathcal{I} is the $N_b \times N_b$ block diagonal matrix

$$\mathcal{I} = \begin{bmatrix} 1_{m_-} & 0 \\ 0 & 0_{m_+} \end{bmatrix}.$$

The matrix of H_{per}^0 in the basis (φ_i) is of the form

$$H^h = \begin{bmatrix} H^{--} & 0 \\ 0 & H^{++} \end{bmatrix}.$$

For Q of the form (5.15), it holds

$$\mathcal{E}^\nu(Q) = \mathcal{E}_h^\nu(Q^h)$$

with

$$\rho_{Q^h}(r) = \sum_{i,j=1}^{N_b} Q_{ij}^h \varphi_i(r) \varphi_j(r)$$

and

$$\mathcal{E}_h^\nu(Q^h) = \text{Tr}(H^h Q^h) - D(\nu, \rho_{Q^h}) + \frac{1}{2} D(\rho_{Q^h}, \rho_{Q^h}) + \epsilon^{\text{xc}}(\rho_{Q^h}).$$

We then end up with the finite-dimensional optimization problem

$$E_h^\nu(q) = \inf \{ \mathcal{E}_h^\nu(Q^h), Q^h \in \mathcal{K}^h, \text{Tr}(Q^h) = q \} \quad (5.16)$$

which is a variational approximation of (5.8) :

$$E_h^\nu(q) \geq E^\nu(q).$$

As $Q^h \in \mathcal{K}^h$ with $\text{Tr}(Q^h) = q$ if and only if

$$\mathcal{I} + Q^h \in \{D = D^T \in \mathbb{R}^{2N_b}, D^2 \leq D, \text{Tr}(D) = q + N_-\},$$

problem (5.16) can be solved using relaxed constrained algorithms [55, 59].

The question is now to build spaces V_-^h and V_+^h that provide good approximations to (5.8) and (5.9). A natural choice is to use the maximally localized (generalized) Wannier functions [64] (MLWFs) of the reference perfect crystal. A very interesting feature of these basis functions is that they can be precalculated once and for all for a given host crystal, independently of the local defect under consideration. To construct V_-^h , one can select the maximally localized (generalized) Wannier functions

of the occupied bands, that overlap with e.g. some ball B_{R_c} of radius R_c centered on the nuclear charge defect. Note that due to the variational nature of the approximation scheme, enlarging the radius R_c systematically improves the quality of the approximation. To obtain a basis set for V_+^h , one can select a number of active (unoccupied) bands using an energy cut-off and retain the maximally localized (generalized) Wannier functions of the active bands that overlap with the same ball B_{R_c} . The so-obtained basis set of the virtual space can be enriched by adding projected atomic orbitals of the atoms and ghost atoms involved in ν (using the localized Wannier functions of the occupied bands to project out the \mathcal{H}_- component of atomic orbitals preserves the locality of these orbitals).

5.5 Numerical results

In order to illustrate the efficiency of the variational approximation presented above, we take the example of a one-dimensional (1D) model with Yukawa interaction potential, for which the energy functional reads

$$E_{1D}(\gamma) = \text{Tr} \left(-\frac{1}{2} \frac{d^2\gamma}{dx^2} \right) - D_\kappa(\rho_{\text{nuc}}, \rho_\gamma) + \frac{1}{2} D_\kappa(\rho_\gamma, \rho_\gamma)$$

with

$$D_\kappa(f, g) = (A/2\kappa) \int_{\mathbb{R}} \int_{\mathbb{R}} f(x) e^{-\kappa|x-x'|} g(x') dx dx'.$$

In the numerical examples reported below, the host crystal is \mathbb{Z} -periodic and the nuclear density is a Dirac comb, i.e.

$$\rho_{\text{nuc}} = Z \sum_{j \in \mathbb{Z}} \delta_j,$$

with Z a positive integer. The values of the parameters ($A = 10$ and $\kappa = 5$) have been chosen in such a way that the ground state kinetic and potential energies are of the same order of magnitude.

The nuclear local defect is taken of the form

$$\nu = (Z - 1)\delta_{0.25} - Z\delta_0.$$

This corresponds to moving one nucleus and lowering its charge by one unit.

The first stage of the calculation consists in solving the cell problem. For simplicity, we use a uniform discretization of the Brillouin zone $(-\pi, \pi]$, and a plane wave expansion of the crystalline orbitals.

The second stage is the construction of MLWFs. For this purpose, we make use of an argument specific to the one-dimensional case [67] : the MLWFs associated with the spectral projector γ are the eigenfunctions of the operator $\gamma x \gamma$. One first constructs N_e mother MLWFs (taking $\gamma = \gamma_{\text{per}}^0$), then N_a mother MLWFs corresponding to the lowest N_a virtual bands (taking for γ the spectral projector associated

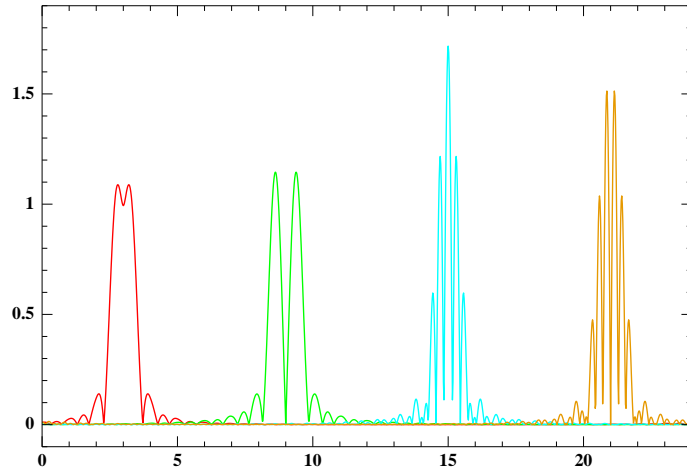


FIG. 5.3 – Modulus of MLWFs associated with the two occupied bands (left) and with the lowest two virtual bands (right).

with the lowest N_a virtual bands). The so-obtained mother MLWFs are represented on Fig. 5.3.

The third stage consists in constructing a basis set $(\varphi_j)_{1 \leq j \leq N_b}$ of $N_b = N_v(N_e + N_a)$ MLWFs by selecting the N_v translations of the $(N_e + N_a)$ mother MLWFs that are closest to the local defect, and in computing the first-order density matrix of the form (5.15) which satisfies the constraints and minimizes the energy. The profile of the density ρ_{Q^h} obtained with $Z = 2$, $N_e = 2$, $N_a = 2$ and $N_b = 28$ is displayed on Fig. 5.4. It is compared with a reference supercell calculation with 1224 plane wave basis functions. A fairly good agreement is obtained with very few MLWFs.

The implementation of our method in the Quantum Espresso suite of programs [78], in the true 3D Kohn-Sham setting, is work in progress [56].

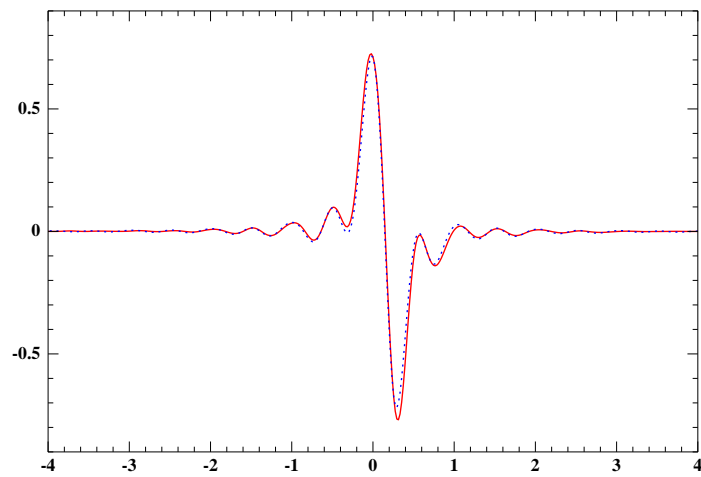


FIG. 5.4 – Density ρ_{Q^h} obtained with 28 MLWFs (line in red). The reference is a supercell calculation in a basis set of size 1224 (dashed line in blue).

Chapitre 6

Annexes : Notions d'analyse fonctionnelle

6.1 Notions d'analyse fonctionnelle

Nous présentons dans cette annexe les définitions et résultats élémentaires que nous utilisons tout au long de ce document. Les preuves des résultats énoncés peuvent être lues dans [6, 35, 40–42, 45].

6.1.0.1 Quelques inégalités d'analyse fonctionnelle

Théorème 1 (Inégalité de Hölder [6]). *Soit Ω un ouvert de \mathbb{R}^n muni de la mesure de Lebesgue.*

Soient $f \in L^p(\Omega)$ et $g \in L^{p'}(\Omega)$ avec $1 \leq p, p' \leq +\infty$ et $1/p + 1/p' = 1$. Alors $fg \in L^1(\Omega)$ et

$$\int_{\Omega} |fg| \leq \|f\|_{L^p(\Omega)} \|g\|_{L^{p'}(\Omega)}.$$

Corollaire 1 (Inégalité d'interpolation [6]). *Soit Ω un ouvert de \mathbb{R}^3 muni de la mesure de Lebesgue.*

Si $f \in L^p(\Omega) \cap L^q(\Omega)$ avec $1 \leq p \leq q \leq +\infty$, alors $f \in L^r(\Omega)$ pour tout $p \leq r \leq q$ et l'on a l'inégalité d'interpolation

$$\|f\|_{L^r} \leq \|f\|_{L^p}^{\alpha} \|f\|_{L^q}^{1-\alpha} \quad \text{où} \quad \frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q} \quad (0 \leq \alpha \leq 1).$$

Théorème 2 (Injections continues de Sobolev en dimension 3 [6]). *Soit $\Omega = \mathbb{R}^3$ ou un ouvert borné "suffisamment régulier" de \mathbb{R}^3 .*

Alors l'espace $H^1(\Omega)$ s'injecte de façon continue dans $L^p(\Omega)$ pour tout $2 \leq p \leq 6$. Il existe donc une constante $c \in \mathbb{R}_+$ telle que

$$\forall u \in H^1(\Omega), \quad \|u\|_{L^p(\Omega)} \leq c \|u\|_{H^1(\Omega)}.$$

Théorème 3 (Inégalité de Hardy-Littlewood-Sobolev [41]). Soit $1 < p, r < +\infty$ et $0 < \lambda < n$ avec $1/p + \lambda/n + 1/r = 2$. Soit $f \in L^p(\mathbb{R}^n)$ et $h \in L^r(\mathbb{R}^n)$. Alors il existe une constante $C(n, \lambda, p)$ indépendante de f et h telle que

$$\left| \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(x) |x - y|^{-\lambda} h(y) dx dy \right| \leq C(n, \lambda, p) \|f\|_{L^p} \|h\|_{L^r}.$$

On se reportera à [35] pour plus de précisions sur les bornes de la constante $C(n, \lambda, p)$.

Pour le potentiel Coulombien en dimension trois ($\lambda = 1, n = 3$), on obtient en particulier

$$\left| \int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \frac{f(x)h(y)}{|x - y|} dx dy \right| \leq C \|f\|_{L^{\frac{6}{5}}} \|h\|_{L^{\frac{6}{5}}}.$$

6.1.0.2 Transformée de Fourier

Dans ce document, nous utilisons comme convention de normalisation pour la transformée de Fourier d'une fonction $f \in \mathfrak{S}(\mathbb{R}^3)$, l'espace de Schwartz des fonctions \mathcal{C}^∞ à décroissance rapide ainsi que leurs dérivées,

$$\begin{aligned} \hat{f}(k) &= (2\pi)^{-\frac{3}{2}} \int_{\mathbb{R}^3} f(x) e^{-ik \cdot x} dx, \\ f(x) &= (2\pi)^{-\frac{3}{2}} \int_{\mathbb{R}^3} \hat{f}(k) e^{ik \cdot x} dk. \end{aligned}$$

Avec cette convention, la transformée de Fourier est une isométrie de $L^2(\mathbb{R}^3)$ dans lui-même, et on a donc les propriétés suivantes :

– formule de Plancherel

$$\forall f \in L^2(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} |f(x)|^2 dx = \int_{\mathbb{R}^3} |\hat{f}(k)|^2 dk,$$

– formule de Parseval

$$\forall (f, g) \in L^2(\mathbb{R}^3) \times L^2(\mathbb{R}^3), \quad \int_{\mathbb{R}^3} \overline{f(x)} g(x) dx = \int_{\mathbb{R}^3} \overline{\hat{f}(k)} \hat{g}(k) dk.$$

Enfin, nous utiliserons régulièrement la propriété sur la convolution

$$\widehat{fg} = (2\pi)^{-\frac{3}{2}} \hat{f} * \hat{g},$$

et

$$\widehat{f * g} = (2\pi)^{\frac{3}{2}} \hat{f} \hat{g}.$$

On utilisera indifféremment les notations \hat{f} et $\mathcal{F}f$ pour désigner la transformée de Fourier de f , et on notera \mathcal{F}^{-1} la transformée de Fourier inverse.

6.1.0.3 Opérateurs linéaires

Soient X et Y deux espaces de Banach. Un opérateur de X dans Y est une application linéaire T définie sur un sous-espace vectoriel $D(T)$ de X , à valeurs dans Y . $D(T)$ est appelé le *domaine* de l'opérateur T . L'opérateur est à domaine dense si $D(T)$ est dense dans X .

Un opérateur *borné* de X dans Y est un opérateur T de domaine $D(T) = X$ pour lequel il existe $c \in \mathbb{R}_+$ tel que la condition de continuité suivante soit vérifiée

$$\forall x \in X, \|Tx\|_Y \leq c \|x\|_X.$$

Muni de la norme définie par

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|_Y}{\|x\|_X},$$

l'espace $\mathcal{B}(X, Y)$ des opérateurs bornés de X dans Y est un espace de Banach, la topologie induite par cette norme s'appelant la *topologie uniforme*. On dit que $(T_n)_{n \in \mathbb{N}}$ converge vers T en norme d'opérateur si $\|T_n - T\|_{\mathcal{B}(X, Y)} \rightarrow_{n \rightarrow +\infty} 0$. Deux autres topologies vont nous être utiles. La *topologie forte* est la topologie la plus faible sur $\mathcal{B}(X, Y)$ telle que les applications du type $E_x : \mathcal{B}(X, Y) \rightarrow Y$ définies par $E_x(T) = Tx$ soient continues pour tout $x \in X$. Pour cette topologie, une famille d'opérateurs $(T_n)_{n \in \mathbb{N}}$ converge vers un opérateur T , ce qu'on note $s\text{-lim } T_n = T$ (*strong limit*) si et seulement si $\|T_n x - Tx\| \rightarrow 0$ pour tout $x \in X$. La *topologie faible des opérateurs* sur $\mathcal{B}(X, Y)$ est la topologie la plus faible pour laquelle les applications $E_{x,l} : \mathcal{B}(X, Y) \rightarrow \mathbb{C}$, données par $E_{x,l}(T) = l(Tx)$ soient continues pour tout $x \in X$ et $l \in Y^*$. La convergence d'une famille d'opérateurs pour cette topologie faible, notée $w\text{-lim } T_n = T$ (*weak limit*) est alors définie par $|l(T_n x) - l(Tx)| \rightarrow 0$ pour tout $x \in X$ et $l \in Y^*$.

Dans le cas particulier où l'on s'intéresse à un espace de Hilbert séparable \mathfrak{H} et aux opérateurs de $\mathcal{B}(\mathfrak{H}, \mathfrak{H}) \equiv \mathcal{B}(\mathfrak{H})$, nous avons les propriétés suivantes :

1. pour une suite $(T_n)_{n \in \mathbb{N}}$ d'opérateurs bornée sur $\mathcal{B}(\mathfrak{H})$, si on suppose que la suite $((T_n x, y))_{n \in \mathbb{N}}$ converge pour tout $x, y \in \mathfrak{H}$, alors il existe $T \in \mathcal{B}(\mathfrak{H})$ tel que $w\text{-lim } T_n = T$.
2. si l'on suppose que $(T_n x)_{n \in \mathbb{N}}$ converge pour tout $x \in \mathfrak{H}$, alors il existe $T \in \mathcal{B}(\mathfrak{H})$ tel que $s\text{-lim } T_n = T$.

Un opérateur est dit *non borné* s'il n'admet pas d'extension bornée.

6.1.0.4 Opérateurs auto-adjoints

Soit \mathfrak{H} un espace de Hilbert et T un opérateur sur \mathfrak{H} de domaine dense $D(T)$.

On pose $D(T^*) = \{u \in \mathfrak{H}, \exists v \in \mathfrak{H} / (Tw, u) = (w, v) \forall w \in D(T)\}$ et on définit l'*adjoint* de T , noté T^* , par $\forall u \in D(T^*), T^*u = v$.

On dit que T est symétrique si

$$\forall (u, v) \in D(T) \times D(T), (Tu, v) = (u, T^*v)$$

et que T est auto-adjoint si $T^* = T$. Notons que si T est un opérateur borné, il en va de même pour T^* et qu'un opérateur borné est *auto-adjoint* si et seulement si il est symétrique.

Les projecteurs orthogonaux forment une classe importante d'opérateurs bornés auto-adjoints : si $P \in \mathcal{B}(\mathfrak{H})$ et $P^2 = P$, alors P est un projecteur et si en outre $P^* = P$, P est un *projecteur orthogonal*.

Donnons à présent deux exemples d'opérateurs auto-adjoints non-bornés que nous utiliserons par la suite. Nous définissons d'abord l'opérateur $T = -\Delta$ sur $L^2(\mathbb{R}^n)$ comme l'opérateur sur $L^2(\mathbb{R}^n)$ de domaine $D(T) = H^2(\mathbb{R}^n)$ pour lequel $(T\varphi)(x) = -\Delta\varphi(x)$ dans $L^2(\mathbb{R}^n)$ pour tout $\varphi \in H^2(\mathbb{R}^n)$. On peut montrer directement, en utilisant le fait que $\mathcal{F}(-\Delta\varphi)(\xi) = |\xi|^2\hat{\varphi}(\xi)$, que cet opérateur est auto-adjoint sur $L^2(\mathbb{R}^n)$. Le second opérateur est l'opérateur $|\nabla|$: il s'agit de l'opérateur auto-adjoint sur $L^2(\mathbb{R}^n)$ de domaine $H^1(\mathbb{R}^n)$, de symbole $|p|$ dans l'espace de Fourier, défini par

$$\forall u \in H^1(\mathbb{R}^n), \quad (|\nabla|u)(x) = \mathcal{F}^{-1}(|k|\hat{u}(k))(x).$$

Notons que $|\nabla|^2 = -\Delta$.

Si T est un opérateur auto-adjoint positif (ou borné inférieurement) et $D(T)$ son domaine de définition, nous pouvons définir pour $(\varphi, \psi) \in D(T) \times D(T)$ la forme quadratique $q(\varphi, \psi) = \langle \varphi, T\psi \rangle$. Cette forme quadratique admet une plus grande extension fermée appelée *forme quadratique associée à T* , le domaine de cette plus grande extension fermée étant alors appelé le *domaine de forme* de l'opérateur T . Ainsi par exemple, pour l'opérateur $-\Delta$ sur \mathbb{R}^n , on a $q(\varphi) = \langle \varphi, -\Delta\psi \rangle = \int_{\mathbb{R}^n} \nabla\varphi \cdot \nabla\psi$, de telle sorte que le domaine de forme de $-\Delta$ est $H^1(\mathbb{R}^n)$.

Nous allons avoir besoin par la suite d'opérateur positif et de racine carrée d'un opérateur borné positif. Nous définissons donc d'abord un opérateur *positif* comme un opérateur $T \in \mathcal{B}(\mathfrak{H})$, tel que $\langle Tx, x \rangle \geq 0$ pour tout $x \in \mathfrak{H}$. On écrit alors $T \geq 0$. La racine carrée est quant à elle définie grâce au

Théorème 4. *Soit $A \in \mathcal{B}(\mathfrak{H})$, avec $A \geq 0$. Alors il existe un unique $B \in \mathcal{B}(\mathfrak{H})$ tel que $B \geq 0$ et $B^2 = A$. De plus, B commute avec tout opérateur borné qui commute avec A .*

Nous introduisons à présent la notion d'opérateur borné relativement à un opérateur auto-adjoint :

Définition 1. *Soit H un opérateur auto-adjoint et W un opérateur non borné symétrique. On dit que W est H -borné si $D(H) \subset D(W)$ et si W restreint à $D(H)$ est borné sur $D(H)$ muni de la norme du graphe à valeurs dans \mathfrak{H} , c'est-à-dire si*

$$\exists a \in \mathbb{R}_+, b \in \mathbb{R}_+ / \forall u \in D(H), \quad \|Wu\| \leq a\|Hu\| + b\|u\|. \quad (6.1)$$

La borne inférieure des réels positifs a telle que (6.1) est vérifiée pour un certain b est appelée la borne relative de W par rapport à H .

Dans ce cadre, deux résultats nous seront utiles :

Théorème 5 (Théorème de Kato-Rellich). *Soit H_0 un opérateur auto-adjoint et W un opérateur symétrique H_0 -borné de borne relative strictement inférieure à 1. Alors l'opérateur $H = H_0 + W$ défini sur le domaine $D(H) = D(H_0)$ est auto-adjoint.*

Dans ce document, nous manipulons à plusieurs reprises des opérateurs sur $L^2(\mathbb{R}^3)$ de la forme $-\Delta + V$ où V est l'opérateur de multiplication par la fonction V . De tels opérateurs sont appelés *opérateurs de Schrödinger*. Nous savons déjà que l'opérateur $-\Delta$ est auto-adjoint sur $L^2(\mathbb{R}^3)$, de domaine $H^2(\mathbb{R}^3)$. On dit qu'une fonction mesurable $V : \mathbb{R}^3 \rightarrow \mathbb{R}$ est *localement uniformément L^2 sur \mathbb{R}^3* si

$$\exists A \in \mathbb{R}_+, / \forall y \in \mathbb{R}^3, \int_{y+[0,1]^3} |V(x)|^2 dx \leq A,$$

On note $L^2_{\text{unif}}(\mathbb{R}^3)$ l'espace des fonctions localement uniformément L^2 sur \mathbb{R}^3 .

Nous avons le résultat suivant, dont la deuxième assertion résulte directement du théorème de Kato-Rellich :

Proposition 6. *Si $V \in L^2_{\text{unif}}(\mathbb{R}^3)$ est une fonction à valeurs réelles, l'opérateur de multiplication par V est un opérateur $-\Delta$ -borné, de borne relative nulle. En conséquence, l'opérateur de Schrödinger $H = -\Delta + V$ est auto-adjoint sur $L^2(\mathbb{R}^3)$, de domaine $H^2(\mathbb{R}^3)$.*

6.1.0.5 Opérateurs compacts

Un ensemble est *relativement compact* si sa fermeture est compacte. Un opérateur T de $\mathcal{B}(\mathfrak{H})$ est *compact* si l'image d'un borné de \mathfrak{H} par T est un ensemble relativement compact dans \mathfrak{H} . De manière équivalente, T est compact si et seulement si pour toute suite bornée $(x_n)_{n \in \mathbb{N}}$ de \mathfrak{H} , $(Tx_n)_{n \in \mathbb{N}}$ admet une sous-suite convergente dans \mathfrak{H} .

On note \mathfrak{S}_∞ (où $\mathfrak{S}_\infty(\mathfrak{H})$ s'il y a risque d'ambiguïté) l'ensemble des opérateurs compacts de \mathfrak{H} dans \mathfrak{H} . \mathfrak{S}_∞ est un sous-espace fermé de $\mathcal{B}(\mathfrak{H})$. De plus, c'est un idéal bilatère :

$$\text{si } T \in \mathfrak{S}_\infty \text{ et } A \in \mathcal{B}(\mathfrak{H}), \text{ alors } TA \in \mathfrak{S}_\infty \text{ et } AT \in \mathfrak{S}_\infty.$$

Tout opérateur compact peut être vu comme la limite uniforme d'une suite d'opérateurs de rang fini.

Définition 2. *Soit H_0 un opérateur auto-adjoint et W un opérateur non borné symétrique. On dit que W est H_0 -compact si W restreint à $D(H_0)$ est un opérateur compact de $D(H_0)$ muni de la norme du graphe à valeurs dans \mathfrak{H} .*

Nous donnons à présent deux résultats importants sur les opérateurs compacts.

Théorème 6 (Théorème de Hilbert-Schmidt). *Soit A un opérateur compact auto-adjoint sur un espace de Hilbert séparable \mathfrak{H} . Alors il existe une base hilbertienne $(\varphi_n)_{n \in \mathbb{N}}$ de \mathfrak{H} telle que $A\varphi_n = \lambda_n \varphi_n$ avec $\lambda_n \in \mathbb{R}$ et $\lambda_n \rightarrow 0$ lorsque $n \rightarrow \infty$.*

Théorème 7 (Forme canonique pour les opérateurs compacts). *Soit A un opérateur compact sur un espace de Hilbert séparable \mathfrak{H} . Alors il existe des bases hilbertiennes $(\psi_n)_{n \in \mathbb{N}}$ et $(\varphi_n)_{n \in \mathbb{N}}$ et des réels positifs $(\lambda_n)_{n \in \mathbb{N}}$, avec $\lambda_n \rightarrow 0$ tels que*

$$A = \sum_{n \in \mathbb{N}} \lambda_n (\psi_n, \cdot) \varphi_n. \quad (6.2)$$

Les (λ_n) sont appelées les valeurs singulières de A .

6.1.0.6 Opérateurs à trace

Introduisons tout d'abord la notion de trace d'opérateur. C'est une généralisation de la notion usuelle de somme des éléments diagonaux d'une matrice, à ceci près qu'un opérateur borné générique d'un espace de Hilbert n'a pas forcément de trace étant donné qu'il s'agit d'une somme infinie.

Théorème 8. *Soit \mathfrak{H} un espace de Hilbert séparable et $(\varphi_n)_{n \in \mathbb{N}}$ une base hilbertienne de \mathfrak{H} . Alors pour tout opérateur positif $A \in \mathcal{B}(\mathfrak{H})$ on pose $\text{Tr}(A) = \sum_{n \in \mathbb{N}} (\varphi_n, A\varphi_n)$. Le nombre $\text{Tr}(A)$ (défini dans $\mathbb{R}_+ \cup \{0\}$) est appelé trace de A et est indépendant de la base hilbertienne choisie. Il vérifie en outre les propriétés suivantes :*

1. $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$,
2. $\text{Tr}(\lambda A) = \lambda \text{Tr}(A) \quad \forall \lambda \geq 0$,
3. $\text{Tr}(UAU^{-1}) = \text{Tr}(A)$ pour tout opérateur U unitaire,
4. Si $0 \leq A \leq B$ alors $0 \leq \text{Tr}(A) \leq \text{Tr}(B)$.

Il en découle la définition suivante :

Définition 3. *Un opérateur $A \in \mathcal{B}(\mathfrak{H})$ est à trace ("trace class") si et seulement si $\text{Tr}(|A|) < \infty$ où $|A| = \sqrt{A^*A}$. La famille des opérateurs à trace est notée \mathfrak{S}_1 (ou $\mathfrak{S}_1(\mathfrak{H})$ s'il y a risque d'ambiguïté).*

\mathfrak{S}_1 est un espace de Banach pour la norme $\|\cdot\|_{\mathfrak{S}_1}$ définie par $\|A\|_{\mathfrak{S}_1} = \text{Tr}(|A|)$ et c'est un idéal $*$, c'est-à-dire un idéal bilatère dans $\mathcal{B}(\mathfrak{H})$, qui vérifie en outre que $A^* \in \mathfrak{S}_1$ si $A \in \mathfrak{S}_1$.

Passons à présent à quelques propriétés concernant l'espace \mathfrak{S}_1 .

Théorème 9. *Tout opérateur $A \in \mathfrak{S}_1$ est compact. Réciproquement, un opérateur A compact appartient à \mathfrak{S}_1 si et seulement si $\sum_{n \in \mathbb{N}} \lambda_n < \infty$ où les $(\lambda_n)_{n \in \mathbb{N}}$ sont les valeurs singulières de A .*

Corollaire 2. *Les opérateurs de rang fini sont denses dans \mathfrak{S}_1 .*

Théorème 10. *\mathfrak{S}_1 est le dual de l'espace \mathfrak{S}_∞ des opérateurs compacts et le dual de \mathfrak{S}_1 est l'ensemble des opérateurs bornés $\mathcal{B}(\mathfrak{H})$.*

Théorème 11. *Soit $A \in \mathfrak{S}_1$ et $(\varphi_n)_{n \in \mathbb{N}}$ une base orthonormée. Alors $\sum_{n \in \mathbb{N}} (\varphi_n, A\varphi_n)$ converge absolument et sa limite est indépendante du choix de la base orthonormée. On définit la trace de A comme*

$$\text{Tr}(A) = \sum_{n \in \mathbb{N}} (\varphi_n, A\varphi_n)$$

où $(\varphi_n)_{n \in \mathbb{N}}$ est une base hilbertienne quelconque de \mathfrak{H} .

6.1.0.7 Opérateurs de Hilbert-Schmidt

Nous utilisons la notion de trace pour définir une seconde classe d'opérateurs :

Définition 4. *Un opérateur $A \in \mathcal{B}(\mathfrak{H})$ est appelé opérateur de Hilbert-Schmidt si et seulement si $\text{Tr}(A^*A) < \infty$. Cette famille d'opérateurs, notée \mathfrak{S}_2 (ou $\mathfrak{S}_2(\mathfrak{H})$ s'il y a risque d'ambiguïté), a une structure d'espace de Hilbert lorsqu'on la munit du produit scalaire $(\cdot, \cdot)_{\mathfrak{S}_2}$ défini par*

$$(A, B)_{\mathfrak{S}_2} = \text{Tr}(A^*B) = \sum_{n=1}^{\infty} (\varphi_n, A^*B\varphi_n),$$

avec $A, B \in \mathfrak{S}_2$ et $(\varphi_n)_{n \in \mathbb{N}}$ base orthonormée quelconque. La norme qui en découle est $\|A\|_{\mathfrak{S}_2} = \sqrt{(A, A)_{\mathfrak{S}_2}} = (\text{Tr}(A^*A))^{\frac{1}{2}}$, et on a $\|A\|_{\mathfrak{S}_2} \leq \|A\|_{\mathfrak{S}_1}$.

Théorème 12.

- \mathfrak{S}_2 est un idéal-*,
- tout opérateur $A \in \mathfrak{S}_2$ est compact et réciproquement tout opérateur A compact appartient à \mathfrak{S}_2 si et seulement si $\sum_{n \in \mathbb{N}} \lambda_n^2 < \infty$ où les $(\lambda_n)_{n \in \mathbb{N}}$ sont les valeurs singulières de A ,
- Les opérateurs de rang fini sont denses dans \mathfrak{S}_2 .

En outre,

Théorème 13. *A appartient à \mathfrak{S}_1 si et seulement si A s'écrit comme le produit de deux opérateurs B et C appartenant à \mathfrak{S}_2 .*

Théorème 14. Soit $\langle \Omega, \mu \rangle$ un espace mesuré et $\mathfrak{H} = L^2(\Omega, d\mu)$. Alors $A \in \mathcal{B}(\mathfrak{H})$ est un opérateur de Hilbert-Schmidt si et seulement si il existe une fonction

$$a \in L^2(\Omega \times \Omega, d\mu \otimes d\mu),$$

telle que

$$\forall f \in \mathfrak{H}, (Af)(x) = \int_{\Omega} a(x, y) f(y) d\mu(y), \quad \mu - \text{presque partout.}$$

De plus,

$$\|A\|_{\mathfrak{S}_2}^2 = \iint_{\Omega \times \Omega} |a(x, y)|^2 d\mu(x) d\mu(y) = \|a\|_{L^2(\Omega \times \Omega)}^2.$$

Ainsi, dans le cas particulier où l'on considère un opérateur $A \in \mathcal{B}(L^2(\Omega))$ dont le noyau s'écrit $a(x, y) = G(x - y)V(y)$ avec $G \in L^2(\Omega)$ et $V \in L^2(\Omega)$, on a $A \in \mathfrak{S}_2$ et

$$\begin{aligned} \|A\|_{\mathfrak{S}_2}^2 &= \|a\|_{L^2(\Omega \times \Omega)}^2 = \iint_{\Omega \times \Omega} |G(x - y)|^2 |V(y)|^2 dx dy, \\ &= \int_{\Omega} \left(\int_{\Omega} |G(x - y)|^2 dx \right) |V(y)|^2 dy \quad \text{par Fubini,} \\ &= \int_{\Omega} \left(\int_{\Omega} |G(z)|^2 dz \right) |V(y)|^2 dy, \\ &= \|G\|_{L^2(\Omega)}^2 \|V\|_{L^2(\Omega)}^2. \end{aligned}$$

Ce résultat est utilisé au chapitre 4.

6.1.0.8 Opérateurs de classe \mathfrak{S}_p

Soit \mathfrak{H} un espace de Hilbert séparable. Pour $1 \leq p < +\infty$ nous définissons la classe de Schatten

$$\mathfrak{S}_p = \{A \in \mathcal{B}(\mathfrak{H}) \mid |A|^p \in \mathfrak{S}_1\},$$

(notée $\mathfrak{S}_p(\mathfrak{H})$ s'il y a risque d'ambiguïté) que l'on munit de la norme $\|A\|_{\mathfrak{S}_p} = (\text{Tr}(|A|^p))^{1/p} = (\text{Tr}(|A^*A|^{p/2}))^{1/p}$.

Nous avons les propriétés suivantes [41]

- Soit $1 \leq p \leq +\infty$ et p' tel que $1/p + 1/p' = 1$. Si $A \in \mathfrak{S}_p$ et $B \in \mathfrak{S}_{p'}$ alors $AB \in \mathfrak{S}_1$ et l'on a $\|AB\|_{\mathfrak{S}_1} \leq \|A\|_{\mathfrak{S}_p} \|B\|_{\mathfrak{S}_{p'}}$;
- Soit $1 \leq p \leq +\infty$. Alors \mathfrak{S}_p , muni de la norme $\|\cdot\|_{\mathfrak{S}_p}$, est un espace de Banach. Si $A \in \mathfrak{S}_p$ alors $A^* \in \mathfrak{S}_p$ et $\|A\|_{\mathfrak{S}_p} = \|A^*\|_{\mathfrak{S}_p}$. Enfin, \mathfrak{S}_p est la fermeture des opérateurs de rang fini pour la norme $\|\cdot\|_{\mathfrak{S}_p}$;
- Soit $1 \leq p \leq +\infty$ alors $\mathfrak{S}_1 \subset \mathfrak{S}_p \subset \mathfrak{S}_\infty$ et de manière générale pour $p \leq q$, $\mathfrak{S}_p \subset \mathfrak{S}_q$.

Théorème 15 (Inégalité de Kato-Seiler-Simon). *Nous nous restreignons ici au cas tridimensionnel (pour le cas général, voir [44] et [45, Thm 4.1]).*

Si $f \in L^p(\mathbb{R}^3)$ et $g \in L^p(\mathbb{R}^3)$ avec $2 \leq p < +\infty$, alors $f(x)g(-i\nabla)$ est dans $\mathfrak{S}_p(L^p(\mathbb{R}^3))$ et

$$\|f(x)g(-i\nabla)\|_{\mathfrak{S}_p} \leq (2\pi)^{-3/p} \|f\|_{L^p(\mathbb{R}^3)} \|g\|_{L^p(\mathbb{R}^3)}. \quad (6.3)$$

$f(x)g(-i\nabla)$ est en fait une notation formelle permettant de représenter facilement une classe d'opérateurs linéaires sur $L^2(\mathbb{R}^3)$ [45, Chapter 4]. Pour deux fonctions f et g mesurables et finies presque partout sur \mathbb{R}^3 nous introduisons les ensembles $D_f = \{h \in L^2 \mid fh \in L^2\}$ et $D_g = \{h \in L^2 \mid gh \in L^2\}$ (denses dans L^2). Pour $\varphi \in D_f$ et $\hat{\psi} \in D_g$, on définit $\langle \varphi, f(x)g(-i\nabla)\psi \rangle$ par

$$\langle \varphi, f(x)g(-i\nabla)\psi \rangle = \int_{\mathbb{R}^3} \overline{f(x)\varphi(x)} \mathcal{F}^{-1}(g\hat{\psi})(x) dx.$$

En pratique, nous utiliserons l'inégalité de Kato-Seiler-Simon pour majorer la norme $\|\cdot\|_{\mathfrak{S}_p}$ d'opérateurs du type $V(x)(1 - \Delta)^{-1}$ qui est bien de la forme $f(x)g(-i\nabla)$ avec $f(x) = V(x)$ et $g(p) = \frac{1}{1+p^2}$.

6.1.0.9 Bases de théorie spectrale

On appelle *ensemble résolvant* d'un opérateur linéaire T sur l'espace de Hilbert \mathfrak{H} , l'ensemble noté $\rho(T)$ des nombres complexes λ tels que $\lambda - T$ admette un inverse borné, noté $R_\lambda(T)$. La famille $(R_\lambda(T))_{\lambda \in \rho(T)}$ est la résolvante de T . L'application $\lambda \mapsto R_\lambda(T)$ est analytique sur l'ouvert $\rho(T)$, à valeurs dans $\mathcal{L}(\mathfrak{H})$ et on a l'*identité de la résolvante*

$$\forall (\lambda, \mu) \in \rho(T) \times \rho(T), \quad -R_\lambda(T) + R_\mu(T) = (\lambda - \mu)R_\lambda(T)R_\mu(T).$$

Le *spectre* de T , noté $\sigma(T)$, est le sous-ensemble fermé complémentaire de $\rho(T)$ dans \mathbb{C} . Il se décompose en l'union disjointe de trois sous-ensembles

- le *spectre ponctuel* $\sigma_p(T)$ est l'ensemble des $\lambda \in \mathbb{C}$ tels que $\lambda - T$ est non injectif; les λ correspondants sont des valeurs propres de T ,
- le *spectre continu* $\sigma_c(T)$ est l'ensemble des $\lambda \in \mathbb{C}$ tels que $\lambda - T$ est injectif et $\text{Ran}(\lambda - T) \neq \overline{\text{Ran}(\lambda - T)} = \mathfrak{H}$,
- le *spectre résiduel* est l'ensemble des $\lambda \in \mathbb{C}$ qui ne sont pas valeurs propres et pour lesquels $\text{Ran}(\lambda - T)$ n'est pas dense dans \mathfrak{H} .

L'analyse spectrale joue un rôle central en mécanique quantique. Ainsi par exemple, l'Hamiltonien d'un système moléculaire est un opérateur auto-adjoint non borné sur un certain espace de Hilbert et son spectre ponctuel correspond aux niveaux d'énergie des états liés du système.

Nous énonçons ici deux premiers résultats pour les spectres d'opérateurs auto-adjoints.

Théorème 16. *Soit T un opérateur auto-adjoint sur un espace de Hilbert \mathfrak{H} . Alors,*

1. T n'a pas de spectre résiduel,
2. $\sigma(T)$ est un sous-ensemble de \mathbb{R} ,
3. les vecteurs propres correspondant à des valeurs propres de T distinctes sont orthogonaux.

Nous définissons à présent d'autres sous-ensembles du spectre.

Définition 5. *On appelle spectre discret de l'opérateur auto-adjoint T , noté $\sigma_d(T)$, l'ensemble des valeurs propres isolées de T de multiplicité finie.*

On appelle alors spectre essentiel l'ensemble $\sigma_{\text{ess}}(T) = \sigma(T) \setminus \sigma_d(T)$.

Le spectre essentiel contient donc le spectre continu de T , les valeurs propres non isolées et les valeurs propres isolées de multiplicité infinie. Le théorème suivant est utilisé à plusieurs reprises dans le chapitre 3.

Théorème 17 (Théorème de Weyl). *Soit H_0 un opérateur auto-adjoint sur \mathfrak{H} de domaine $D(H_0)$ et W un opérateur symétrique sur \mathfrak{H} , H_0 -compact. Alors l'opérateur $H = H_0 + W$ de domaine $D(H) = D(H_0)$ est auto-adjoint et $\sigma_{\text{ess}}(H) = \sigma_{\text{ess}}(H_0)$.*

Grâce au théorème de Hilbert-Schmidt (ou théorème de diagonalisation des opérateurs auto-adjoints compacts) vu à la Section 6.1.0.5, nous savons que pour un opérateur auto-adjoint compact T , il existe une suite de réels $(\lambda_n)_{n \in \mathbb{N}}$ tendant vers zéro et une base hilbertienne $(e_n)_{n \in \mathbb{N}}$ de \mathfrak{H} tels que

$$\forall n \in \mathbb{N}, \quad T e_n = \lambda_n e_n.$$

Nous pouvons à présent ajouter que $\sigma(T) = \{\lambda_n\}_{n \in \mathbb{N}} \cup \{0\}$ et que $\forall \lambda \in \sigma(T) \setminus \{0\}$, $\text{Ker}(\lambda - T)$ est de dimension finie, soit en d'autres termes que $\lambda \in \sigma_d(T)$. Enfin $\sigma_{\text{ess}}(T) = \{0\}$, le réel 0 pouvant être soit dans le spectre ponctuel soit dans le spectre continu, selon que T est, ou n'est pas, injectif.

Terminons cette section par quelques mots sur la décomposition spectrale des opérateurs auto-adjoints et le calcul fonctionnel.

Une première manière d'appréhender la décomposition spectrale d'un opérateur consiste à considérer le cas où le spectre est purement ponctuel. Dans ce cas, un opérateur auto-adjoint T sur un espace de Hilbert \mathfrak{H} est décomposable sous la forme

$$T = \sum_{n \in \mathbb{N}} \lambda_n |e_n\rangle \langle e_n|$$

où $(e_n)_{n \in \mathbb{N}}$ désigne une base orthonormale de \mathfrak{H} . Dans l'équation ci-dessus, on a utilisé la notation bra-ket : le "ket" $|e_n\rangle$ désigne le vecteur e_n et le "bra" $\langle e_n|$ symbolise la forme linéaire (e_n, \cdot) . En utilisant ces notations, le domaine de l'opérateur T est

$D(T) = \{x \in \mathfrak{H}, x = \sum_{n \in \mathbb{N}} x_n e_n, \sum_{n \in \mathbb{N}} |\lambda_n|^2 |x_n|^2 < \infty\}$. Enfin pour une fonction mesurable $f : \mathbb{R} \rightarrow \mathbb{C}$, on définit l'opérateur $f(T)$ par :

$$D(f(T)) = \{x \in \mathfrak{H}, x = \sum_{n \in \mathbb{N}} x_n e_n, \sum_{n \in \mathbb{N}} |f(\lambda_n)|^2 |x_n|^2 < \infty\},$$

$$\forall x \in D(f(T)), f(T)x = \sum_{n \in \mathbb{N}} f(\lambda_n) x_n e_n. \quad (6.4)$$

Cette formule constitue la base du calcul fonctionnel (pour les opérateurs auto-adjoints à spectre purement ponctuel). Si l'on désigne par χ_I la fonction caractéristique de l'ensemble $I \subset \mathbb{R}$, l'opérateur $\chi_I(T)$ défini par

$$\chi_I(T) = \sum_{n \in \mathbb{N}} \chi_I(\lambda_n) |e_n\rangle \langle e_n| = \sum_{n/\lambda_n \in I} |e_n\rangle \langle e_n|$$

est appelé le projecteur spectral sur I de l'opérateur auto-adjoint T .

Pour définir la décomposition spectrale d'un opérateur auto-adjoint quelconque (non nécessairement compact), nous avons d'abord besoin d'introduire la notion de *famille spectrale*.

Soit \mathfrak{H} un Hilbert et $(P_\lambda)_{\lambda \in \mathbb{R}}$ une famille de projecteurs orthogonaux sur \mathfrak{H} vérifiant

1. $P_\lambda P_\mu = P_{\inf(\lambda, \mu)}$,
2. $s\text{-}\lim_{\substack{\mu > \lambda \\ \mu \rightarrow \lambda}} P_\mu = P_\lambda$,
3. $s\text{-}\lim_{\lambda \rightarrow -\infty} P_\lambda = 0$ et $s\text{-}\lim_{\lambda \rightarrow +\infty} P_\lambda = 1$ (ici 1 désigne l'opérateur identité sur \mathfrak{H}),

$(P_\lambda)_{\lambda \in \mathbb{R}}$ est appelée une famille spectrale.

Nous pouvons énoncer un premier théorème concernant le lien entre famille spectrale et opérateur auto-adjoint :

Théorème 18. *Soit \mathfrak{H} un espace de Hilbert et $(P_\lambda)_{\lambda \in \mathbb{R}}$ une famille spectrale sur \mathfrak{H} . On pose $D(T) = \{x \in \mathfrak{H}, \int_{-\infty}^{+\infty} |\lambda|^2 d\|P_\lambda x\|^2 < \infty\}$ et $\forall x \in D(T), Tx = \int_{-\infty}^{+\infty} \lambda dP_\lambda x$. Alors T définit un opérateur auto-adjoint sur \mathfrak{H} . On note souvent $1 = \int_{-\infty}^{+\infty} dP_\lambda$ et $T = \int_{-\infty}^{+\infty} \lambda dP_\lambda$.*

Nous définissons à présent deux ensembles :

Soit \mathcal{S} l'ensemble des familles spectrales sur \mathfrak{H} , \mathcal{A} l'ensemble des opérateurs auto-adjoints sur \mathfrak{H} .

Le théorème précédent permet de donner un sens à l'application

$$\mathcal{F} : \mathcal{S} \rightarrow \mathcal{A},$$

$$(P_\lambda) \rightarrow T : D(T) = \{x \in \mathfrak{H}, \int_{-\infty}^{+\infty} |\lambda|^2 d\|P_\lambda x\|^2 < \infty\} \rightarrow \mathfrak{H},$$

$$\bar{x} \rightarrow \int_{-\infty}^{+\infty} \lambda dP_\lambda x.$$

et nous avons par ailleurs [40]

Théorème 19 (Théorème spectral). *L'application \mathcal{F} définit une correspondance biunivoque entre \mathcal{S} et \mathcal{A} , autrement dit à tout opérateur auto-adjoint T sur \mathfrak{H} , on peut associer une famille spectrale $(P_\lambda^T)_{\lambda \in \mathbb{R}}$ telle que $\mathcal{F}((P_\lambda^T)) = T$. On peut construire $(P_\lambda^T)_{\lambda \in \mathbb{R}}$ à partir de la résolvante de T à l'aide de la formule suivante : si $\lambda < \mu$ sont deux réels, on a*

$$\frac{1}{2} (P_\mu^T + P_{\mu-0}^T) - \frac{1}{2} (P_\lambda^T + P_{\lambda-0}^T) = \frac{1}{2i\pi} \int_{\mathcal{C}} (z - T)^{-1} dz$$

où \mathcal{C} est une courbe fermée bornée régulière du plan complexe telle que $\mathcal{C} \cap \mathbb{R} = \{\lambda, \mu\}$ et telle que les tangentes à \mathcal{C} en λ et en μ ne soient pas parallèles à l'axe réel.

Notons que si T est un opérateur auto-adjoint compact de la forme (6.4), la famille spectrale qui lui est associée est donnée par $P_\lambda = \sum_{\lambda_n \leq \lambda} |e_n\rangle\langle e_n|$.

Pour en revenir au calcul fonctionnel, pour un opérateur auto-adjoint T sur \mathfrak{H} et une fonction mesurable $f : \mathbb{R} \rightarrow \mathbb{C}$, on définit l'opérateur $f(T)$ par :

$$D(f(T)) = \left\{ x \in \mathfrak{H}, \int_{-\infty}^{+\infty} |f(\lambda)|^2 d\|P_\lambda\|^2 < \infty \right\},$$

$$\forall x \in D(f(T)), f(T)x = \int_{-\infty}^{+\infty} f(\lambda) dP_\lambda x.$$

En particulier, si $I \subset \mathbb{R}$, le projecteur spectral sur I associé à l'opérateur auto-adjoint T est défini par

$$\chi_I(T) = \int_{\mathbb{R}} \chi_I(\lambda) dP_\lambda^T = \int_I dP_\lambda^T.$$

Enfin, nous pouvons établir un lien entre le spectre d'un opérateur et la famille spectrale :

Théorème 20. *Soit \mathfrak{H} un espace de Hilbert séparable et T un opérateur auto-adjoint sur \mathfrak{H} . On a :*

1. $\sigma(T) \subset \mathbb{R}$,
2. Si $T \geq c$ (T borné inférieurement), alors $\sigma(T) \subset [c, +\infty)$,
3. Si $T \in \mathcal{B}(\mathfrak{H})$, $\sigma(T) \subset [-\|T\|, \|T\|]$. De plus au moins l'un des deux réels $-\|T\|$ et $\|T\|$ est dans le spectre de T ,
4. $\lambda_0 \in \sigma_p(T)$ si et seulement si $P_{\lambda_0} \neq P_{\lambda_0-0}$ et $\text{Ker}(\lambda_0 - T) = \text{Ran}(P_{\lambda_0} - P_{\lambda_0-0})$,
5. $\lambda_0 \in \sigma_c(T)$ si et seulement si $P_{\lambda_0} = P_{\lambda_0-0}$ et $\forall \epsilon > 0, P_{\lambda_0+\epsilon} \neq P_{\lambda_0-\epsilon}$,
6. $\sigma_r(T) = \emptyset$,
7. $\lambda_0 \in \sigma_d(T)$ si et seulement si il existe $\epsilon > 0$ tel que $\dim \text{Ran}(P_{\lambda_0+\epsilon} - P_{\lambda_0-\epsilon}) < \infty$,
8. $\lambda_0 \in \sigma_{\text{ess}}(T)$ si et seulement si pour tout $\epsilon > 0 \dim \text{Ran}(P_{\lambda_0+\epsilon} - P_{\lambda_0-\epsilon}) = \infty$.

Bibliographie

Sur la modélisation théorique des cristaux

- [1] V. Bach, J.-M. Barbaroux, B. Helffer and H. Siedentop. On the Stability of the Relativistic Electron-Positron Field, *Comm. Math. Phys.* **201** (1999), no. 2, 445–460.
- [2] R. Benguria, H. Brezis and E.H. Lieb. The Thomas-Fermi-von Weizsäcker theory of atoms and molecules, *Comm. Math. Phys.* **79** (1981), no. 2, 167–180.
- [3] R. Bhatia. *Matrix analysis*, Graduate Texts in Mathematics, **169**. Springer-Verlag, New York, 1997.
- [4] F. Bloch. "Über die Quantenmechanik der Elektronen in Kristallgittern", *Z. Phys.* **52** (1928), 555–560.
- [5] M. Born and R. Oppenheimer. *Zur Quantentheorie der Molekeln*, *Ann. Phys.* **84** (1927), 457–484.
- [6] H. Brezis. *Analyse fonctionnelle. Théorie et applications*, Dunod 2002.
- [7] E. Cancès, A. Deleurence and M. Lewin. A new approach to the modelling of local defects in crystals : the reduced Hartree-Fock case, *Comm. Math. Phys.* **281** (2008), no. 1, 129–177.
- [8] I. Catto and P.-L. Lions. *Binding of atoms and stability of molecules in Hartree and Thomas-Fermi type theories*, Parts I, II, III, IV, *Comm. Part. Diff. Equ.*, 17 and 18 (1992 and 1993).
- [9] I. Catto, C. Le Bris and P.-L. Lions. *Limite thermodynamique pour des modèles de type Thomas-Fermi*, *C.R. Acad. Sci. Paris Série I* **322** (1996) 357–364.
- [10] I. Catto, C. Le Bris and P.-L. Lions. *Mathematical theory of thermodynamic limits : Thomas-Fermi type models*, Oxford University Press 1998.
- [11] I. Catto, C. Le Bris and P.-L. Lions. Sur la limite thermodynamique pour des modèles de type Hartree et Hartree-Fock, *C.R. Acad. Sci. Paris, Série I* **327** (1998), 259–266.
- [12] I. Catto, C. Le Bris and P.-L. Lions. On the thermodynamic limit for Hartree-Fock type problems, *Ann. I. H. Poincaré* **18** (2001), 687–760.

- [13] P. Chaix and D. Iracane. *From quantum electrodynamics to mean field theory : I. The Bogoliubov-Dirac-Fock formalism*, *J. Phys. B.* **22** (1989), 3791–3814.
- [14] P. Chaix, D. Iracane, and P.L. Lions, *From quantum electrodynamics to mean field theory : II. Variational stability of the vacuum of quantum electrodynamics in the mean-field approximation*, *J. Phys. B.* **22** (1989), 3815–3828.
- [15] R. Dautray and J.-L. Lions. *Analyse mathématique et calcul numérique pour les sciences et les techniques*, volumes 1-3 (Masson, 1985).
- [16] R. Dovesi, R.Orlando, C. Roetti, C. Pisani and V.R. Saunders. The periodic Hartree-Fock method and its implementation in the Crystal code, *Phys. Stat. Sol. (b)* **217** (2000), 63–88.
- [17] R.M. Dreizler and E.K.U. Gross, *Density functional theory*, Springer, 2000.
- [18] C. Fefferman. The Thermodynamic Limit for a Crystal, *Comm. Math. Phys.* **98** (1985), no. 3, 289–311.
- [19] G. Floquet. Sur les équations différentielles linéaires à coefficients périodiques, *Ann. Ecole Norm. Sér. 2* **12** (1883) 47–89.
- [20] D. Gilbarg, N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, Reprint of the 1998 edition.
- [21] P. Gravejat, M. Lewin, E. Séré. Ground state and charge renormalization in a nonlinear model of relativistic atoms, To appear in *Comm. Math. Phys.* (2008).
- [22] Ch. Hainzl, M. Lewin and E. Séré. Existence of a stable polarized vacuum in the Bogoliubov-Dirac-Fock approximation, *Comm. Math. Phys.* **257** (2005), no. 3, 515–562.
- [23] Ch. Hainzl, M. Lewin and E. Séré. Self-consistent solution for the polarized vacuum in a no-photon QED model, *J. Phys. A : Math & Gen.* **38** (2005), no 20, 4483–4499.
- [24] Ch. Hainzl, M. Lewin and E. Séré. Existence of atoms and molecules in the mean-field approximation of no-photon quantum electrodynamics, To appear in *Arch. Rational Mech. Anal.*
- [25] Ch. Hainzl, M. Lewin and J.P. Solovej. The mean-field approximation in quantum electrodynamics. The no-photon case, *Comm. Pure Applied Math.* **60** (2007), no. 4, 546–596.
- [26] D. Hasler and J.P. Solovej. The Independence on Boundary Conditions for the Thermodynamic Limit of Charged Systems, *Comm. Math. Phys.* **261** (2006), no. 3, 549–568.
- [27] P. Hohenberg and W. Kohn. *Inhomogeneous electron gas*, *Phys. Rev. B* **136** (1964) 864–871.
- [28] W. Hunziker. On the Spectra of Schrödinger Multiparticle Hamiltonians. *Helv. Phys. Acta* **39** (1966), 451–462.
- [29] M. Klein, A. Martinez, R. Seiler and X.P. Wang. On the Born-Oppenheimer expansion for polyatomic molecules, *Comm. Math. Phys.* **143** (1992), 607–639.

- [30] W. Kohn and L.J. Sham. *Self-consistent equations including exchange and correlation effects*, Phys. Rev. A 140 (1965) 1133–1138.
- [31] E.H. Lieb. The Stability of Matter. *Rev. Mod. Phys.* **48** (1976), 553–569.
- [32] E.H. Lieb. Thomas-Fermi and related theories of atoms and molecules. *Rev. Mod. Phys.* **53** (1981), 603–641.
- [33] E.H. Lieb. *Density functional for Coulomb systems*, Int. J. Quant. Chem. 24 (1983) 243–277.
- [34] E.H. Lieb and J.L. Lebowitz. The constitution of matter : existence of thermodynamics for systems composed of electrons and nuclei. *Adv. Math.* **9** (1972), 316–398.
- [35] E.H. Lieb and M. Loss. *Analysis*, Second Edition. Graduate Studies in Mathematics, Vol. 14. American Mathematical Society, Providence, Rhode Island, 2001.
- [36] E.H. Lieb and B. Simon. The Hartree-Fock theory for Coulomb systems. *Comm. Math. Phys.* **53** (1977), 185–194.
- [37] E.H. Lieb and B. Simon. The Thomas-Fermi theory of atoms, molecules and solids. *Adv. Math.* **23** (1977), 22–116.
- [38] E.H. Lieb, J.P. Solovej and J. Yngvason. Asymptotics of heavy atoms in high magnetic fields. I. Lowest Landau band regions. *Comm. Pure Appl. Math.* **47** (1994), no. 4, 513–591.
- [39] P.-L. Lions. The concentration-compactness method in the Calculus of Variations. The locally compact case. Part. I : Anal. non-linéaire, Ann. IHP **1** (1984), 109–145. Part. II : Anal. non-linéaire, Ann. IHP **1** (1984), 223–283.
- [40] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, Vol I, Functional Analysis, Second Ed. Academic Press, New York, 1980.
- [41] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, Vol II, Fourier Analysis, Self-Adjointness, Second Ed. Academic Press, New York, 1975.
- [42] M. Reed and B. Simon. *Methods of Modern Mathematical Physics*, Vol IV, Analysis of Operators. Academic Press, New York, 1978.
- [43] D. Ruelle. *Statistical Mechanics. Rigorous results*. Imperial College Press and World Scientific Publishing, 1999.
- [44] E. Seiler and B. Simon. Bounds in the Yukawa₂ Quantum Field Theory : Upper Bound on the Pressure, Hamiltonian Bound and Linear Lower Bound. *Comm. Math. Phys.* **45** (1975), 99–114.
- [45] B. Simon. *Trace Ideals and their Applications*. Vol 35 of *London Mathematical Society Lecture Notes Series*. Cambridge University Press, 1979.
- [46] J.P. Solovej. Proof of the ionization conjecture in a reduced Hartree-Fock model, *Invent. Math.* **104** (1991), no. 2, 291–311.
- [47] L.E. Thomas. Time-Dependent Approach to Scattering from Impurities in a Crystal. *Comm. Math. Phys.* **33** (1973), 335–343.

- [48] C. Van Winter. Theory of Finite Systems of Particles. I. The Green function. *Mat.-Fys. Skr. Danske Vid. Selsk.* **2(8)** (1964).
- [49] G. M. Zhislin. A study of the spectrum of the Schrödinger operator for a system of several particles. (Russian) *Trudy Moskov. Mat. Obšč.* **9** (1960), 81–120.

Sur la modélisation numérique des cristaux

- [50] V. Bach, E.H. Lieb and J.P. Solovej. *General Hartree-Fock theory and the Hubbard model*, J. Statist. Phys. **76** 1-2 (1994), 3–89.
- [51] G.A. Baraff and M. Schlüter. *New self-consistent approach to the electronic structure and localized defects in solids*, Phys. Rev. B **19** 10 (1979), 4965–4979
- [52] F. A. Berezin. *The method of second quantization*, Academic Press, 1966.
- [53] J. Bernholc, N.O. Lipari and S.T. Pantelides. *Self-consistent method for point-defects in semi-conductors - Application to vacancy in silicon*, Phys. Rev. Lett. **41** 13 (1978), 895–899.
- [54] L. Boulton, *Limiting set of second order spectra*, Math. Comp. **75** (2006) 1367–1382.
- [55] E. Cancès. *Self-consistent field algorithms for Kohn-Sham models with fractional occupation numbers*, J. Chem. Phys. **114** 24 (2001) 10616–10622.
- [56] E. Cancès, I. Dabo and M. Lewin. In preparation.
- [57] E. Cancès, M. Defranceschi, W. Kutzelnigg, C. Le Bris and Y. Maday, *Computational Quantum Chemistry : a primer*, Handbook of Numerical Analysis, vol. X, Elsevier 2003, 3–270.
- [58] E. Cancès, A. Deleurence and M. Lewin. *Non-perturbative embedding of local defects in crystalline materials*, J. Phys. : Condens. Matter **20** (2008), 294213.
- [59] E. Cancès, K.N. Kudin and G.E. Scuseria. *A black-box self-consistent field convergence algorithm : One step closer*, J. Chem. Phys. **116** 19 (2002), 8255–8261.
- [60] Ch. Hainzl, M. Lewin, E. Séré and J.P. Solovej. *Minimization method for relativistic electrons in a mean-field approximation of quantum electrodynamics*, Phys. Rev. A **76** 5 (2007), no. 052104.
- [61] Ch. Kittel. *Quantum Theory of Solids*, Second Edition, Wiley, 1987.
- [62] W. Kohn and L.J. Sham. *Self-consistent equations including exchange and correlation effects*, Phys. Rev. A **140** (1965), 1133–1138.
- [63] G. Makov and M.C. Payne. *Periodic boundary conditions in ab initio calculations*, Phys. Rev. B **51** 7 (1995), 4014–4022
- [64] N. Marzari and D. Vanderbilt. *Maximally localized generalized Wannier functions for composite energy bands*, Phys. Rev. B **56** 20 (1997), 12847–12865.
- [65] R. Nada, C.R.A. Catlow, C. Pisani and R. Orlando. *An ab-initio Hartree-Fock perturbed-cluster study of neutral defects in LiF*, Modelling Simul. Mater. Sci. Eng. **1** (1993), 165–187.
- [66] M.C. Payne and M.I.J. Probert. *Improving the convergence of defect calculations in supercells : An ab initio study of the neutral silicon vacancy*, Phys. Rev. B **67** 7 (2003), no. 075204.

- [67] M. Peressi, R. Resta and C. Sgiarovello. *Electron localization in the insulating state : Application to crystalline semiconductors*, Phys. Rev. B **64** 11 (2001), no. 115202.
- [68] C. Pisani. *Quantum-mechanical treatment of the energetics of local defects in crystals : a few answers and many open questions*, Phase Transitions **52** (1994), 123–136.
- [69] C. Pisani. *Quantum-mechanical ab-initio calculation of the properties of crystalline materials*, Lecture Notes in Chemistry, Springer, 1996.
- [70] C. Pisani, R. Dovesi, R. Nada and L.N. Kantorovich, *Ab initio Hartree-Fock perturbed-cluster treatment of local defects in crystals*, J. Chem. Phys. **92** (12) (1990), 7448–7460.
- [71] S.N.M. Ruijsenaars. *Bogoliubov transformations for systems of relativistic charged particles*, J. Math. Phys. **18** 3 (1977), 517–526.
- [72] G. Scharf and H.P. Seipp. *Charged vacuum, spontaneous positron production and all that*, Phys. Lett. **108B** 3 (1982), 196–198.
- [73] P.A. Schultz. *Charged local defects in extended systems*, Phys. Rev. Lett. **84** 9 (2000), 1942–1945.
- [74] D. Shale and W.F. Stinespring. *Spinor representations of infinite orthogonal groups*, J. Math. and Mech. **14** (1965), 315–322.
- [75] S. Soussi. Convergence of the supercell method for defect modes calculations in photonic crystals, *SIAM J. Numer. Anal.* **43** (2005), no. 3, 1175–1201.
- [76] A.M. Stoneham. *Theory of Defects in Solids - Electronic Structure of Defects in Insulators and Semiconductors*. Oxford University Press, 2001.
- [77] P.H. Dederichs and R. Zeller. *Electronic structure of impurities in Cu, calculated self-consistently by Korringa-Kohn-Rostocker Greens function method*, Phys. Rev. Lett. **42** 25 (1979), 1713–1716.
- [78] <http://www.quantum-espresso.org/>