HABILITATION À DIRIGER LES RECHERCHES

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à l'Université Paris-Dauphine

Periodic and Half-Periodic Fermionic Systems

Spécialité MATHÉMATIQUES APPLIQUÉES

COMPOSITION DU JURY :

Mathieu Lewin
CNRS & Université Paris-Dauphine
Coordinateur

Phan Thánh Nam
Ludwig Maximilians Universität München
Rapporteur

Gianluca Panati
Sapienza Universitá di Roma
Rapporteur

Benjamin Schlein
Universität Zurich
Rapporteur

Clotilde Fermanian
Université Paris-Est

Yvon Maday
Sorbonne Université

Simona Rota Nodari
Université Côte d'Azur

Soutenue par **David Gontier**

le 15/01/2024

Coordonnée par **Mathieu Lewin**

Dauphine | PSL
UNIVERSITÉ PARIS

Université Paris-Dauphine

Unité de Recherche : **Centre de Recherche en Mathématiques de la Décision (CEREMADE, UMR 7534)**

Place de Marécal de Lattre de Tassigny, 75016 Paris

http://ceremade.dauphine.fr

Habilitation à Diriger les Recherches, présentée par **David Gontier**

Soutenue le **15/01/2024**

Discipline : **Mathématiques Appliquées**

Periodic and Half-Periodic Fermionic Systems

Composition du jury

Coordinateur	Mathieu Lewin	Université Paris–Dauphine
Rapporteurs	Phan Thánh Nam	LMU München
	Gianluca Panati	Sapienza Universitá di Roma
	Benjamin Schlein	Universität Zurich
Examinateurs.trices	Clotilde Fermanian	Université Paris-Est
	Yvon Maday	Sorbonne Université
	Simona Rota Nodari	Université Côte d'Azur

Résumé

Ce mémoire d'HDR comporte quatre parties :

- Nous faisons le lien entre l'existence des fonctions de Wannier en matière condensée, et l'existence d'homotopies pour une famille de projecteurs.
- Nous étudions l'apparition de mode de bords lorsqu'un matériau est coupé (systèmes semi-périodiques), et donnons un cadre général pour l'étude du spectre des opérateurs correspondants.
- Nous calculons quelques propriétés du diagramme de phase du gaz de fermions, dans l'approximation de Hartree-Fock.
- Nous étudions les inégalités de Lieb-Thirring de rang fini, et montrons un phénomène de cristallisation pour ses minimiseurs.

Abstract

This Habilitation consists in four parts:

- We link the existence of Wannier functions in condensed matter with the existence of homotopies for a family of projectors.
- We consider the emergence of edge modes when one cuts a material (half-periodic systems), and we give a general framework to study the spectrum of such operators.
- We describe some properties of the phase diagram of the fermionic gas, in the Hartree-Fock approximation.
- We study finite rank Lieb-Thirring inequalities, and exhibit a cristallisation phenomenon for its minimisers.

Mots clés : Analyse non linéaire, théorie spectrale, inégalités fonctionnelles, physique mathématiques, matière condensée, systèmes fermioniques

Keywords: Nonlinear analysis, spectral theory, fonctional inequality, mathematical physics, condensed matter, fermionic systems

CONTENTS

1 Low energy spectrum of periodic systems **8**
1.1 Introduction . . . 8
1.2 Applications . . . 9
1.2.1 Wannier functions . . . 9
1.2.2 Construction of reduced model . . . 10
1.2.3 Tight-binding models . . . 10
1.3 Construction of Wannier functions for insulators . . . 10
1.3.1 Results . . . 11
1.3.2 Proof of Theorem 1.1 with the column interpolation method . . . 12
1.3.3 Symmetries, and a simple derivation of the Kitaev's table . . . 14
1.4 Construction of Wannier functions in the general case . . . 16
1.4.1 Metallic systems . . . 16
1.4.2 Removing topological obstructions . . . 18
1.5 Brillouin zone integration . . . 20
1.5.1 The insulating case . . . 21
1.5.2 The metallic case . . . 21
1.6 Perspectives . . . 23

2 Semi-periodic systems **24**
2.1 Introduction . . . 24
2.2 From the cut Hamiltonian to the junction Hamiltonian . . . 25
2.2.1 Spectral flows . . . 25
2.2.2 The bulk/edge index . . . 26
2.2.3 Junctions between two Hamiltonians . . . 27
2.3 The dislocated model . . . 28
2.3.1 Exact computation in the case of dislocations . . . 28
2.3.2 The Grand Hilbert Hotel . . . 29
2.3.3 The spectrum of a general junction between two–dimensional Hamiltonians . . . 30
2.3.4 Application: two–dimensional materials cut with a general angle . . . 31
2.4 Perspectives . . . 34
2.4.1 Soft wall models . . . 34
2.4.2 Lagrangian setting . . . 34

3 The Hartree–Fock gas **37**
3.1 Introduction . . . 37
3.2 Phase diagram in the translation–invariant setting . . . 38
3.2.1 Reduction to the no-spin case . . . 39

3.2.2 The free Fermi gas at null temperature . . . 40
3.2.3 Phase transitions with temperature . . . 41
3.3 The Hartree–Fock gas versus the free Fermi gas . . . 44
3.3.1 A lower bound involving a degenerate operator . . . 45
3.3.2 Study of the degenerate operator . . . 47
3.4 Phase diagram in the Peierls model for polyacetylene . . . 50
3.5 Perspectives . . . 53

4 Lieb-Thirring (and related) inequalities **55**
4.1 Introduction . . . 55
4.2 Lieb-Thirring inequality . . . 56
4.2.1 The finite rank Lieb-Thirring inequality . . . 56
4.2.2 The periodic Lieb-Thirring inequality . . . 57
4.2.3 What is known and conjectured about Lieb–Thirring «best scenarii» . . . 58
4.3 Fermionic non-linear Schrödinger . . . 61
4.3.1 Fermionic minimization problems . . . 61
4.3.2 Dual formulation of Lieb-Thirring and NLS inequalities . . . 64
4.4 Perspectives . . . 65

Bibliography **66**

REMERCIEMENTS

Mes premiers remerciements vont à Mathieu Lewin, qui m'a très bien accueilli au Ceremade, aussi bien humainement que scientifiquement. Je lui suis très reconnaissant pour les nombreuses pauses café et les discussions autour du tableau noir, et d'avoir bien voulu partager avec moi aussi bien son savoir que sa vision de la recherche.

J'ai eu la chance de pouvoir discuter et travailler avec de nombreux collègues pendant ces dernières années. J'ai beaucoup appris lors de toutes ces rencontres. Je tiens notemment à remercier mes co-auteurs : Athman Bakhta, Horia Cornean, Jean Dolbeault, Virginie Ehrlacher, Rupert F. Frank, Louis Garrigue, Christian Hainzl, Salma Lahbabi, Antoine Levitt, Mathieu Lewin, Damiano Lombardi, Abdallah Maichine, Domenico Monaco, Faizan Q. Nazar, Fabio Pizzichillo, Éric Séré, Sami Siraj-Dine et Hanne Van Den Bosch. Merci tout particulièrement à Éric Cancès et Habib Ammari de m'avoir dirigé en thèse et en post-doc respectivement. Merci aussi à Adéchola Kouandé, Camilo Gómez Araya, Alfred Kirsch, Solal Perrin-Roussel et Thaddeus Roussigné de m'avoir choisi pour co-diriger leur thèse, officiellement ou non.

Ma recherche n'aurait pas été possible sans les conditions de travail exceptionnelles qu'offre l'Université Paris–Dauphine, et plus particulièrement mon laboratoire, le Ceremade. Je pense évidemment à la très bonne ambiance générale du laboratoire, aux généreuses décharges de cours et aux moyens humains et financiers mis en place pour les jeunes chercheurs et chercheuses à Dauphine, mais aussi aux excellents étudiants et étudiantes que j'ai eu la chance de rencontrer dans cette université. Merci aussi au DMA de l'ENS de m'avoir accueilli en demi-délégation pendant ces trois dernières années.

Merci à Étienne Lécroart pour son illustration du Grand Hilbert Hotel.

Je suis très reconnaissant envers Gianluca Panati, Phan Thành Nam et Benjamin Schlein, pour avoir accepté de rapporter mon habilitation, ainsi qu'à Clotilde Fermanian, Yvon Maday et Simona Rota Nodari, qui participent au jury.

Je remercie enfin ma femme Julia, pour son soutien constant, ainsi que mes deux enfants Félix et Valentin pour leurs sourires.

INTRODUCTION

Some researchers are able to present their work in a way that tells a cohesive story, with a clear motivation and logical sequence of questions and answers. In this case, the Habilitation becomes a story: there is a red threat motivating all the research presented in the manuscript, and the results are organised in a logical sequence. Unfortunately, we could not find a single red thread for what we wanted to present, so we decided to write four different chapters. They are all about infinite fermionic systems in mean–fields regimes, but they all tell a different story:

- In **Chapter 1**, we study crystalline structures from condensed matter. The goal is to describe the low energy spectrum of periodic Schrödinger operators. We explain the construction of Wannier functions, review part of the Kitaev table for topological insulators, and give methods to numerically compute physical quantities for such crystals.
- In **Chapter 2**, we study the spectrum of half materials, described by «semi»–periodic linear Schrödinger operators. We show that an *edge spectrum* appears when one cuts a crystalline structure, and that this one is linked to the presence of a *spectral flow*. We then give a general framework to study this edge spectrum in various contexts.
- In **Chapter 3**, we focus on the fermionic electron gas, under the Hartree–Fock approximation, and study its phase diagram. We show that there is a spin–symmetry breaking at low density and low temperature. We also show that there is no symmetry breaking at high density or high temperature.
- In **Chapter 4**, we study the existence of optimizers in Lieb-Thirring (and related) inequalities. We prove that, for some values of the parameters, a fermionic state with a finite number of particles can never be optimal. We provide numerical simulations suggesting that the optimal state might describe an infinite crystalline structure.

The main goal of the present manuscript is to introduce the works done by the author after his PhD. In order to avoid repeating already written articles, and to keep a good reading flow for the reader, we decided to skip most details of the proofs, and to only provide some ideas behind them, and some flavours of the techniques used. The reader can always refer to the corresponding articles for the details of the proofs, and for further remarks on the results. The style is therefore not rigorous, and we will prefer analogies and pictures rather than formal proofs to explain our results.

List of publications of the author

Articles described in this manuscript

Chapter 1:

- [G1] *Localised Wannier functions in metallic systems.* Horia Cornean, David Gontier, Antoine Levitt, Domenico Monaco. Annales Henri Poincaré 20(4), 1367–1391 (2019).
- [G2] *Numerical construction of Wannier functions through homotopy.* David Gontier, Antoine Levitt, Sami Siraj-Dine. Journal of Mathematical Physics 60(3), 031901 (2019).
- [G3] *Numerical quadrature in the Brillouin zone for periodic Schrödinger operators.* Éric Cancès, Virginie Ehrlacher, David Gontier, Antoine Levitt, Damiano Lombardi. Numerische Mathematik 144(3), 479–526 (2020).
- [G4] *Symmetric Fermi projections and Kitaev's table: topological phases of matter in low dimensions.* David Gontier, Domenico Monaco, Solal Perrin-Roussel. Journal of Mathematical Physics 63(4), 041902 (2022).

Chapter 2:

- [G5] *Edge states in ordinary differential equations for dislocations.* David Gontier. Journal of Mathematical Physics 61(4), 043507 (2020).
- [G6] *Edge states for second order elliptic operators in a channel.* David Gontier. Journal of Spectral Theory 12(3), 1155–1202 (2023).
- [G7] *Spectral properties of periodic systems cut with an angle.* David Gontier. Comptes Rendus Mathématique, 359(8), 949-958 (2021).

Chapter 3:

- [G8] *Spin symmetry breaking in the translation-invariant Hartree-Fock uniform electron gas.* David Gontier, Mathieu Lewin. SIAM Journal of Mathematical Analysis 51(4), 3388–3423 (2019).
- [G9] *Lower bound on the Hartree-Fock energy of the electron gas.* David Gontier, Christian Hainzl, Mathieu Lewin. Physical Review A 99(5), 052501 (2019).
- [G10] *Phase transition in the Peierls model for polyacetylene.* David Gontier, Adéchola Kouande, Éric Séré. Annales Henri Poincaré 24(11), 3945–3966 (2023).

Chapter 4:

- [G11] *The nonlinear Schrödinger equation for orthonormal functions: I. Existence of ground states.* David Gontier, Mathieu Lewin, Faizan Q. Nazar. Archive for Rational Mechanics and Analysis 240(3), 1203–1254 (2021).
- [G12] *The nonlinear Schrödinger equation for orthonormal functions: II. Application to Lieb-Thirring inequalities.* Rupert L. Frank, David Gontier, Mathieu Lewin. Communications in Mathematical Physics 384(3), 1783–1828 (2021).
- [G13] *The periodic Lieb-Thirring inequality.* R.L. Frank, D. Gontier, M. Lewin. In: Partial Differential Equations, Spectral Theory, and Mathematical Physics. The Ari Laptev Anniversary Volume, volume 18. EMS Publishing House (2021).
- [G14] *Optimizers for the finite-rank Lieb-Thirring inequality.* Rupert L. Frank, David Gontier, and Mathieu Lewin. accepted in American Journal of Mathematics.

Other articles and preprints

I will not mention the other following works (done after my PhD):

- [G15] *A mathematical and numerical framework for bubble meta-screens.* Habib Ammari, Brian Fitzpatrick, David Gontier, Hyundae Lee, Hai Zhang. SIAM Journal of Applied Mathematics 77(5), 1827-1850 (2017).
- [G16] *Sub-wavelength focusing of acoustic waves in bubbly media.* Habib Ammari, Brian Fitzpatrick, David Gontier, Hyundae Lee, Hai Zhang. Procceedings of the Royal Society A 473(2208), 20170469 (2017).
- [G17] *Minnaert resonances for acoustic waves in bubbly media.* Habib Ammari, Brian Fitzpatrick, David Gontier, Hyundae Lee, Hai Zhang. Annales de l'Institut Henri Poincaré C 35(7), 1975-1998 (2018).
- [G18] *The reduced Hartree-Fock model with self-generated magnetic fields.* David Gontier, Salma Lahbabi. Journal of Mathematical Physics 60(8), 081902 (2019).
- [G19] *Numerical reconstruction of the first band(s) in an inverse Hill's problem.* Athman Bakhta, Virginie Ehrlacher, David Gontier. ESAIM: Control, Optimisation and Calculus of Variations 26, 59 (2020).
- [G20] *Density Functional Theory for two-dimensional homogeneous materials.* David Gontier, Salma Lahbabi, Abdallah Maichine. Communications in Mathematical Physics 388(3), 1475-1505 (2021).
- [G21] *Density Functional Theory for two-dimensional homogeneous materials with magnetic fields.* David Gontier, Salma Lahbabi, Abdallah Maichine. Journal of Functional Analysis 285(9), 110100 (2023).
- [G22] *A simple derivation of moiré-scale continuous models for twisted bilayer graphene.* Éric Cancès, Louis Garrigue, David Gontier. Physical Review B 107(15), 155403 (2023).
- [G23] *Second-order homogenization of periodic Schrödinger operators with highly oscillating potentials.* Éric Cancès, Louis Garrigue, David Gontier. SIAM Journal on Mathematical Analysis 55(3), 2288-2323 (2023).
- [G24] *Keller and Lieb-Thirring estimates of the eigenvalues in the gap of Dirac operators.* Jean Dolbeault, David Gontier, Fabio Pizzichillo, Hanne Van Den Bosch. Revista Matemática Iberoamericana (2023).

CHAPTER 1

LOW ENERGY SPECTRUM OF PERIODIC SYSTEMS

1.1 Introduction

In this chapter, we study numerical methods to compute the properties of operators which commute with lattice translations. Such operators appear naturally in the context of condensed matter, where one studies the microscopic feature of crystalline structure. In most applications, we are only interested in the low-energy spectrum of such operators, so we can ask the following questions:

- How to encode efficiently the information contained in the low-energy spectrum?
- How to numerically compute some physical quantities, such as the energy per unit cell of a crystal?

The first question is linked to the construction of *Wannier functions*, and will be the main object of this Chapter. The second is linked to *Brillouin zone integration*, and will be studied in Section 1.5

Notation

Let us fix the notation of this Chapter. Our main object of interest is a self-adjoint operator H acting on $\mathcal{H} := \ell^2(\mathbb{Z}^d, \mathbb{C}^M)$ with $M \in \mathbb{N}$. We restrict ourselves to tight-binding models, as it is enough for our purpose, and simplifies some arguments, while keeping all the main ideas of the proofs. Most of our results can be generalized to the infinite dimensional case where the Hilbert space is $L^2(\mathbb{R}^d)$ (which somehow corresponds to $M = \infty$). After an undisplayed Bloch transform, we are lead to study an analytic family

$$\mathbb{T}^d \ni \mathbf{k} \mapsto H(\mathbf{k}) \in \mathcal{S}(\mathbb{C}^M).$$

Here, $\mathbb{T}^d \approx [0,1]^d$ denotes the d-dimensional torus and $\mathcal{S}(\mathcal{H})$ the set of self–adjoint operators on the Hilbert space $\mathcal{H}$. The set $\mathcal{S}(\mathbb{C}^M)$ can be identified with the set of $M \times M$ hermitian matrices.

Let $\varepsilon_{1,\mathbf{k}} \leq \varepsilon_{2,\mathbf{k}} \leq \cdots$ be the eigenvalues of $H(\mathbf{k})$, ranked in increasing order, and counting multiplicities. Let also $\{\psi_{n,\mathbf{k}}\}_{1\leq n\leq M}$ be a corresponding orthonormal basis of eigenvectors. The map $\mathbf{k} \mapsto \varepsilon_n(\mathbf{k})$ is called the n-th **Bloch band**. It is a classical result these maps are continuous [Kat13], and are analytic outside the band crossings. The set of crossings between the N-th and $(N+1)$-th Bloch bands is

$$K_N := \left\{ \mathbf{k} \in \mathbb{T}^d, \quad \varepsilon_{N,\mathbf{k}} = \varepsilon_{N+1,\mathbf{k}} \right\}.$$

If $\mathbf{k} \notin K_N$, then we have the strict inequality $\varepsilon_{N,\mathbf{k}} < \varepsilon_{N+1,\mathbf{k}}$, and one can define the spectral projector $P_N(\mathbf{k})$ onto the N first eigenvectors, namely

$$P_N(\mathbf{k}) := \sum_{n=1}^{N} |\psi_{n,\mathbf{k}}\rangle\langle\psi_{n,\mathbf{k}}| = \frac{1}{2\mathrm{i}\pi} \oint_{\mathscr{C}} \frac{\mathrm{d}z}{z - H_{\mathbf{k}}},$$

where we used the Cauchy residual formula in the last equality. Here, $\mathscr{C}$ is a positively oriented loop of the form $\mathscr{C} = (\Sigma+\mathrm{i}) \to (\Sigma-\mathrm{i}) \to (E-\mathrm{i}) \to (E+\mathrm{i}) \to (\Sigma+\mathrm{i})$, where $\Sigma < \varepsilon_{1,\mathbf{k}}$, and $\varepsilon_{N,\mathbf{k}} < E < \varepsilon_{N+1,\mathbf{k}}$. The curve $\mathscr{C}$ depends on $\mathbf{k}$, but, for a fixed curve $\mathscr{C}_0$ valid for $\mathbf{k}_0$, the Cauchy residual formula remains valid in a neighbourhood of $\mathbf{k}_0$, proving that $\mathbf{k} \mapsto P_N(\mathbf{k})$ is analytic on $\mathbb{T}^d \setminus K_N$.

For $\mathbf{k} \in K_N$, the projector $P_N(\mathbf{k})$ is not well-defined, due to the degeneracy of the N-th eigenvalue. On the other hand, if $K_N = \emptyset$, there is a (relative) gap above the N-th Bloch band. This is sometimes called the (relative) **insulating case**. In this case, P_N is analytic and periodic on the whole torus $\mathbb{T}^d$.

For all $\mathbf{k} \in \mathbb{T}^d \setminus K_N$, $P_N(\mathbf{k})$ is an orthogonal projector of rank N, hence an element of the *Grassmannian*

$$\mathcal{G}_N^M := \left\{ P \in \mathcal{S}(\mathbb{C}^M), \quad P^2 = P, \quad \mathrm{Tr}(P) = N \right\}.$$

Let us also introduce the set of *frames*

$$\mathcal{F}_N^M := \left\{ \Phi \in \mathcal{M}_{M\times N}(\mathbb{C}), \quad (\Phi^*)\Phi = \mathbb{I}_N \right\}.$$

A matrix Φ is in $\mathcal{F}_N^M$ iif its column vectors $\Phi = (\phi_1, \cdots, \phi_N)$ are orthonormal in $\mathbb{C}^M$. For $\Phi = (\phi_1, \cdots, \phi_N) \in \mathcal{F}_N^M$, one can associate the orthogonal projector P_Φ onto the vectorial space spanned by $\{\phi_1, \cdots, \phi_N\}$. Explicitly, $P_\Phi = \sum_{i=1}^N |\phi_i\rangle\langle\phi_i|$, which we write in matrix form as $P_\Phi = \Phi\Phi^*$. The map $\Phi \mapsto P_\Phi$ from $\mathcal{F}_N^M$ to $\mathcal{G}_N^M$ is well-defined and onto. However, it is not one-to-one: if $\Phi_0 \in \mathcal{F}_N^M$ is a frame for P, then, for any unitary $U \in \mathrm{U}(N)$, $\Phi = \Phi_0 U$ is also a frame for P.

1.2 Applications

Given a smooth and periodic family $P : \mathbb{T}^d \to \mathcal{G}_N^M$ of projectors, we say that P admits a global smooth frame $\Phi = (\phi_1, \cdots, \phi_N)$ if $P(\mathbf{k}) = \sum_{i=1}^N |\phi_i(\mathbf{k})\rangle\langle\phi_i(\mathbf{k})|$ for all $\mathbf{k} \in \mathbb{T}^d$. The main goal of this section is to construct such frames.

Before we turn to the results, let us explain why the construction of frames can provide efficient numerical tools to study the low-energy spectrum of a family of Hamiltonians.

1.2.1 Wannier functions

First, the constructions of frames gives a powerful way to numerically store the map $\mathbf{k} \mapsto P(\mathbf{k})$. Indeed, given $P : \mathbb{T}^d \to \mathcal{G}_N^M$, one could record the Fourier coefficients of this map, and write

$$P(\mathbf{k}) = \sum_{\mathbf{R}\in\mathbb{Z}^d} P_{\mathbf{R}} \mathrm{e}^{\mathrm{i}2\pi\mathbf{k}\cdot\mathbf{R}},$$

where each coefficient $P_{\mathbf{R}}$ is an $M \times M$ matrix. Using the relation $P(\mathbf{k}) = \Phi(\mathbf{k})\Phi^*(\mathbf{k})$, one can instead store the Fourier coefficients of the frame Φ itself, that is

$$\Phi(\mathbf{k}) = \sum_{\mathbf{R}\in\mathbb{Z}^d} W(\mathbf{R}) \mathrm{e}^{\mathrm{i}2\pi\mathbf{k}\cdot\mathbf{R}},$$

where each coefficient $W(\mathbf{R})$ is an $M \times N$ matrix. This is more efficient that storing $P_{\mathbf{R}}$ in the regime where $N \ll M$ (one stores N vectors of size M instead of a full $M \times M$ matrix), which is usually the case for real-life systems.

The columns of $W(\mathbf{R}) = (w_1(\mathbf{R}), \cdots, w_N(\mathbf{R}))$ are called the **Wannier functions**. Explicitly,

$$\forall \mathbf{R} \in \mathbb{Z}^d, \quad W(\mathbf{R}) = \int_{\mathbb{T}^d} \Phi(\mathbf{k}) \mathrm{e}^{-\mathrm{i}2\pi\mathbf{k}\cdot\mathbf{R}} \mathrm{d}\mathbf{k}, \quad \text{or} \quad w_n(\mathbf{R}) = \int_{\mathbb{T}^d} \phi_{n,\mathbf{k}} \mathrm{e}^{-\mathrm{i}2\pi\mathbf{k}\cdot\mathbf{R}} \mathrm{d}\mathbf{k}.$$

In the case where $\mathbf{k} \mapsto \Phi(\mathbf{k})$ is smooth, the Fourier coefficients $\mathbf{R} \mapsto W(\mathbf{R})$ (hence the Wannier functions) decay fast. They decay faster than any polynomial when $\mathbf{k} \mapsto \Phi(\mathbf{k})$ is C^∞, and exponentially fast when it is analytic.

1.2.2 Construction of reduced model

The construction of Wannier functions has applications in condensed matter physics, where one is interested in the low-energy spectrum of self-adjoint operators.

Assume one can construct a smooth frame $\Phi(\mathbf{k}) = \{\phi_{1,\mathbf{k}}, \cdots, \phi_{n,\mathbf{k}}\}$ for $P_N(\mathbf{k})$. Then one can also construct the $N \times N$ hermitian matrix $M(\mathbf{k})$ with coefficients

$$\forall 1 \leq i, j \leq N, \quad M_{ij}(\mathbf{k}) := \langle \phi_{i,\mathbf{k}}, H(\mathbf{k})\phi_{j,\mathbf{k}} \rangle .$$

This is the operator $H(\mathbf{k})$ projected onto $P_N(\mathbf{k})$. By construction, the map $\mathbf{k} \mapsto M(\mathbf{k})$ is smooth, and the min-max principle shows that the eigenvalues of $M(\mathbf{k})$ are exactly $\varepsilon_{1,\mathbf{k}} \leq \cdots \leq \varepsilon_{N,\mathbf{k}}$. In other words, the (small) matrix M reproduces exactly the low-energy spectrum of the (large) matrix H. The map $\mathbf{k} \mapsto M(\mathbf{k})$ therefore provides a **reduced model** for the low-energy properties of $\mathbf{k} \mapsto H(\mathbf{k})$.

To numerically compute the Fourier coefficients of the smooth periodic map $M(\mathbf{k})$, one can discretize the integral over $\mathbb{T}^d$ using a uniform coarse grid of size L^d. This provides accurate coefficients, with an error exponentially small in L (see Lemma 1.16 below). This is sometimes called *Wannier interpolation* [Yat+07].

1.2.3 Tight-binding models

Once the smooth and periodic map $\mathbf{k} \mapsto M(\mathbf{k})$ has been constructed, one can consider its Fourier coefficients, and write

$$M(\mathbf{k}) = \sum_{\mathbf{R} \in \mathbb{Z}^d} M_{\mathbf{R}} \mathrm{e}^{\mathrm{i}2\pi \mathbf{k}\cdot\mathbf{R}}.$$

Since $\mathbf{k} \mapsto M(\mathbf{k})$ is smooth, the Fourier coefficients $M_{\mathbf{R}}$, which are now $N \times N$ matrices, decay fast in $\mathbf{R}$. There is a nice interpretation for the coefficients $M_{\mathbf{R}}$. Indeed, consider the convolution operator $M : \ell^2(\mathbb{Z}^d, \mathbb{C}^N) \to \ell^2(\mathbb{Z}^d, \mathbb{C}^N)$ defined by

$$\forall \Psi = (\psi_{\mathbf{R}})_{\mathbf{R}\in\mathbb{Z}^d}, \quad (M\psi)_{\mathbf{R}_0} := \sum_{\mathbf{R}\in\mathbb{Z}^d} M_{\mathbf{R}}\psi_{\mathbf{R}_0-\mathbf{R}} = \sum_{\mathbf{R}\in\mathbb{Z}^d} M_{\mathbf{R}_0-\mathbf{R}}\psi_{\mathbf{R}}. \tag{1.1}$$

The operator M acting on the discrete Hilbert space $\ell^2(\mathbb{Z}^d, \mathbb{C}^N)$ is called a **tight binding model**. It models a periodic lattice system in which, at each location $\mathbf{R}_0 \in \mathbb{Z}^d$, a «particle» with N degrees of freedom can hop at a location $\mathbf{R} \in \mathbb{Z}^d$, with a transition matrix $M_{\mathbf{R}-\mathbf{R}_0}$. When the sum in (1.1) is truncated to $|\mathbf{R}| < R_0$, one obtains a model which is *local*. For instance, when the sum is truncated to $|\mathbf{R}|_\infty \leq 1$, we get a next-to-nearest neighbour model. This approximation is accurate whenever the coefficients $M_{\mathbf{R}}$ with $|\mathbf{R}|_\infty > R_0$ can be neglected.

Since the operator M is a discrete convolution operator, it can be diagonalized in Fourier space, and, if $\mathcal{F} : \ell^2(\mathbb{Z}^d, \mathbb{C}^N) \to L^2(\mathbb{T}^d, \mathbb{C}^N)$ denotes the Fourier transform, we have

$$\mathcal{F}M\mathcal{F}^* = \int_{\mathbb{T}^d}^{\oplus} M(\mathbf{k})\mathrm{d}\mathbf{k}, \quad \text{in the sense,} \quad \mathcal{F}(M\psi)(\mathbf{k}) = M(\mathbf{k})(\mathcal{F}\psi)(\mathbf{k}).$$

In particular, the tight-binding model reproduces the same low energy band diagrams as the original operator H.

1.3 Construction of Wannier functions for insulators

This section is based on ideas developed in [G2], in collaboration with Antoine Levitt and Sami Siraj–Dine.

We now explain how to construct fast decaying Wannier functions, or equivalently smooth frames $\Phi(\mathbf{k})$ for a given $P(\mathbf{k})$. Let us first focus on the insulating case. In this case, the map $P(\mathbf{k}) := P_N(\mathbf{k})$ is well–defined, analytical and periodic. So we consider a general smooth and periodic map of projectors

$$P = P(\mathbf{k}) \in C^\infty(\mathbb{T}^d, \mathcal{G}_N^M). \tag{1.2}$$

Here, we consider C^∞ functions instead of analytical, as we will construct C^∞ functions in the case of metallic system (Section 1.4). Some results of this chapter are valid in the analytical case. More specifically, if P admits a continuous frame, then it also admits a frame which has the same smoothness as P. One way to see this is as follows: if $\Phi_0 \in \mathcal{F}_N^M$ is a continuous frame for P, we first consider Ψ an analytical map on $\mathcal{M}_{M,N}(\mathbb{C})$ close to Φ_0 (take Ψ a trigonometric polynomial for instance, with $\|\Phi_0 - \Psi\|_\infty \leq \varepsilon$). Then, one can check that $\Phi_1 := P\Psi(\Psi^*\Psi)^{-1/2}$ is a frame for P, and has the same regularity as P.

1.3.1 Results

Our main result is a constructive proof of the following. This proof is given in the next section.

Theorem 1.1. *Let $d \geq 1$, let $1 \leq N < M < \infty$, and let $P \in C^\infty(\mathbb{T}^d, \mathcal{G}_N^M)$. Consider the following two assertions:*

(i) The map P is smoothly contractible: there is a smooth map $[0,1] \ni s \mapsto P_s \in C^\infty(\mathbb{T}^d, \mathcal{G}_N^M)$ so that $P_{s=0}(\mathbf{k}) = P_0$ and $P_{s=1}(\mathbf{k}) = P(\mathbf{k})$, where P_0 is any constant projector in $\mathcal{G}_N^M$.

(ii) There exists a global smooth frame for P, that is a smooth map $\Phi = \Phi(\mathbf{k}) = C^\infty(\mathbb{T}^d, \mathcal{F}_N^M)$ so that, in matrix form,

$$\forall \mathbf{k} \in \mathbb{T}^d, \quad P(\mathbf{k}) = \Phi(\mathbf{k})\Phi^*(\mathbf{k}), \quad \text{and} \quad \Phi^*(\mathbf{k})\Phi(\mathbf{k}) = \mathbb{I}_N.$$

Then we always have $(i) \implies (ii)$, with a frame Φ which is smoothly contractible. If in addition $d \leq 2(M-N)$, then $(ii) \implies (i)$ as well.

In practice, we have $M \gg N$, and $d \in \{1,2,3\}$, in which case the condition $d \leq 2(M-N)$ is satisfied. This leaves open the case $d = 3$ and $M = N+1$. Actually, the proof below will show that $(i) \iff (ii)$ also in the case $d = 3$ and $M \geq 3$. As for the remaining case $d = 3$ and $(N, M) = (1, 2)$, we have

Lemma 1.2 (Case $d = 3$ and $(N, M) = (1, 2)$)**.** *There exists $P \in C^\infty(\mathbb{T}^3, \mathcal{G}_1^2)$ which admits a global smooth frame, but which is not contractible.*

Proof. Let $\mathbf{v}(\mathbf{k}) : \mathbb{T}^3 \to \mathbb{S}^3$ be a smooth map which has a non-vanishing degree and let $\phi := (v_1 + \mathrm{i}v_2, v_3 + \mathrm{i}v_4)^T \in \mathbb{C}^2$ be its representation in the unit sphere of $\mathbb{C}^2$. The projector $P(\mathbf{k}) := |\phi(\mathbf{k})\rangle\langle\phi(\mathbf{k})|$ is a smooth rank-one projector in $\mathbb{C}^2$, which, by construction, admits the global smooth frame ϕ.

Assume P is contractible, and let $s \mapsto P_s$ be a contraction from $P(\mathbf{k})$ to the constant map $P_0 = |\phi_0\rangle\langle\phi_0|$. Using the map $s \mapsto P_s(\mathbf{k})$ to transport the frame $\phi(\mathbf{k})$ (see Lemma 1.4 below), we obtain a contraction $s \mapsto \phi_s$ from $\phi(\mathbf{k})$ to a frame of the form $\theta(\mathbf{k})\phi_0$, where $\mathbf{k} \mapsto \theta(\mathbf{k})$ is a smooth phase from $\mathbb{T}^3 \to \mathbb{S}^1$. Since the image of $\theta(\cdot)$ is one–dimensional, it cannot contribute to the three–dimensional degree, so $\deg_3(\mathbf{v}) = \deg_3(\phi) = \deg_3(\phi_0) = 0$, a contradiction. □

In low dimensions $d \in \{1,2,3\}$, one can characterize the maps P which are contractible. Let us recall the definition of the (first) Chern number. Let S be a compact connected manifold of dimension $d = 2$. The **Chern number** of $P \in C^\infty(S, \mathcal{G}_N^M)$ is

$$\mathrm{Ch}\,(P, S) := \frac{1}{2\mathrm{i}\pi} \int_S \mathrm{Tr}(P\mathrm{d}P \wedge \mathrm{d}P). \tag{1.3}$$

It is a well-known fact that $\mathrm{Ch}\,(P, S)$ is integer valued. It equals the winding number of an obstruction matrix, see for instance [G1].

Theorem 1.3.

- *(Case $d = 1$). Any $P(\mathbf{k}) \in C^\infty(\mathbb{T}^1, \mathcal{G}_N^M)$ is contractible.*
- *(Case $d = 2$). A map $P(\mathbf{k}) \in C^\infty(\mathbb{T}^2, \mathcal{G}_N^M)$ is contractible iff $\mathrm{Ch}\,(P, \mathbb{T}^2) = 0$*
- *(Case $d = 3$). If $P(\mathbf{k}) \in C^\infty(\mathbb{T}^2, \mathcal{G}_N^M)$ is contractible, then*

$$\mathrm{Ch}\,(P(0,\cdot,\cdot), \mathbb{T}^2) = \mathrm{Ch}\,(P(\cdot,0,\cdot), \mathbb{T}^2) = \mathrm{Ch}\,(P(\cdot,\cdot,0), \mathbb{T}^2) = 0.$$

If in addition $(N, M) \neq (1, 2)$, the converse holds.

Proof. Since the Chern number is integer valued, and has an expression continuous in P, this integer depends only on the homotopy class of P. In particular, if P is smoothly contractible, we must have $\mathrm{Ch}\,(P, S) = \mathrm{Ch}\,(P_0, S) = 0$.

The converse is proved by constructing a global frame for P. A complete construction can be found in [G2], following ideas from [CHN16; FMP16; Can+17]. □

Together with Theorem 1.1, we deduce that P admits global frame iff the Chern numbers of P vanishes (and $(d, N, M) \neq (3, 1, 2)$). This happens for instance whenever P is time-reversal symmetric (see Definition 1.6 below). Hence, for time–reversal symmetric systems, one can always construct Wannier functions.

This result appears first in the work of Panati [Pan07; Bro+07] in the context of condensed matter, following some preliminary works by Kohn [Koh59], des Cloizeaux [Clo64a; Clo64b], Nenciu [Nen83] and Hellfer–Sjöstrand [HS89] (where the authors studied the case $N = 1$). The first algorithm to construct such frames (or, equivalently, Wannier functions) is the Marzari–Vanderbild algorithm [MV97]. It was known to fail for some ill-prepared initial conditions. A method to construct well-prepared initial conditions was then proposed in [CHN16; FMP16], and implemented in [Can+17]. As we explain below, it uses the technique of parallel transport, and reduces the problem to contracting a family of unitaries $U(\mathbf{k}) : \mathbb{T}^{d-1} \to \mathrm{U}(N)$. This last problem was only partially answered in these previous works (but covers the important case of time-reversal symmetric systems). In our work [G2], we gave a general algorithm, that we called *column interpolation method*, to contract such families of unitaries in dimension $d \in \{1, 2, 3\}$. We illustrate this method in the next section.

1.3.2 Proof of Theorem 1.1 with the column interpolation method

We first record the following well-known Lemma by Levi-Civita.

Lemma 1.4 (Parallel transport of frames)**.** *Let $[0, 1] \ni s \mapsto P(s)$ be a C^1 curve of projectors in $\mathcal{G}_N^M$, and let Φ_0 be any matrix in $\mathcal{M}_{M,K}(\mathbb{C})$. Then, the (unique) solution to the Cauchy problem*

$$\begin{cases} \Phi'(s) & = [P'(s), P(s)]\,\Phi(s), \\ \Phi(s=0) & = \Phi_0, \end{cases} \tag{1.4}$$

satisfies $\Phi^*(s)\Phi(s) = \Phi^*(0)\Phi(0)$ and $\Phi^*(s)P(s)\Phi(s) = \Phi^*(0)P(0)\Phi(0)$. In particular,

(a) If Φ_0 is a frame for P_0, then $\Phi(s)$ is a frame for $P(s)$ for all s.

(b) If Φ_0 is an orthonormal family in $\mathrm{Ran}\, P_0^\perp$, then Φ_s is an orthonormal family in $\mathrm{Ran}\, P(s)^\perp$ for all s.

Proof. We first note that the adjoint matrix Φ^* satisfies $(\Phi^*)' = -\Phi^*[P', P]$. This gives

$$(\Phi^*\Phi)' = (\Phi^*)'\Phi + \Phi^*\Phi' = -\Phi^*[P', P]\Phi + \Phi^*[P', P]\Phi = 0,$$

which proves the first point. For the second point, we recall that $P^2 = P$, so $P'P + PP' = P'$. Multiplying by P on the left and on the right, we get $PP'P = 0$. This gives

$$\begin{aligned}(\Phi^* P\Phi)' &= (\Phi^*)'P\Phi + \Phi^* P'\Phi + \Phi^* P\Phi' = \left(-\Phi^*[P', P]\right) P\Phi + \Phi^* P'\Phi + \Phi^* P[P', P]\Phi \\ &= \Phi^* \left(-P'P + PP'P + P' + PP'P - PP'\right)\Phi = 0.\end{aligned}$$

We now prove (a). We have $\Phi_0^*\Phi_0 = \mathbb{I}_N$, hence $\Phi(s)^*\Phi(s) = \mathbb{I}_N$, so $\Phi(s)$ is a frame. We also have $\Phi_0^* P_0 \Phi_0 = \mathbb{I}_N$, hence $\Phi(s)^* P(s)\Phi(s) = \mathbb{I}_N = \Phi(s)^*\Phi(s)$. This implies $\Phi^*(s)(1 - P(s))\Phi(s) = 0$, so the matrix $A := (1 - P(s))\Phi(s)$ satisfies $A^*A = 0$. We deduce that $A = 0$, that is $P(s)\Phi(s) = \Phi(s)$. So $\Phi(s)$ is a frame for $P(s)$. The proof of (b) is similar. □

The results below will not use the fact that the transport is *parallel* (any transport would do). However, Equation (1.4) has at least two interesting features. First, the Cauchy problem can be solved numerically. In addition, by usual regularity theory of ordinary differential equations, if $(s, \mathbf{k}) \mapsto P_s(\mathbf{k})$ is smooth both in s and $\mathbf{k}$, and if $\Phi_0(\mathbf{k})$ is a smooth family of initial conditions, then the constructed frame $(s, \mathbf{k}) \mapsto \Phi(s, \mathbf{k})$ is also smooth in s and $\mathbf{k}$.

We also record the following Lemma.

Lemma 1.5. *If $N > d$, then any smooth map $\mathbf{v} : \mathbb{T}^d \to \mathbb{S}^N$ is smoothly contractible.*

Since the sphere is path–connected, $\mathbf{v}$ can be contracted to any vector of the sphere.

Proof. According to Sard's Lemma, and since $N > d$, the map $\mathbf{v}$ cannot be onto, and we can find $\mathbf{v}_0 \in \mathbb{S}^N$ so that $\mathbf{v} \neq -\mathbf{v}_0$. We set

$$\mathbf{v}_s := \frac{(1-s)\mathbf{v} + s\mathbf{v}_0}{\|(1-s)\mathbf{v} + s\mathbf{v}_0\|}.$$

The denominator never cancels since $\mathbf{v} \neq -\mathbf{v}_0$. Hence the map $s \mapsto \mathbf{v}_s$ is well-defined and provides a smooth contraction. □

We are now in position to prove Theorem 1.1.

Proof of Theorem 1.1. Assume first that there is a smooth map $P_s(\mathbf{k})$ connecting P_0 and $P(\mathbf{k})$. Choose a constant frame $\Phi_0(\mathbf{k}) = \Phi_0$ for P_0. Transporting the frame $\Phi_0(\mathbf{k})$ along $P_s(\mathbf{k})$, we obtain, at $s = 1$, a frame $\Phi_{s=1}(\mathbf{k})$ for $P(\mathbf{k})$, which is smooth in $\mathbf{k}$. So (i) always implies (ii), with a contractible frame Φ for P.

Conversely, let $\Phi(\mathbf{k})$ be a smooth frame for $P(\mathbf{k})$. We construct a contraction for $\Phi(\mathbf{k})$ by induction on N. Assume we already know how to contract a frame $\widetilde{\Phi}(\mathbf{k})$ of rank K to the constant frame $\widetilde{\Phi}_0 := \{\mathbf{e}_1, \cdots, \mathbf{e}_K\}$, for all $0 \leq K \leq N - 1$. We isolate the last vector (column) of $\Phi(\mathbf{k})$, and write

$$\Phi(\mathbf{k}) = \left(\widetilde{\Phi}(\mathbf{k}), \phi_N(\mathbf{k})\right).$$

By induction, there is a smooth map $\widetilde{\Phi}_s(\mathbf{k})$ connecting $\widetilde{\Phi}(\mathbf{k})$ to $\widetilde{\Phi}_0$. Denote by $\widetilde{P}_s(\mathbf{k}) := \widetilde{\Phi}_s(\mathbf{k})\widetilde{\Phi}_s^*(\mathbf{k})$ the corresponding projector of rank $N - 1$. For all $\mathbf{k} \in \mathbb{T}^d$, we use the family $s \mapsto \widetilde{P}_s(\mathbf{k})$ to transport the last vector $\phi_N(\mathbf{k})$. According to Lemma 1.4 (b), we obtain a map $\phi_{N,s}(\mathbf{k})$ so that

$$\Phi_s(\mathbf{k}) := \left(\widetilde{\Phi}_s(\mathbf{k}), \phi_{N,s}(\mathbf{k})\right)$$

is a frame for all $s \in [0, 1]$ and all $\mathbf{k} \in \mathbb{T}^d$. At $s = 1$, this frame is of form

$$\Phi_{s=1}(\mathbf{k}) = \{\mathbf{e}_1, \cdots, \mathbf{e}_{N-1}, \phi_{N,1}(\mathbf{k})\}.$$

It remains to contract $\phi_{N,1}(\mathbf{k})$ to the constant vector $\mathbf{e}_N$, while keeping orthogonality with the first $(N-1)$ other vectors. To do so, we notice that since $\phi_{N,1}(\mathbf{k})$ is orthogonal to $\{\mathbf{e}_1, \cdots, \mathbf{e}_{N-1}\}$, it is of the form

$$\Phi_{N,1}(\mathbf{k}) = (0, 0, \cdots, 0, \mathbf{v}(\mathbf{k})),$$

where $\mathbf{v}(\mathbf{k})$ is a smooth map from $\mathbb{T}^d$ to the unit sphere of $\mathbb{C}^{M-(N-1)}$, isomorphic to $\mathbb{S}^{2(M-N)+1}$. Since $d \leq 2(M-N)$, one can smoothly contract $\mathbf{v}(\mathbf{k})$ to $\mathbf{v}_0 = (1, 0, \cdots, 0) \in \mathbb{S}^{2(M-N)+1}$ by Lemma 1.5. This map contracts $\Phi_{N,1}(\mathbf{k})$ to $\mathbf{e}_N$ as wanted. Concatenating the homotopies, we obtain a smooth contraction of frames $\Phi_s(\mathbf{k})$ connecting $\Phi(\mathbf{k})$ to the constant frame Φ_0. The map $s \mapsto P_s(\mathbf{k}) := \Phi_s(\mathbf{k})\Phi_s^*(\mathbf{k})$ gives the desired homotopy.

Case $d = 3$. We now focus on the case $d = 3$. Our goal is to prove that if $M \geq 3$, then the existence of a frame implies that P is contractible. We already proved that it was the case if $d \geq 2(M-N)$. When $d = 3$, this is not satisfied for $N = M$ or $M = N+1$. In the first case $M = N$, there is a unique projector $P \in \mathcal{G}_M^M$, which is the identity operator, and the result is trivial. It remains to study the case $M = N+1$. Assume $P(\mathbf{k})$ has a smooth frame $\Phi(\mathbf{k}) = (\phi_1(\mathbf{k}), \cdots, \phi_{M-1}(\mathbf{k}))$. Consider the (Hodge dual of the) exterior product

$$\phi_M(\mathbf{k}) := \phi_1(\mathbf{k}) \wedge \phi_2(\mathbf{k}) \wedge \cdots \wedge \phi_{M-1}(\mathbf{k}).$$

It is the element in $\mathbb{C}^M$ so that, for all $v \in \mathbb{C}^M$, $\langle \phi_M, v\rangle_{\mathbb{C}^M} = \det(\phi_1, \cdots, \phi_{M-1}, v)$. The map $\mathbf{k} \mapsto \phi_M(\mathbf{k})$ is a smooth frame for the rank–1 projector $Q(\mathbf{k}) = 1 - P(\mathbf{k}) = P^\perp(\mathbf{k})$. Note that P is contractible iff $Q := Q^\perp$ is contractible. We now apply the previous result for the rank–1 projector $Q(\mathbf{k})$. When $d = 3 < 2(M-1)$, that is when $M \geq 3$, $Q(\mathbf{k})$ admits the frame $\phi_M(\mathbf{k})$, hence is smoothly contractible. So $P(\mathbf{k}) = 1 - Q(\mathbf{k}) = Q^\perp(\mathbf{k})$ is also contractible. The case $d = 3$, $M = 2$ and $N = 1$ was considered in Lemma 1.2. □

The idea to transport the columns one by one to construct homotopies has numerous applications. In [G2], we used it to give a constructive proof of the well-known fact that a map $\mathbb{T}^1 \ni t \mapsto U(t)$ of unitaries is contractible iff the winding number of $t \mapsto \det(U(t))$ vanishes.

1.3.3 Symmetries, and a simple derivation of the Kitaev's table

In this section, we describe the results in [G4]. This is joint work with Domenico Monaco and Solal Perrin-Roussel.

In the previous section, we proved that the existence of global frames is linked to the homotopy class of P. We now study these classes, and ask under which conditions a family of projectors P_0 can be smoothly deformed in another family of projectors P_1. In practice, the projectors P come from condensed matter problems, and the variable $\mathbf{k}$ is the Bloch momentum. In this case, some symmetries of the system induce symmetries on the map P. There are three important symmetries that one often encounters. Recall that our Hilbert space is $\mathcal{H} \sim \mathbb{C}^M$, and that a map $T : \mathcal{H} \to \mathcal{H}$ is anti-unitary if it is anti-linear $T(\lambda x) = \overline{\lambda} T(x)$ and

$$\forall x, y \in \mathcal{H}, \quad \langle Tx, Ty\rangle_{\mathcal{H}} = \langle y, x\rangle_{\mathcal{H}} \qquad (= \overline{\langle x, y\rangle_{\mathcal{H}}}).$$

Definition 1.6 (Time-reversal symmetry)**.** *Let $T : \mathcal{H} \to \mathcal{H}$ be an anti-unitary operator such that $T^2 = \varepsilon_T \mathbb{I}_{\mathcal{H}}$ with $\varepsilon_T \in \{-1, 1\}$. We say that a $P : \mathbb{T}^d \to \mathcal{G}_N^M$ satisfies* **time-reversal symmetry** *if*

$$T^{-1} P(\mathbf{k}) T = P(-\mathbf{k}), \qquad \textit{(T-symmetry)}.$$

If $\varepsilon_T = 1$, this T-symmetry is said to **even***, and if $\varepsilon_T = -1$ it is* **odd***.*

Definition 1.7 (Charge-conjugation/particle-hole symmetry)**.** *Let $C : \mathcal{H} \to \mathcal{H}$ be an anti-unitary operator such that $C^2 = \varepsilon_C \mathbb{I}_{\mathcal{H}}$ with $\varepsilon_C \in \{-1, 1\}$. We say that a continuous map $P : \mathbb{T}^d \to \mathcal{G}_n(\mathcal{H})$ satisfies* **charge-conjugation symmetry** *(also called* **particle-hole symmetry***) if*

$$C^{-1}P(\mathbf{k})C = \mathbb{I}_{\mathcal{H}} - P(-\mathbf{k}), \quad \textit{(C-symmetry)}.$$

If $\varepsilon_C = 1$, this C-symmetry is said to **even***, and if $\varepsilon_C = -1$ it is* **odd***.*

Definition 1.8 (Chiral symmetry)**.** *Let $S : \mathcal{H} \to \mathcal{H}$ be a unitary operator such that $S^2 = \mathbb{I}_{\mathcal{H}}$. We say that $P : \mathbb{T}^d \to \mathcal{G}_n(\mathcal{H})$ satisfies* **chiral** *or* **sublattice symmetry***, or in short* **S-symmetry***, if*

$$S^{-1}P(\mathbf{k})S = \mathbb{I}_{\mathcal{H}} - P(\mathbf{k}), \quad \textit{(S-symmetry)}.$$

In general, it is assumed that when two symmetries are present, then so is the third one. For instance, when $T-$ and $C-$ symmetries are both present, we assume that the product $S := TC$ is an S-symmetry. This is equivalent to require that the operators T and C commute or anti-commute among each other, depending on their even/odd nature. Indeed, the product $S := TC$ satisfies $S^2 = \mathbb{I}_{\mathcal{H}}$ iff

$$TCTC = \mathbb{I}_{\mathcal{H}} \quad \Longleftrightarrow \quad TC = C^{-1}T^{-1} = \varepsilon_T \varepsilon_C CT.$$

Taking into account all possible types of symmetries leads to 10 symmetry classes for maps $P : \mathbb{T}^d \mapsto \mathcal{G}_N(\mathcal{H})$, the famous *tenfold way of topological insulators* [Ryu+10], following [AZ97; HHZ05]. The names of these classes are given in Table 1.1, and are taken from the original works of E. Cartan [Car26; Car27] for the classification of symmetric spaces. For a dimension $d \in \mathbb{N}_0$, a rank $N \in \mathbb{N}$, and a Cartan label X of one of these 10 symmetry classes, we denote by $\mathrm{X}(d, N, M)$ the set of continuous maps $P\colon \mathbb{T}^d \to \mathcal{G}_N^M$ respecting the symmetry requirements of class X.

Given two continuous maps $P_0, P_1 \in \mathrm{X}(d, N, M)$, we ask whether we can find explicit index maps from $\mathrm{X}(d, N, M)$ to $\mathbb{Z}$ or $[\mathbb{Z} \mod 2]$, so that $\mathrm{Index}(P_0) = \mathrm{Index}(P_1)$ iff P_0 and P_1 are path-connected in $\mathrm{X}(d, N, M)$?

In [G4], we answer this problem (and construct explicit homotopies when possible) for all 10 symmetry classes, and for $d \in \{0, 1\}$. The topological indices that we find are summarized in Table 1.1.

Symmetry				Constraints		Indices		Cartan symmetric space
Cartan label	T	C	S	N	M	$d=0$	$d=1$	
A	0	0	0			0	0	$U(M)/\mathrm{U}(N) \times \mathrm{U}(M-N)$
AIII	0	0	1		$M = 2N$	0	$\mathbb{Z}$	$\mathrm{U}(N)$
AI	1	0	0			0	0	$\mathrm{O}(M)/\mathrm{O}(N) \times \mathrm{O}(M-N)$
BDI	1	1	1		$M = 2N$	$\mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}$	$\mathrm{O}(N)$
D	0	1	0		$M = 2N$	$\mathbb{Z}_2$	$\mathbb{Z}_2 \times \mathbb{Z}_2$	$\mathrm{O}(2N)/\mathrm{U}(N)$
DIII	-1	1	1	$N = 2n$	$M = 2N = 4n$	0	$\mathbb{Z}_2$	$\mathrm{U}(2n)/\mathrm{Sp}(n)$
AII	-1	0	0	$N = 2n$	$M = 2m \in 2\mathbb{N}$	0	0	$\mathrm{Sp}(m)/\mathrm{Sp}(n) \times \mathrm{Sp}(m-n)$
CII	-1	-1	1	$N = 2n$	$M = 2N = 4n$	0	$\mathbb{Z}$	$\mathrm{Sp}(n)$
C	0	-1	0		$M = 2N$	0	0	$\mathrm{Sp}(N)/\mathrm{U}(N)$
CI	1	-1	1		$M = 2N$	0	0	$\mathrm{U}(N)/\mathrm{O}(N)$

Table 1.1: The Kitaev table in dimension $d \in \{0, 1\}$. For $d = 0$, we can identify each class with a *Cartan symmetric space*.

The results of [G4] are not new, but we gave a comprehensive and constructive proof for this table, using simple linear algebra. As an example of the strategy used, let us prove the AIII($d = 1$) part of the table. We assume that there is a unitary $S : \mathcal{H} \to \mathcal{H}$ with $S^2 = \mathbb{I}_{\mathcal{H}}$ so that $S^{-1}P(\mathbf{k})S = \mathbb{I}_{\mathcal{H}} - P(\mathbf{k})$ for all $\mathbf{k} \in \mathbb{T}^d$.

Lemma 1.9. *If* AIII(d, N, M) *is non empty, then* $M = 2N$*, and there is a basis of* $\mathcal{H}$ *in which* S *has the block structure*

$$S = \begin{pmatrix} \mathbb{I}_N & 0 \\ 0 & -\mathbb{I}_N \end{pmatrix}. \tag{1.5}$$

In this basis, a projection $P \in \mathcal{G}_N^{2N}$ *satisfies* $S^{-1}PS = \mathbb{I} - P$ *iff it has the block structure*

$$P = \frac{1}{2}\begin{pmatrix} \mathbb{I}_N & Q \\ Q^* & \mathbb{I}_N \end{pmatrix}, \quad \textit{with} \quad Q \in \mathrm{U}(N).$$

Proof. If AIII(d, N, M) is non empty, there is $P_0 \in \mathcal{G}_N^M$ so that $S^{-1}P_0S = \mathbb{I} - P_0$. Since P_0 is unitarily equivalent to $\mathbb{I} - P_0$, we have $M = 2N$. Let $(\psi_1, \cdots, \psi_N)$ be an orthonormal basis for $\mathrm{Ran}(P_0)$. We set, for $i \in \{1, \cdots, N\}$,

$$\phi_i := \frac{1}{\sqrt{2}}\big(\psi_i + S\psi_i\big), \qquad \phi_{N+i} := \frac{1}{\sqrt{2}}\big(\psi_i - S\psi_i\big).$$

We can check that $(\psi_1, \cdots, \psi_{2N})$ is an orthonormal basis for $\mathcal{H}$, and in this basis, S has the form (1.5).

Let P be any projector in $\mathcal{G}_N^{2N}$ so that $S^{-1}PS = \mathbb{I} - P$, and write

$$P = \frac{1}{2}\begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}.$$

The equation $S^{-1}PS = \mathbb{I} - P$ implies that $P_{11} = P_{22} = \mathbb{I}_N$, and the equation $P^2 = P$ shows that $Q := P_{12}$ is unitary. □

Since the map $P \mapsto Q$ is one-to-one, any map $P : \mathbb{T}^d \to$ AIII$(d = 0, N, M)$ translates into a map $Q : \mathbb{T}^d \to \mathrm{U}(N)$ (this is the $\mathrm{U}(N)$ appearing in the last column of our Table 1.1 for AIII). In dimension $d = 1$, it is a classical result that such maps are contractible iff the Winding number of $\det(Q)$ vanishes. Actually, the $\mathbb{Z}$ appearing in the Kitaev table for AIII$(1, N, 2N)$ corresponds to the index

$$\text{Index} : P \in \text{AIII}(d = 1, N, M) \mapsto \text{Winding}\,\Big(\det(Q), \mathbb{T}^1\Big) \in \mathbb{Z}.$$

The other cells of the table are constructed in a similar way.

1.4 Construction of Wannier functions in the general case

We now generalize the previous results to metallic systems and to projectors which are not contractible. The main message of this section is that, by allowing P to be over–represented, that is by allowing extra vectors in the frames, one can lift band crossings and/or topological obstructions.

1.4.1 Metallic systems

In this section, we describe the results in [G1]. This is joint work with Horia D. CORNEAN, Domenico MONACO and Antoine LEVITT.

We now come back to the setting described in the introduction. Recall that $P_N(\mathbf{k})$ is the projection on the N first eigenvectors. It is not defined on the crossing set $K_N := \Big\{\mathbf{k} \in \mathbb{T}^d,\ \varepsilon_{N,\mathbf{k}} = \varepsilon_{N+1,\mathbf{k}}\Big\}$, but is smooth (analytic) on $\mathbb{T}^d \setminus K_N$.

We say that K_N satisfies **Assumption A** if K_N is a finite union of isolated points and piecewise smooth curves. This assumption is satisfied for most systems in practice. The main result of [G1] is the following.

Theorem 1.10 ([G1], Theorem 2.1)**.** *Let $d = 3$. Assume K_N and K_{N+1} satisfies Assumption A, and that $K_N \cap K_{N+1} = \emptyset$. Then there exists $P \in C^\infty(\mathbb{T}^3, \mathcal{G}_{N+1}^M)$ so that, for all $\mathbf{k} \notin K_N$, we have $P_N(\mathbf{k}) \subset P(\mathbf{k})$.*

In other words, we add one vector to gain smoothness.

Remark 1.11. *As will be clear from the proof, if $K_N \cap K_{N+1} \neq \emptyset$, then we look for an integer $L \in \mathbb{N}$ so that $K_N \cap K_{N+L} \neq \emptyset$, and we need to add L extra vectors.*

Remark 1.12. *We do not know* a priori *that the constructed map $P(\mathbf{k})$ is smoothly contractible. However, in the important* time-reversal symmetric case *(TRS), it is possible to construct P to be TRS as well, in which case it is always contractible.*

As discussed in Section 1.2, this Theorem has direct numerical applications. Indeed, for numerical purpose, the cost of adding a vector and to represent P instead of P_N is irrelevant, and since the information of P_N is contained in P, one does not loose any information. In the case where P admits a global frame, one can create a reduced model $M(\mathbf{k})$ of size $(N+1) \times (N+1)$. By the min–max principle, the N lowest eigenvalues of $M(\mathbf{k})$ coincide with the ones of the initial operator $H(\mathbf{k})$.

The idea to *add functions* to handle metallic systems was first considered algorithmically by Souza, Marzari and Vanderbild [SMV01], and later by Damle, Levitt, Lin and Ying in their SCDM algorithm [DLY17; DLL19]. In some sense, our Theorem shows that it is indeed possible to smooth the projectors by adding functions («the set of solutions is non empty»).

Before we turn to the proof, we need a Lemma. Let Ω be a compact manifold of dimension $d \geq 2$, and let $P : \Omega \to \mathcal{G}_N^M$ be smooth. The **Berry curvature** of P is the two-form

$$\mathcal{F}[P] := -\mathrm{i}\mathrm{Tr}\,(P\mathrm{d}P \wedge \mathrm{d}P) = -\mathrm{i} \sum_{1 \leq \alpha < \beta \leq d} \mathrm{Tr}\,(P[\partial_\alpha P, \partial_\beta P])\,\mathrm{d}k_\alpha \wedge \mathrm{d}k_\beta.$$

When $\Omega = S$ is two-dimensional, the Chern number of P defined in (1.3) is the integral of $\mathcal{F}[P]$ over S, up to a 2π factor. Let $\Phi(\mathbf{k})$ be a local frame for $P(\mathbf{k})$ around some $\mathbf{k}_0$ (local frames always exist), a computation shows that

$$\mathcal{F}[P] = \mathrm{d}\mathcal{A}[\Phi], \quad \text{where} \quad \mathcal{A}[\Phi] := -\mathrm{i}\mathrm{Tr}\,(\Phi^* \mathrm{d}\Phi)$$

is the (trace of the) **Berry connection** of Φ. Although the Berry connection depends on the frame Φ, the Berry curvature $\mathrm{d}\mathbf{A}[\Phi]$ only depends on $P = \Phi\Phi^*$, the corresponding projector. We deduce the following facts.

Lemma 1.13.

- *(Bianchi's identity) The Berry curvature is a closed form:* $\mathrm{d}\mathcal{F}[P] = 0$.
- *If P, Q are projectors so that $P \perp Q$, then $\mathcal{F}[P+Q] = \mathcal{F}[P] + \mathcal{F}[Q]$.*

The first point comes from the fact that $\mathrm{d}\mathcal{F} = \mathrm{d}^2\mathcal{A} = 0$. For the second point, if Φ_P and Φ_Q are (local) frames for P and Q respectively, then (Φ_P, Φ_Q) is a local frame for $P+Q$, and we get $\mathcal{A}[(\Phi_P, \Phi_Q)] = \mathcal{A}[\Phi_P] + \mathcal{A}[\Phi_Q]$.

We can now provide some ideas behind the proof of Theorem 1.10.

Proof. If $N = M - 1$, we can choose $P = \mathbb{I}_{\mathbb{C}^M}$, so we only need to focus on the case $N \leq M - 2$.

The main idea is to cut the torus $\mathbb{T}^d$ appropriately around K_{N+1}. Since $K_{N+1} \cap K_N = \emptyset$, we can find a smooth connected open set $\Omega \subset \mathbb{T}^3$ so that $K_{N+1} \subset \Omega$ and $K_N \subset \mathbb{T}^3 \setminus \Omega$. Our Assumption A ensures that Ω can be chosen to be a δ–neighbourhood of some one-dimensional connected skeleton G (see Figure 1.1).

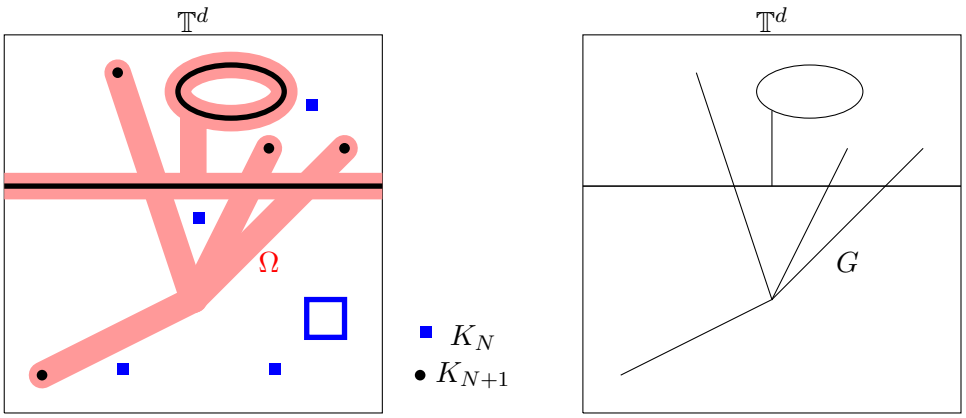


Figure 1.1: 2D sketch of the set $\Omega \subset \mathbb{T}^d$ enclosing the set K_{N+1} but avoiding the set K_N, and its skeleton G.

Since $K_{N+1} \subset \Omega$, the projector P_{N+1} is well-defined and smooth on $\mathbb{T}^3 \setminus \Omega$. On the other hand, since $K_N \subset \mathbb{T}^3 \setminus \Omega$, the map P_N is well defined and smooth on Ω. We therefore look for an operator $P(\mathbf{k})$ of the form

$$P(\mathbf{k}) := \begin{cases} P_{N+1}(\mathbf{k}) & \text{for} \quad \mathbf{k} \notin \Omega \\ P_N(\mathbf{k}) + p(\mathbf{k}) & \text{for} \quad \mathbf{k} \in \Omega. \end{cases}$$

Our goal is to construct a smooth map $p(\mathbf{k})$ on Ω, which is orthogonal to $P_N(\mathbf{k})$, and which equals $\widetilde{p}(\mathbf{k}) := P_{N+1}(\mathbf{k}) - P_N(\mathbf{k})$ at the boundary $\partial\Omega$.

Let us explain how to construct an extension of $\widetilde{p}$ in Ω. First, we have, by Lemma 1.13,

$$\mathrm{Ch}\,(\widetilde{p}, \partial\Omega) = \mathrm{Ch}\,(P_{N+1}, \partial\Omega) - \mathrm{Ch}\,(P_N, \partial\Omega)\,.$$

Since P_{N+1} is well-defined on $\mathbb{T}^3 \setminus \Omega$, we have, using Stokes' theorem and the fact that $\mathrm{d}\mathcal{F} = 0$ (see again Lemma 1.13),

$$\mathrm{Ch}\,(P_{N+1}, \partial\Omega) = \frac{1}{2\pi} \int_{\partial(\mathbb{T}^3 \setminus \Omega)} \mathcal{F}[P_{N+1}] = \frac{\pm 1}{2\pi} \int_{\mathbb{T}^3 \setminus \Omega} \mathrm{d}\mathcal{F}[P_{N+1}] = 0.$$

Similarly, since P_N is well-defined on Ω, we have $\mathrm{Ch}\,(P_N, \partial\Omega) = 0$, and we deduce that $\mathrm{Ch}\,(\widetilde{p}, \partial\Omega) = 0$ as well. This implies that there is a smooth frame $\widetilde{\phi}$ for $\widetilde{p}$ on $\partial\Omega$, which is contractible (this is a combination of Theorems 1.1 and 1.3). It remains to extend this frame to the interior of Ω, while keeping orthogonality with P_N. We refer to [G1] for the end of the proof.

□

1.4.2 Removing topological obstructions

In the previous section, we explained how to handle band crossing problems by adding vectors to the frame. For completeness, we now show how to use similar ideas to lift topological obstruction by adding one vector and two dimensions. This idea is not new, and be found for instance in [CMM19]. Our method is different, and relies on the following result (stated here in dimension $d = 2$).

Theorem 1.14. *For all $m \in \mathbb{Z}$, there is an analytic map $P_m : \mathbb{T}^2 \to \mathcal{G}_1^2$ so that $\mathrm{Ch}\,(P_m, \mathbb{T}^2) = m$.*

The fact that such a map exists follows from general considerations. Here, we provide a formula for P_m, which can be implemented easily.

Proof. If $P_1(k_1, k_2)$ satisfies $\mathrm{Ch}\,(P_1, \mathbb{T}^2) = 1$, then $P_m(k_1, k_2) := P_1(k_1, mk_2)$ satisfies $\mathrm{Ch}\,(P_m, \mathbb{T}^2) = m$ for all $m \in \mathbb{Z}$. So we only need to focus on the case $m = 1$.

Let us first present an idea that will fail, but will guide us in the construction of P_1. We define

$$\phi(k_1,k_2) := \sum_{n\in\mathbb{Z}} \varphi(k_1-n)\mathrm{e}^{-\mathrm{i}2\pi k_2(k_1-n)}, \quad \text{with} \quad \varphi(x) := \mathrm{e}^{-\pi x^2}.$$

One recognizes a Bloch transform of φ, where k_1 plays the role of x, and $(-2\pi k_2)$ the role of the Bloch momentum. By construction, ϕ is analytic and satisfies

$$\phi(k_1+1,k_2) = \phi(k_1,k_2), \quad \text{and} \quad \phi(k_1,k_2+1) = \mathrm{e}^{-\mathrm{i}2\pi k_1}\phi(k_1,k_2). \tag{1.6}$$

Assume that $\|\phi\| \neq 0$ on $\mathbb{T}^1\times[0,1]$, and let P be the 1-dimensional projector on Ran $\{\phi\}$. Then $\phi/\|\phi\|$ would be a continuous frame for P on the two–dimensional cut torus $\mathbb{T}^1\times[0,1]$. Note that this frame is not continuous over the full torus $\mathbb{T}^2$, since $\phi(k_1,1)\neq\phi(k_1,0)$. It is a classical result that the Chern of P is the winding of the obstruction of ϕ (see for instance [G1]), that is

$$\mathrm{Ch}\,(P,\mathbb{T}^2) = \mathrm{Winding}\,(U(k_1),\mathbb{T}^1), \quad \text{with} \quad U(k_1) = \left\langle \frac{\phi(k_1,1)}{\|\phi(k_1,1)\|}, \frac{\phi(k_1,0)}{\|\phi(k_1,0)\|}\right\rangle.$$

Using (1.6) we find that $U(k_1) = \mathrm{e}^{\mathrm{i}2\pi k_1}$, which has a winding equal to 1. So $\mathrm{Ch}\,(P,\mathbb{T}^2) = 1$ as wanted.

Unfortunately, we cannot have $\|\phi\| \neq 0$ everywhere on $\mathbb{T}^1\times[0,1]$. One reason is that P is a projector of rank 1 on $\mathbb{C}$, so is the identity, hence has $\mathrm{Ch}\,(P,\mathbb{T}^2) = 0$. One therefore rather considers the two–dimensional vector

$$\widetilde{\phi}(\mathbf{k}) := \begin{pmatrix} \phi(\mathbf{k}) \\ \phi(\mathbf{k}+\mathbf{K}) \end{pmatrix},$$

where $\mathbf{K}$ is a fixed vector of the form $\mathbf{K} = (K_1,0)$, chosen so that $\|\widetilde{\phi}(\mathbf{k})\| \neq 0$ over the full space $\mathbf{k}\in\mathbb{R}^2$. Now, the projector $P_1 : \mathbb{T}^2 \to \mathcal{G}_1^2$ onto $\mathrm{Vect}(\widetilde{\phi})$ is a well-defined rank–1 projector on $\mathbb{C}^2$, and has $\mathrm{Ch}\,(P_1,\mathbb{T}^2) = 1$ by the previous reasoning. □

Remark 1.15. *One can start with any initial function φ which is analytic and decays fast on $\mathbb{R}$. In the case where $\varphi(x) = \mathrm{e}^{-\pi x^2}$, one gets*

$$\begin{aligned}\phi(k_1,k_2) &= \sum_{n=-\infty}^{\infty} \mathrm{e}^{-\pi(k_1-n)^2}\mathrm{e}^{-2\mathrm{i}\pi k_2(k_1-n)} = \mathrm{e}^{-\pi k_1^2}\mathrm{e}^{-2\mathrm{i}\pi k_1 k_2}\sum_{n=-\infty}^{\infty} \mathrm{e}^{-\pi n^2+2\pi n(k_1+\mathrm{i}k_2)} \\ &= \mathrm{e}^{-\pi k_1^2}\mathrm{e}^{-2\mathrm{i}\pi k_1 k_2}\vartheta\left(-\mathrm{i}k_1+k_2,\mathrm{i}\right),\end{aligned}$$

where ϑ is the usual Jacobi ϑ–function, defined by

$$\vartheta(z,\tau) = \sum_{n\in\mathbb{Z}} \exp\left(\mathrm{i}\pi n^2\tau + \mathrm{i}2\pi n z\right).$$

The map $z\mapsto\vartheta(z,\mathrm{i})$ vanishes only for $z = \frac{1}{2}(1+\mathrm{i})$ in the first square $[0,1]+\mathrm{i}[0,1]$. So we can take any vector $\mathbf{K}$ of the form $\mathbf{K} = (K_1,0)$ with $K_1\notin\mathbb{Z}$.

With this at hand, one can augment the Hilbert space $\mathcal{H} = \mathbb{C}^M$ and add two dimensions in order to lift topological obstructions. We set $\widetilde{\mathcal{H}} = \mathcal{H}\oplus\mathbb{C}^2 \sim \mathbb{C}^{M+2}$. If $P\in C^\infty(\mathbb{T}^2,\mathcal{G}_N^M)$ has a non vanishing Chern $m\in\mathbb{Z}$, one considers the projector

$$\widetilde{P} := P\oplus P_{-m} \quad \in C^\infty(\mathbb{T}^2,\mathcal{G}_{N+1}^{M+2}).$$

Using Lemma 1.13, $\widetilde{P}$ has Chern 0 (hence admits frames). In matrix form, this is also

$$\widetilde{P} = \begin{pmatrix} P & 0 \\ 0 & P_{-m}\end{pmatrix} \quad \in \mathcal{M}_{M+2,M+2}.$$

Let $\widetilde{\Phi} \in \mathcal{F}_{N+1}^{M+2}$ be a frame for $\widetilde{P}$, and write it of the form

$$\widetilde{\Phi} = \begin{pmatrix} \Phi & v \end{pmatrix}, \quad \text{with} \quad \Phi \in \mathcal{M}_{M,N+2}, \quad v \in \mathcal{M}_{1,N+2}.$$

Then Φ is not necessarily a frame (it usually does not satisfy $\Phi^*\Phi = \mathbb{I}_N$), but it satisfies $\Phi\Phi^* = P$. For numerical purpose, it is therefore enough to store Φ (or $\widetilde{\Phi}$) to encode the family of projectors P (see Section 1.2.1).

In the case where P represents the low-energy spectrum of a Hamiltonian, we have $P(\mathbf{k}) = \mathbb{1}(H(\mathbf{k}) < \varepsilon(\mathbf{k}))$. Then, for Σ negative enough (lower than the $\inf_{\mathbf{k}} \inf \sigma(H(\mathbf{k}))$), we have $\widetilde{P}(\mathbf{k}) = \mathbb{1}(\widetilde{H}(\mathbf{k}) < \varepsilon(\mathbf{k}))$, where we set

$$\widetilde{H}(\mathbf{k}) = H(\mathbf{k}) \oplus (\Sigma P_{-m}(\mathbf{k})) = \begin{pmatrix} H(\mathbf{k}) & 0 \\ 0 & \Sigma P_{-m}(\mathbf{k}) \end{pmatrix}.$$

We can introduce the reduced model $\widetilde{M}(\mathbf{k}) \in \mathcal{M}_{N+1,N+1}$ with coefficients

$$\widetilde{M}_{ij} := \langle \widetilde{\phi}_i, \widetilde{H}\widetilde{\phi}_j \rangle = \langle \phi_i, H\phi_j \rangle + \Sigma \langle v_i, v_j \rangle.$$

By the min-max principle, the spectrum of $\widetilde{M}(\mathbf{k})$ is exactly the spectrum of $H(\mathbf{k})$ below $\varepsilon(\mathbf{k})$, plus the extra eigenvalue Σ. This method therefore gives an efficient way to encode the low-energy spectrum of topologically obstructed Hamiltonians.

1.5 Brillouin zone integration

In this section, we describe the results in [G3]. This is joint work with Éric CANCÈS, Antoine LEVITT, Damiano LOMBARDI and Virginie EHRLACHER.

We now explain how to numerical compute some physical properties from the family $H(\mathbf{k})$ (which can represent either a general Hamiltonian or a reduced model $M(\mathbf{k})$ obtained from Wannier interpolation). For simplicity, we assume again that $H(\mathbf{k})$ acts on a finite dimensional space $\mathbb{C}^M$ (the parameter M can be small in the case of reduced models).

Our main goal is to compute the *total energy per unit cell*. To define this quantity, we introduce the *integrated density of states* $\mathcal{N} : \mathbb{R} \to \mathbb{R}$ and the *integrated density of energy* $\mathcal{E} : \mathbb{R} \to \mathbb{R}$ defined respectively by

$$\mathcal{N}(\varepsilon) := \int_{\mathbb{T}^d} \mathrm{Tr}\Big(\mathbb{1}(H(\mathbf{k}) < \varepsilon)\Big) \mathrm{d}\mathbf{k} = \sum_{n=1}^{M} \int_{\mathbb{T}^d} \mathbb{1}(\varepsilon_{n,\mathbf{k}} \leq \varepsilon) \mathrm{d}\mathbf{k},$$

$$\mathcal{E}(\varepsilon) := \int_{\mathbb{T}^d} \mathrm{Tr}\Big(H(\mathbf{k})\mathbb{1}(H(\mathbf{k}) < \varepsilon)\Big) \mathrm{d}\mathbf{k} = \sum_{n=1}^{M} \int_{\mathbb{T}^d} \varepsilon_{n,\mathbf{k}} \mathbb{1}(\varepsilon_{n,\mathbf{k}} \leq \varepsilon) \mathrm{d}\mathbf{k}.$$

The map $\mathcal{N}$ is continuous non–decreasing, with $\mathcal{N}(-\infty) = 0$, and $\mathcal{N}(\infty) = M$. For any $0 < N < M$ a fixed chosen number (usually an integer), representing the *number of particles per unit cell*, there is $\varepsilon_F \in \mathbb{R}$ so that $\mathcal{N}(\varepsilon_F) = N$. This number may not be unique, and any solution $\varepsilon \in \mathbb{R}$ to $\mathcal{N}(\varepsilon) = N$ is called a *Fermi level*. If $\mathcal{N}^{-1}(N) = \{\varepsilon_F\}$ is unique, the system is said to be *metallic* while if $\mathcal{N}^{-1}(N) = [a, b]$ with $a < b$, the system is *insulating*. In the latter case, N is always an integer. It is the index for which we have the inequalities

$$\forall \mathbf{k} \in \mathbb{T}^d, \quad \varepsilon_{N,\mathbf{k}} \leq a, \quad \text{while} \quad b \leq \varepsilon_{N+1,\mathbf{k}}.$$

In particular, $\mathcal{E}(\varepsilon)$ is independent of $\varepsilon \in \mathcal{N}^{-1}(N)$. This quantity $E = \mathcal{E}(\varepsilon_F)$ is called the *total energy per unit cell*.

Our goal is to numerically estimate this energy. More specifically, we ask how to discretize the integrals appearing in the definition of $\mathcal{N}$ and $\mathcal{E}$. Note that to compute E, one should first estimate $\mathcal{N}(\cdot)$, in order to approximate the Fermi level ε_F, to finally compute $\mathcal{E}(\varepsilon_F)$.

1.5.1 The insulating case

Let us first recall what is known in the insulating case. We already mentioned that in this case, N is an integer, and $\varepsilon_{N,\mathbf{k}} < \varepsilon_F < \varepsilon_{N+1,\mathbf{k}}$. In particular, the projector $P_N(\mathbf{k})$ on the N first bands is well-defined and smooth, and equals $P_N(\mathbf{k}) = \mathbb{1}(H(\mathbf{k}) < \varepsilon_F)$. The total energy simplifies into

$$E = \mathcal{E}(\varepsilon_F) = \int_{\mathbb{T}^d} \mathrm{Tr}\Big(H(\mathbf{k})P_N(\mathbf{k})\Big)\mathrm{d}\mathbf{k}.$$

The key remark is that the integrand $f(\mathbf{k}) := \mathrm{Tr}\Big(H(\mathbf{k})P_N(\mathbf{k})\Big)$ is analytic and periodic. In particular, its integral can be well approximated by a Riemann sum on a uniform grid. For $L \in \mathbb{N}$ a discretization parameter, we define the uniform grid with L points per direction as

$$\mathbb{T}^d_L := \mathbb{T}^d \cap L^{-1}\mathbb{Z}^d. \tag{1.7}$$

We recall the following well-known result (see for instance [GL16, Lemma 5.1])

Lemma 1.16. *There is $C \geq 0$ and $\alpha > 0$ so that, for all $Y > 0$, all $L \in \mathbb{N}$ and all functions $f : \mathbb{T}^d \to \mathbb{R}$ periodic and analytic on $S_Y := \mathbb{T}^d + \mathrm{i}[-Y, Y]$, we have*

$$\left| \int_{\mathbb{T}^d} f(\mathbf{k})\mathrm{d}\mathbf{k} - \frac{1}{L^d} \sum_{\mathbf{k}\in\mathbb{T}^d_L} f(\mathbf{k}) \right| \leq \frac{C}{Y^d} \left(\sup_{\mathbf{z}\in S_Y} |f(\mathbf{z})| \right) \mathrm{e}^{-\alpha Y L}.$$

We deduce that there is a constant $C' \geq 0$ and $\alpha' > 0$ so that, for all $L \in \mathbb{N}$, we have

$$\Big|E - E^L\Big| \leq C'\mathrm{e}^{-\alpha' L}, \qquad \text{with} \quad E^L := \frac{1}{L^d} \sum_{\mathbf{k}\in\mathbb{T}^d_L} \mathrm{Tr}\Big(H(\mathbf{k})P_N(\mathbf{k})\Big).$$

In other words, we have exponential speed of convergence in L when we approximate the Brillouin zone integration by the corresponding Riemann sum on a regular mesh. In practice, this means that we only need to compute the spectrum of $H(\mathbf{k})$ for $\mathbf{k}$ on a coarse grid $\mathbb{T}^d_L$. This grid contains L^d points. Using the symmetries of the system, one can restrict the computations for $\mathbf{k}$ in the so-called reduced Brillouin zone.

The idea to take uniform grids originates from the work of Monkhorst and Pack [MP76], and the fact that it leads to exponential speed of convergence is folklore in the community. This speed is also attained in some non–linear Kohn–Sham models, such as the reduced Hartree–Fock model [GL16].

1.5.2 The metallic case

We now turn to the metallic case. In our work [G3], we presented two different methods to compute the Fermi level ε_F and the total energy energy per unit cell E. The first one follows an idea of Blöchl, Jepsen and Andersen [BJA94] (see also [Zah05]), and consists into splitting the Brillouin zone $\mathbb{T}^d$ into a union of simplices (tetrahedra if $d = 3$) $\mathbb{T}^d = \bigcup_j T_j$, interpolating the maps $\mathbf{k} \mapsto \varepsilon_{n,\mathbf{k}}$ by piece–wise polynomial functions on these simplices, and evaluating quantities of the form

$$\int_T p(\mathbf{k})\mathbb{1}\left(q(\mathbf{k}) \leq \varepsilon\right)\mathrm{d}\mathbf{k},$$

where p and q are (low–order) polynomials approximating the maps $\mathbf{k} \mapsto \varepsilon_{n,\mathbf{k}}$ on T. We will not present the results of this first method here.

The idea of the second one is to smear the function $\mathbb{1}(x < 0)$ appearing in the definition of $\mathcal{N}$ and $\mathcal{E}$. Let $p \in \mathbb{N}$ be some chosen order, and let δ^1 be a function which is analytic on some complex

strip of the form $\mathbb{R}+\mathrm{i}[-YY]$, and such that, when restricted to the real line, $\delta^1 : \mathbb{R} \mapsto \mathbb{R}$ is a Schwartz function satisfying

$$\int_{\mathbb{R}} \delta^1 = 1, \quad \text{and} \quad \forall 1 \leq n \leq p, \quad \int_{\mathbb{R}} \delta^1(x) x^n = 0.$$

We say that δ^1 is a *mollifier* of order p. For $p = 1$, one can take $\delta^1 \geq 0$ a positive function, but for $p \geq 2$, δ^1 must alternate sign, since we must have $\int \delta^1(x)x^2 = 0$. Typical choices for the function δ^1 are

- the Fermi–Dirac smearing $\delta^1(x) = \frac{1}{1+\mathrm{e}^x+\mathrm{e}^{-x}}$, which is positive, of order $p = 1$;
- the Gaussian smearing $\delta^1(x) = \pi^{-1/2}\mathrm{e}^{-x^2}$, which is positive, of order $p = 1$;
- smearings of the form $\delta^1(x) = P(x)\mathrm{e}^{-x^2}$ with P a polynomial chosen so that δ^1 is of order p. This includes the Marzari–Vanderbild cold smearings [Mar+99], and the Methfessel–Paxton smearings [MP89].

For $\sigma > 0$ a smearing parameter, we set $\delta^\sigma(x) := \sigma^{-1}\delta^1(\sigma^{-1}x)$, and we approximate the function $f(x) := \mathbb{1}(x < 0)$ by the smooth function

$$f^\sigma(x) := \Big(f * \delta^\sigma\Big)(x) = \int_x^\infty \delta^\sigma(s)\mathrm{d}s = f^1(\sigma^{-1}x).$$

Let us introduce the smoothed quantities

$$\mathcal{N}^\sigma(\varepsilon) := \int_{\mathbb{T}^d} \mathrm{Tr}\Big(f^\sigma(H(\mathbf{k}) - \varepsilon)\Big)\mathrm{d}\mathbf{k} \quad \text{and} \quad \mathcal{E}^\sigma(\varepsilon) := \int_{\mathbb{T}^d} \mathrm{Tr}\Big(H(\mathbf{k})f^\sigma(H(\mathbf{k}) - \varepsilon)\Big)\mathrm{d}\mathbf{k}.$$

A computation shows that $\mathcal{N}^\sigma = \mathcal{N} * \delta^\sigma$ and $\mathcal{E}^\sigma = \mathcal{E} * \delta^\sigma$. This time, the map $\varepsilon \mapsto \mathcal{N}^\sigma(\varepsilon)$ is strictly increasing (also in the insulating case). The approximate Fermi level is ε_F^σ solution to $\mathcal{N}^\sigma(\varepsilon_F^\sigma) = N$ and the approximate total energy is $E^\sigma := \mathcal{E}^\sigma(\varepsilon_F^\sigma)$.

The function $\mathcal{N}^\sigma$ is expected to be a good approximation of $\mathcal{N}$ in the regions where $\mathcal{N}$ is smooth. Unfortunately, the map $\varepsilon \mapsto \mathcal{N}(\varepsilon)$ is not smooth in general, due to the possible band crossings. For our purpose, we will assume smoothness. More specifically, we define the *Fermi surfaces* $\mathcal{S}(\varepsilon) \subset \mathbb{T}^d$ and $\mathcal{S}_n(\varepsilon) \subset \mathbb{T}^d$ as the sets

$$\mathcal{S}(\varepsilon) := \bigcup_{1 \leq n \leq M} \mathcal{S}_n(\varepsilon), \quad \text{where} \quad \mathcal{S}_n(\varepsilon) := \Big\{\mathbf{k} \in \mathbb{T}^d, \quad \varepsilon_{n,\mathbf{k}} = \varepsilon\Big\}.$$

Lemma 1.17 ([G3] Lemmas 5.5-5.6)**.** *Under the following two assumptions:*

- **Assumption 1 (no band crossings at ε_F)**: $\forall n \neq m$, $\mathcal{S}_n(\varepsilon_F) \cap \mathcal{S}_m(\varepsilon_F) = \emptyset$;
- **Assumption 2 (no flat bands at ε_F)**: $\forall 1 \leq n \leq M$, $\forall \mathbf{k} \in \mathcal{S}_n(\varepsilon_F)$, $\nabla_{\mathbf{k}}\varepsilon_{n,\mathbf{k}} \neq 0$,

the maps $\mathcal{N}$ and $\mathcal{E}$ are smooth on a neighbourhood $\mathcal{U}$ of ε_F. There is $C \geq 0$ so that, for all $\sigma > 0$,

$$\begin{cases} \forall \varepsilon \in \mathcal{U}, \qquad |\mathcal{N}(\varepsilon) - \mathcal{N}^\sigma(\varepsilon)| \leq C\sigma^{p+1}, \qquad |\mathcal{E}(\varepsilon) - \mathcal{E}^\sigma(\varepsilon)| \leq C\sigma^{p+1}, \\ |\varepsilon_F - \varepsilon_F^\sigma| \leq C\sigma^{p+1}, \quad \textit{and} \quad |E - E^\sigma| \leq C\sigma^{p+1}. \end{cases}$$

We skip the proof for brevity. Assumption 2 implies in particular that $\mathcal{S}_n(\varepsilon_F)$ is a well-defined smooth manifold of dimension $d-1$. Together with Assumption 1, we get that the Fermi surface $\mathcal{S}(\varepsilon_F)$ is also a smooth manifold of dimension $d-1$. The proof of smoothness uses the co-area formula, and the estimates comes from a Taylor expansion of $\mathcal{N}$ and $\mathcal{E}$, and the fact that δ^1 decays fast, and is of order p (hence cancels polynomials of degree p).

Unfortunately, the quantities $\mathcal{N}^\sigma$ and $\mathcal{E}^\sigma$ can not be computed numerically, as they still involve an integration over the whole Brillouin zone $\mathbb{T}^d$. However, the integrands are now smooth and periodic,

hence the integrals can be well approximated by the corresponding Riemann sums. Recall that the uniform $\mathbb{T}^d_L$ was defined in (1.7). We introduce the approximate quantities $\mathcal{N}^{\sigma,L}$ and $\mathcal{E}^{\sigma,L}$ by

$$\mathcal{N}^{\sigma,L}(\varepsilon) := \frac{1}{L^d}\sum_{\mathbf{k}\in\mathbb{T}^d_L} \mathrm{Tr}\Big(f^\sigma(H(\mathbf{k})-\varepsilon)\Big), \quad \mathcal{E}^{\sigma,L}(\varepsilon) := \frac{1}{L^d}\sum_{\mathbf{k}\in\mathbb{T}^d_L} \mathrm{Tr}\Big(H(\mathbf{k})f^\sigma(H(\mathbf{k})-\varepsilon)\Big).$$

The approximate Fermi level is $\varepsilon_F^{\sigma,L}$, unique solution to $\mathcal{N}^{\sigma,L}(\varepsilon_F^{\sigma,L}) = N$, and the approximate total energy per unit cell is $\mathcal{E}^{\sigma,L} := \mathcal{E}^{\sigma,L}(\varepsilon_F^{\sigma,L})$. All these quantities can be computed numerically. Our main result is the following. We skip its proof, and only emphasize that it uses Lemma 1.16 on Riemann sums, and the fact that δ^1 is analytic on $\mathbb{R}+\mathrm{i}[-Y,Y]$.

Theorem 1.18 ([G3] Lemmas 5.10, Theorem 5.11)**.** *There is a neighbourhood $\mathcal{U}$ of ε_F, and constants $C>0$ and $\eta>0$ so that, for all $\sigma>0$ and all $L\in\mathbb{N}^*$,*

$$\begin{cases} \forall \varepsilon\in\mathcal{U}, \quad \left|\mathcal{N}^\sigma(\varepsilon)-\mathcal{N}^{\sigma,L}(\varepsilon)\right| \le C\sigma^{-(d+1)}\mathrm{e}^{-\eta\sigma L}, \quad \left|\mathcal{E}^\sigma(\varepsilon)-\mathcal{E}^{\sigma,L}(\varepsilon)\right| \le C\sigma^{-(d+1)}\mathrm{e}^{-\eta\sigma L}, \\ \left|\varepsilon_F^\sigma-\varepsilon_F^{\sigma,L}\right| \le C\sigma^{-(d+1)}\mathrm{e}^{-\eta\sigma L}, \quad \textit{and} \quad \left|E^\sigma-E^{\sigma,L}\right| \le C\sigma^{-(d+1)}\mathrm{e}^{-\eta\sigma L}. \end{cases}$$

In particular,

$$\left|\varepsilon_F-\varepsilon_F^{\sigma,L}\right| \le C\left(\sigma^{p+1}+\sigma^{-(d+1)}\mathrm{e}^{-\eta\sigma L}\right), \quad \textit{and} \quad \left|E^\sigma-E^{\sigma,L}\right| \le C\left(\sigma^{p+1}+\sigma^{-(d+1)}\mathrm{e}^{-\eta\sigma L}\right).$$

We can now optimize the parameters σ and L to obtain some desired accuracy. Since the number of grid points grows as L^d, it affects strongly the computational time. On the contrary, the parameter δ is irrelevant for numerical time, and can be tuned as wanted. The choice $\sigma \approx \log(L)L^{-1}$ leads to an error of order $O(L^{-(p+1)})$ up to log factors.

The speed of convergence in the metallic case is therefore worse that in the insulating case, at least for the smearing method: we have polynomial speed of convergence, instead of exponential speed.

1.6 Perspectives

In [G2], we designed an algorithm to explicitly contract a loop $\mathbb{T}^1 \to \mathrm{U}(N)$ when this is possible. A natural next step would be to design an algorithm to contract maps from $\mathbb{T}^d \to \mathrm{U}(N)$. Our proof can handle the case $d=1$ and $d=2$, but when $d=3$, some other obstructions appear. Using our column interpolation method, we can reduce the problem to the following question (see also the recent work [MR23]):

Given a map $U:\mathbb{T}^3\to\mathrm{SU}(2)$ which is contractible, how to numerically construct a contraction?

Using the identification $\mathrm{SU}(2)\sim\mathbb{S}^3$, this is related to the following question: given a contractible map $F:\mathbb{T}^d\to\mathbb{S}^d$ with null degree, how to numerically contract it?

CHAPTER 2

SEMI-PERIODIC SYSTEMS

2.1 Introduction

In this chapter, we study the spectrum of periodic operators, when they are cut. We start from a periodic Schrödinger operator of the form

$$H := -\Delta + V \quad \text{acting on} \quad L^2(\mathbb{R}^d),$$

where V is a bounded $\mathbb{Z}^d$–periodic potential. It is a standard result that the spectrum of H is purely essential, composed of bands and gaps. We then cut this operator, and study the edge Hamiltonian

$$H^\sharp := -\Delta + V \quad \text{acting on} \quad L^2(\mathbb{R}^d_+), \qquad \mathbb{R}^d_+ := \mathbb{R}_+ \times \mathbb{R}^{d-1},$$

with some boundary conditions at the cut $\mathbb{R}^d_0 := \{0\} \times \mathbb{R}^{d-1}$ (typically Dirichlet boundary conditions).

Using Bloch–Floquet theory together with the Weyl's characterization of the essential spectrum, it is not difficult to see that $\sigma(H) = \sigma_{\text{ess}}(H) \subset \sigma(H^\sharp)$ (the bulk spectrum is purely essential, and is contained in the cut spectrum). However, *edge modes* can appear at the cut, so that the *edge spectrum*

$$\sigma_{\text{edge}} := \sigma(H) \setminus \sigma(H^\sharp)$$

is not empty in general. The main goal of this chapter is to describe σ_{edge} in various situations.

Notation, Hamiltonians on a channel

By periodicity of V, both operators H and $H^\sharp$ are periodic in the last $(d-1)$–variables. After a Bloch transform in these directions, we obtain smooth families of Hamiltonians $H_{\mathbf{k}}$ and $H^\sharp_{\mathbf{k}}$ with $\mathbf{k} \in \mathbb{R}^{d-1}$, of the form

$$H_{\mathbf{k}} := -\Delta + V \quad \text{acting on} \quad L^2(\Omega), \quad \text{and} \quad H^\sharp_{\mathbf{k}} := -\Delta + V \quad \text{acting on} \quad L^2(\Omega^+),$$

where $\Omega := \mathbb{R} \times (0,1)^{d-1}$ is a channel in the first direction, and $\Omega^\pm := \mathbb{R}_\pm \times (0,1)^{d-1}$. The operator $H_{\mathbf{k}}$ and $H^\sharp_{\mathbf{k}}$ have domains $\mathcal{D}_{\mathbf{k}}$ and $\mathcal{D}^\sharp_{\mathbf{k}}$ reflecting the $\mathbf{k}$-quasi periodic boundary conditions along the $(d-1)$ last variables. We assume Dirichlet boundary conditions at the cut $\{0\} \times (0,1)^{d-1}$, although our results can be extended to other boundary conditions. The maps $\mathbf{k} \mapsto H_{\mathbf{k}}$ and $\mathbf{k} \mapsto H^\sharp_{\mathbf{k}}$ are $\mathbb{Z}^{d-1}$–periodic, so we write $\mathbf{k} \in \mathbb{T}^{d-1}$.

In what follows, we consider a slightly different setting, and study a family of Hamiltonians $\mathbb{T}^1 \ni t \mapsto H_t$ acting on $L^2(\Omega)$, and with corresponding edge Hamiltonians $H^\sharp_t$ acting on $L^2(\Omega^+)$. The main difference is that t is now a real (one–dimensional) parameter. The previous case falls in this setting

when $d = 2$ and $t = k$. Our results will therefore be appropriate to study edge states in 2-dimensional materials. This setting also allows to study family of Hamiltonians of the form

$$H_t := -\Delta + V_t, \qquad \text{acting on} \quad L^2(\Omega),$$

where $t \mapsto V_t$ is a smooth periodic map of potentials, and where $\Omega := \mathbb{R} \times \mathbb{T}^{d-1}$, that is we assume periodic boundary conditions for the last $(d-1)$ variables. This is not mandatory, but avoids the discussion of boundary conditions at $\partial\Omega$ for the definition of self–adjoint extensions. In particular, the domain of H_t is $H^2(\Omega)$, independent of $t \in \mathbb{T}^1$.

When t is seen as a time variable, H_t models a Thouless pump [Tho83; BGO10]. Note we need not assume that $V_t(\cdot)$ is periodic in this general case. We will extensively study the *dislocated* case where $V_t(x) = V(x - t)$. In this case, the periodicity in t implies that $V(\cdot)$ is 1-periodic. The operator $H_t^\sharp$ can be seen to describe a fixed periodic system $-\Delta + V$ which is cut at $\{t\} \times \mathbb{T}^{d-1}$, for some $t \in \mathbb{R}$.

2.2 From the cut Hamiltonian to the junction Hamiltonian

In this section, we describe the results in [G6].

We first consider the case where $H_t = -\Delta + V_t$. We define the bulk/edge index for such a family as a spectral flow of $t \mapsto H_t^\sharp$, and show that this definition allows to predict the spectrum of general junctions.

2.2.1 Spectral flows

First, we recall some basic facts on the *Spectral flow* [APS76; Phi96].

Let $\mathbb{T}^1 \ni t \mapsto A_t$ be a periodic map of self-adjoint operators acting on the same Hilbert space $\mathcal{H}$. We assume that this map is norm-resolvent continuous, which implies in particular that the spectrum $t \mapsto \sigma(A_t)$ is continuous. We define the spectrum and essential spectrum of the family A_t respectively as the closure of the unions

$$\sigma(\{A_t\}) := \overline{\bigcup_{t\in\mathbb{T}^1} \sigma(A_t)}, \quad \text{and} \quad \sigma_{\text{ess}}(\{A_t\}) := \overline{\bigcup_{t\in\mathbb{T}^1} \sigma_{\text{ess}}(A_t)}.$$

An interval (a, b) in $\mathbb{R} \setminus \sigma_{\text{ess}}(\{A_t\})$ is an *essential gap of the family* A_t. If E is in an essential gap (a, b) of the family A_t, we define the spectral flow of A_t at energy E, noted

$$\mathrm{Sf}(A_t, E, \mathbb{T}^1),$$

as the net number of branches of eigenvalues of $\mathbb{T}^1 \ni t \mapsto A_t$ going downwards in the essential gap where E lies. In particular, the spectral flow is an integer, and is independent of E in the essential gap (a, b) (see Figure 2.1).

The following results are standard (see for instance [Phi96]).

Lemma 2.1.

- *If* $E \notin \sigma(\{A_t\})$, *then* $\mathrm{Sf}\Big(A_t, E, \mathbb{T}^1\Big) = 0$.
- *If* $f : \mathbb{R} \to \mathbb{R}$ *is a strictly increasing function, then* $\mathrm{Sf}\Big(f(A_t), f(E), \mathbb{T}^1\Big) = \mathrm{Sf}\Big(A_t, E, \mathbb{T}^1\Big)$.
- *If* $\mathbb{T}^1 \ni t \mapsto K_t$ *is a norm continuous periodic map of compact operators, then*

$$\mathrm{Sf}\Big(A_t + K_t, E, \mathbb{T}^1\Big) = \mathrm{Sf}\Big(A_t, E, \mathbb{T}^1\Big).$$

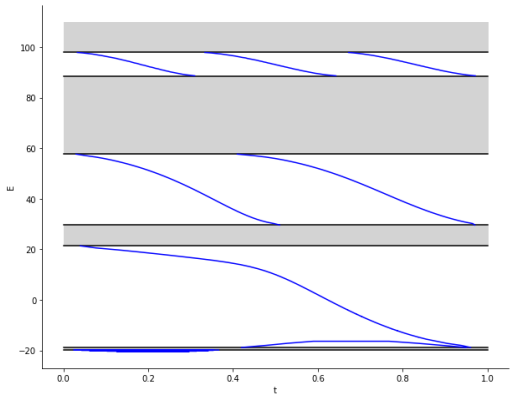


Figure 2.1: The spectral flow is n in the n-th lowest essential gap.

For the first point, we note that no eigenvalue touches E, hence there is no flow. The second point comes from the fact that if $t \mapsto \lambda_t$ is a branch of eigenvalue of A_t crossing E downwards (resp. upwards), then $t \mapsto f(\lambda_t)$ is a branch of eigenvalue of $f(A_t)$ crossing E downwards (resp. upwards). For the last point, we first note that the addition of a compact operator does not change the essential spectrum, so if $E \notin \sigma_{\mathrm{ess}}(\{A_t\})$, then $E \notin \sigma_{\mathrm{ess}}(\{A_t+K_t\})$ as well. By continuity of the spectral flow with respect to the operator norm topology (see [Phi96; G7] for details), the map $s \mapsto \mathrm{Sf}\,(A_t + sK_t, E, \mathbb{T}^1)$ is continuous and integer valued, hence constant.

2.2.2 The bulk/edge index

Let $\mathbb{T}^1 \ni t \mapsto H_t$ be a continuous family of Schrödinger operators acting on the full channel $L^2(\Omega)$. We focus on the case where H_t is of the form $H_t := -\Delta + V_t$, but our results can be extended to other cases. We define the bulk/edge index of H_t as the spectral flow of its cut counterpart $t \mapsto H_t^\sharp$, with Dirichlet boundary conditions. In order to do so, we need a Lemma (see below for the proof).

Lemma 2.2 ([G6] Theorem 43)**.** *For all $t \in \mathbb{T}^1$, we have $\sigma_{\mathrm{ess}}(H_t) = \sigma_{\mathrm{ess}}(H_t^{\sharp,+}) \cup \sigma_{\mathrm{ess}}(H_t^{\sharp,-})$. In particular, if $E \notin \sigma_{\mathrm{ess}}(\{H_t\})$, then $E \notin \sigma_{\mathrm{ess}}(\{H_t^{\sharp,\pm}\})$, and the spectral flows $\mathrm{Sf}(H_t^{\sharp,\pm}, E, \mathbb{T}^1)$ are well-defined. If in addition $E \notin \sigma(\{H_t\})$, then*

$$\mathrm{Sf}(H_t^{\sharp,+}, E, \mathbb{T}^1) + \mathrm{Sf}(H_t^{\sharp,-}, E, \mathbb{T}^1) = 0.$$

Definition 2.3. *For $E \notin \sigma(\{H_t\})$, we define the bulk/edge index of H_t at energy E as the spectral flow of $H_t^{\sharp,+}$, that is*

$$\mathcal{I}\left(H_t, E, \mathbb{T}^1\right) := \mathrm{Sf}(H_t^{\sharp,+}, E, \mathbb{T}^1) \qquad = -\mathrm{Sf}(H_t^{\sharp,-}, E, \mathbb{T}^1).$$

With our definition, the bulk/edge index is an edge quantity. We note however that it only depends on bulk properties, namely on the potential V_t, hence our denomination *bulk/edge*. Here, we made a specific choice for the edge boundary conditions, namely Dirichlet, but we prove below that this index is independent of the chosen (fixed) boundary conditions. In the community of bulk-edge correspondence, the bulk index has many definitions depending on the situations: for two–dimensional periodic systems, it is usually defined as the Chern number of $P(\mathbf{k}) := \mathbb{1}(H(\mathbf{k}) < E)$. This is the case for instance in [Hat93a; Hat93b] (see also [ASV13]).

We note that our Definition 2.3 is independent of the situation, and turns out to be quite flexible. For instance, we do not assume the potentials $V_t(\cdot)$ to be periodic (but only that $t \mapsto V_t$ is periodic). The fact that all known bulk and edge definitions coincide (*«bulk-edge correspondence»*) has been a fertile area of research, starting from the works of Hatsugai [Hat93a; Hat93b] (see [Tho+82; Hal82] for earlier works). One should mention the numerous works by Graf and co-authors [EG02; EGS05; Gra07;

GO08; BGO10; GP13; GS18]. Some proofs of bulk-edge correspondence involve K-theory [SKR00; KRS02; PS16], micro-local analysis [Dro21b; Dro21a], and so on.

Proof of Lemma 2.2. We consider the set $\Omega^{\mathrm{cut}} = \Omega^- \cup \Omega^+ = \Omega \setminus \Gamma$ where $\Gamma := \{0\} \times \mathbb{T}^{d-1}$ is the cut. Note that $L^2(\Omega) = L^2(\Omega^{\mathrm{cut}})$. We introduce the cut operator

$$H_t^{\mathrm{cut}} := -\Delta + V_t \quad \text{acting on } L^2(\Omega), \text{ with domain } H^2(\Omega^{\mathrm{cut}}) \cap H_0^1(\Omega^{\mathrm{cut}}). \tag{2.1}$$

It has an action similar to H_t, but with a domain which reflects Dirichlet boundary conditions at the cut Γ.

Let $\Sigma \in \mathbb{R}$ be a negative enough number so that

$$\forall t \in \mathbb{T}^1, \qquad \Sigma < \sigma(H_t) \quad \text{and} \quad \Sigma < \sigma(H_t^{\mathrm{cut}}).$$

It is a classical result that there is $m \in \mathbb{N}$ large enough, depending on the dimension d, so that, for all $t \in \mathbb{T}^1$, the operator

$$K_t := (\Sigma - H_t)^{-m} - (\Sigma - H_t^{\mathrm{cut}})^{-m} \tag{2.2}$$

is compact, see for instance [RS79, Theorem XI.79]. This already proves that

$$\sigma_{\mathrm{ess}}\Big((\Sigma - H_t)^{-m}\Big) = \sigma_{\mathrm{ess}}\Big((\Sigma - H_t^{\mathrm{cut}})^{-m}\Big), \qquad \text{hence} \qquad \sigma_{\mathrm{ess}}\Big(H_t\Big) = \sigma_{\mathrm{ess}}\Big(H_t^{\mathrm{cut}}\Big).$$

In addition, since the left and right channels are decoupled in H_t^{cut}, we have $\sigma(H_t^{\mathrm{cut}}) = \sigma(H_t^{\sharp,-}) \cup \sigma(H_t^{\sharp,+})$ and $\sigma_{\mathrm{ess}}(H_t^{\mathrm{cut}}) = \sigma_{\mathrm{ess}}(H_t^{\sharp,-}) \cup \sigma_{\mathrm{ess}}(H_t^{\sharp,+})$. This proves the first point.

In addition, the map $t \mapsto K_t$ is continuous and periodic. Using Lemma 2.1, we deduce that, for any $E \notin \sigma_{\mathrm{ess}}(\{H_t\})$, we have

$$\mathrm{Sf}\Big((\Sigma - H_t)^{-1}, (\Sigma - E)^{-1}, \mathbb{T}^1\Big) = \mathrm{Sf}\Big((\Sigma - H_t^{\mathrm{cut}})^{-1}, (\Sigma - E)^{-1}, \mathbb{T}^1\Big).$$

The map $f : x \mapsto (\Sigma - x)^{-1}$ is strictly increasing. By Lemma 2.1, we obtain

$$\mathrm{Sf}(H_t, E, \mathbb{T}^1) = \mathrm{Sf}(H_t^{\mathrm{cut}}, E, \mathbb{T}^1).$$

Again, the Dirichlet boundary conditions at the cut disentangles the left and right sides of H_t^{cut}, so

$$\mathrm{Sf}(H_t^{\mathrm{cut}}, E, \mathbb{T}^1) = \mathrm{Sf}(H_t^{\sharp,+}, E, \mathbb{T}^1) + \mathrm{Sf}(H_t^{\sharp,-}, E, \mathbb{T}^1).$$

If $E \notin \sigma(\{H_t\})$, then $\mathrm{Sf}(H_t, E, \mathbb{T}^1) = 0$, which implies $\mathrm{Sf}(H_t^{\sharp,+}, E, \mathbb{T}^1) + \mathrm{Sf}(H_t^{\sharp,-}, E, \mathbb{T}^1) = 0$. □

2.2.3 Junctions between two Hamiltonians

As we see from the proof of Lemma 2.2, cutting the channel induces a compact perturbation of the resolvent, and does not affect the spectral flow. This allows to compute the spectral flows of general junctions.

Let $t \mapsto V_t^L$ and $t \mapsto V_t^R$ be two periodic continuous families of potentials (L/R stands for Left/Right), and let $t \mapsto H_t^{L/R} := -\Delta + V_t^{L/R}$ be the corresponding bulk operators. Let $\chi : \mathbb{R} \to \mathbb{R}$ be any switch function, so that $\chi(x) = 1$ for $x < -X$ while $\chi(x) = 0$ for $x > X$, for some X large. We consider the junction operator

$$H_t^{\mathrm{junction}} := -\Delta + V_t^L(x)\chi(x) + V_t^R(x)(1 - \chi(x)) \quad \text{acting on } L^2(\Omega), \text{ with domain } H^2(\Omega). \tag{2.3}$$

This operator models a transition between the potential V_t^L on the left, and the potential V_t^R on the right, hence the denomination «junction».

Theorem 2.4 ([G6] Theorem 44)**.** *Let $E \in \mathbb{R}$ be such that $E \notin \sigma(\{H_t^L\})$ and $E \notin \sigma(\{H_t^R\})$. Then $E \notin \sigma_{\rm ess}(\{H_t^{\rm junction}\})$, and*

$$\mathrm{Sf}\left(H_t^{\rm junction}, E, \mathbb{T}^1\right) := \mathcal{I}\left(H_t^R, E, \mathbb{T}^1\right) - \mathcal{I}\left(H_t^L, E, \mathbb{T}^1\right).$$

Proof. The proof is similar to the one of Lemma 2.2. This time, we cut at $x = -X$ and at $x = X$. On the middle part, we have a Schrödinger operator on the relatively compact space $(-X, X) \times \mathbb{T}^{d-1}$. The middle operator is therefore compact resolvent and does not contribute to the spectral flow. The result follows. □

If $E \notin \sigma_{\rm ess}(\{H_t^{\rm junction}\}$ is an eigenvalue of the junction operator $H_t^{\rm junction}$ for some $t \in \mathbb{T}^1$, then the corresponding eigenfunction $\psi_E \in L^2(\Omega)$, called the *edge mode*, must be localised at the junction. In most situations (for instance if V_t^R and V_t^L are bounded functions), the Combes-Thomas estimates [CT73] imply that ψ_E must be exponentially decaying away from the junction.

A corollary of Theorem 2.4 is that the bulk/edge index is independent of the choice of boundary conditions at the cut. We only require that the corresponding operator $\widetilde{K_t}$ in (2.2) is compact for this new boundary conditions. This happens in particular for Neumann boundary conditions, see [RS79, Theorem XI.79]. In this case, one can repeat all the previous arguments. Let us prove that the spectral flow is the same in the Neumann and Dirichlet cases.

Lemma 2.5. *Let $H_t^{\sharp_N,\pm}$ be the edge operators with Neumann boundary conditions at the cut. Then, we have $E \notin \sigma_{\rm ess}(\{H_t^{\sharp_N,\pm}\})$ iff $E \notin \sigma_{\rm ess}(\{H_t^{\sharp,\pm}\})$, and, in this case,*

$$\mathrm{Sf}\Big(H_t^{\sharp_N,\pm}, E, \mathbb{T}^1\Big) = \mathrm{Sf}\Big(H_t^{\sharp,\pm}, E, \mathbb{T}^1\Big).$$

Proof. We prove the result for the right channel. We write $H_t := -\Delta + V_t$, and we set $V_t^R := V_t$ and $V_t^L := \Lambda$, with $\Lambda > E$. We consider the corresponding junction operator $H_t^{\rm junction}$ as in (2.3). Since $\Lambda > E$, the energy E is not in the spectrum of the left operator H_t^L with Dirichlet nor Neumann boundary conditions. So

$$\mathrm{Sf}\Big(H_t^{\rm junction}, E, \mathbb{T}^1\Big) = \mathrm{Sf}\Big(H_t^{\sharp,+}, E, \mathbb{T}^1\Big), \quad \text{and similarly,} \quad \mathrm{Sf}\Big(H_t^{\rm junction}, E, \mathbb{T}^1\Big) = \mathrm{Sf}\Big(H_t^{\sharp_N,+}, E, \mathbb{T}^1\Big).$$

□

One should think of these proofs as a LEGO® game. One has left and right pieces, representing respectively V_t^L and V_t^R, all labelled by some integers, representing their bulk/edge indices. One can assemble a left and a right piece using a switch function χ. We obtain a new integer, which is the difference between the right and the left indices. Now, making specific choices for the left piece, for the right piece and for the switch function χ, allows to identify non trivial identities between these indices.

2.3 The dislocated model

Let us give an application of the previous theory. We focus on the special case of dislocations, where one can compute explicitly the bulk/edge index. This allows to describe precisely the spectrum of a general junction between two materials.

2.3.1 Exact computation in the case of dislocations

We focus on the special case $V_t(x, \mathbf{y}) = V(x - t, \mathbf{y})$, where $V : \Omega \to \mathbb{R}$ is 1-periodic in the first direction: $V(x + 1, \mathbf{y}) = V(x, \mathbf{y})$ for all $(x, \mathbf{y}) \in \Omega = \mathbb{R} \times \mathbb{T}^{d-1}$. Explicitly,

$$H_t := -\Delta + V(x - t, \mathbf{y}), \quad \text{acting on } L^2(\Omega), \text{ with domain } H^2(\Omega). \tag{2.4}$$

We assume for simplicity that V is a continuous potential, which implies in particular that $t \mapsto H_t$ is norm-resolvent continuous. Since V is 1-periodic, the maps $t \mapsto H_t$ and $t \mapsto H_t^{\sharp,\pm}$ are indeed 1–periodic. Actually, H_t is a translated version of $H_0 := H_{t=0}$, so the spectrum of H_t is independent of $t \in \mathbb{T}^1$.

Since V is 1–periodic in the first direction, the operator H_0 can be Bloch diagonalized in this direction. For $E \notin \sigma(H_0)$, we denote by

$$\mathcal{N}(E) := \text{ number of Bloch bands of } H_0 \text{ below } E.$$

It is also the number of Bloch bands of H_t below E for any t.

In this special case, one can compute the bulk/edge index of H_t. The following result was proved in dimension $d = 1$ by Korotyaev [Kor00; Kor05], Drouot [Dro21b] (under an extra assumption), and the author [G5]. The general case $d \geq 1$ was proved by Hempel and Kohlmann in a series of papers [HK11b; HK11a; HK12; Hem+15].

Theorem 2.6. *In the case of dislocations, when H_t is of the form (2.4), we have, for all $E \notin \sigma(H_0)$,*

$$\mathcal{I}(H_t, E, \mathbb{T}^1) = \mathcal{N}(E).$$

The main idea of the proof of Hempel and Kohlmann is to consider a dislocated model, that is a junction between a left (constant) potential $V_t^L(\mathbf{x}) = V(\mathbf{x})$, and a right potential $V_t^R(\mathbf{x}) = V(x_1 - t, \mathbf{y})$, namely

$$H^{\text{disloc}} := -\Delta + V(x_1, \mathbf{y})\mathbb{1}(x_1 < 0) + V(x_1 - t, \mathbf{y})\mathbb{1}(x_1 \geq 0), \quad \text{acting on } L^2(\Omega), \text{ with domain } H^2(\Omega).$$

Consider the state $\gamma_t(E) := \mathbb{1}(H_t^{\text{disloc}} \leq E)$. It corresponds to a Fermi sea of fermions filling the material described by H_t^{disloc} at Fermi level E. At $t = 0$ or $t = 1$, we recover the bulk potential with no dislocation, and the operator $\gamma_t(E)$ describes a Fermi sea with $\mathcal{N}(E)$ particles per unit cell. As t moves from 0 to 1, a new cell is created near the cut $x_1 = 0$, and $\mathcal{N}(E)$ new particles must fill this new cell. The proof of Hempel and Kohlmann shows that these new particles are pumped from upper energies. This reflects the spectral flow of $\mathcal{N}(E)$ eigenvalues coming downwards.

We display below in Figure 2.2 a numerical computation of the spectrum of H_t^{cut}, defined in (2.1), as a function of t, in the case $d = 1$ and $V_t(x) := V(x - t)$ with $V(x) = 50 \cdot \cos(2\pi x) + 10 \cdot \cos(4\pi x)$. In this figure, one clearly sees that the spectral flow equals n in the n-th essential gap.

2.3.2 The Grand Hilbert Hotel

We find the following analogy useful. Let us first recall that the story of the Hilbert Hotel, this famous hotel with infinitely many rooms, numbered 1,2,3, and so on. At the beginning of the story, the hotel is fully occupied, with infinitely many guests. A new guest arrives. To accommodate him, the manager asks everyone to shift one room to the right. So an infinite hotel, even full, can welcome a finite number of new guests. The next day, the manager buys the land next to its hotel, and creates a new "room 0". In order to fill it, he asks everyone to shift one room to the left.

The story goes on, and the Hilbert bus arrives (a bus with infinitely many seats, numbered 1,2,3, and so on). The bus is full of travellers. The manager can still accommodate all these new people: he would ask the guest currently occupying Room n to go to Room $2n$, and the traveller seated in Seat m to go to Room $2m - 1$. So the fully occupied Hilbert hotel can also welcome infinitely (countable) many new guests.

The **Grand Hilbert Hotel** is similar, but it has an infinite number of floors, numbered 1,2,3,... and an infinity of rooms per floor (floor 1 room 1, ...). As in the original story, the hotel is fully occupied at the beginning of the story. When the Hilbert car arrives, the manager can welcome everyone using the previous strategy (*e.g.* moving only the people from floor 1, or shifting everyone to

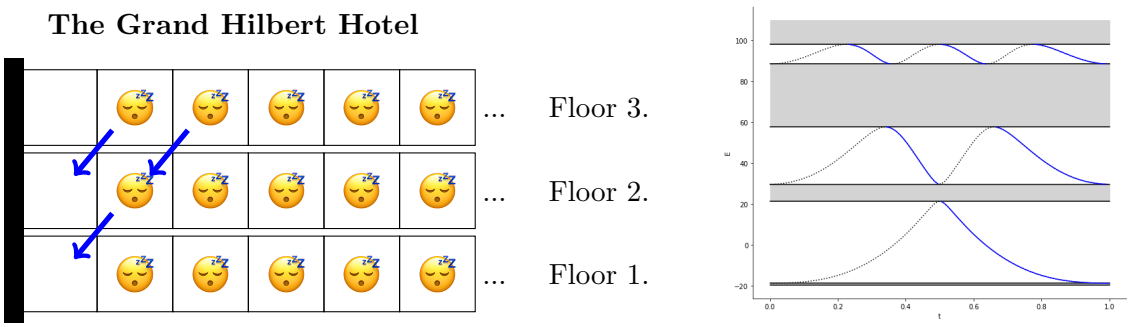


Figure 2.2: The Grand Hilbert Hotel, and the spectral flow of H_t^{cut}. Recall that $\sigma(H_t^{\mathrm{cut}}) = \sigma(H_t^{\sharp,+}) \cup \sigma(H_t^{\sharp,-})$. The blue lines reflect the spectral flow of $H_t^{\sharp,+}$, and the dotted black lines the one of $H_t^{\sharp,-}$. As t moves from 0 to 1, the wall shifts to the right, so one cell is deleted for $H_t^{\sharp,+}$, and one cell is created for $H_t^{\sharp,-}$.

the right). But another strategy is possible. He can ask one guest from the first floor to go upstairs. This frees a room at Floor 1, and allow to welcome the traveller seated in Seat 1. He would then ask two people from Floor 2 to go upstairs, hence freeing two rooms: one for the traveller in Seat 2, and one for the previous guest from Floor 1. And so on: he asks n people from Floor n to go to Floor $n+1$. At the end of the process, the traveller in Seat n occupies Floor n-Room 1, and the original guests either do not move, or move one floor upstairs and shift right. This strategy is somehow better than the one from the usual Hilbert Hotel story, since guests moves at the maximum upstairs and right, and only a finite number of people are moving for each given floor.

The next day, the manager successfully buys the land near its hotel, and constructs a new room per floor on the left. At each Floor n, a new empty "*Floor n- Room* 0" is created. In the classical story, the manager would ask everyone to shift one room to the left in order to fill all these new rooms. But in the Grand Hilbert Hotel, he can also ask someone of the second floor to go down one floor. Then, the first floor is filled, but there are now two empty rooms in the second floor. So the manager asks two people from the third floor to come down, and so on.

By analogy, this second solution explains well the spectral flow that we observe. One should think of the particles as the people in the room, and the Bloch bands as the floors. When t moves from 0 to 1, a new cell is created at the left boundary of the system. This means that a new cell is created *per Bloch band*. One needs particles to fill these new cells, and they can only come from upper energy bands. The motion of people going down reflects the spectral flow, see Figure 2.2.

2.3.3 The spectrum of a general junction between two–dimensional Hamiltonians

We now describe the spectrum of a general two–dimensional Hamiltonian of the form

$$H_{s,t} := -\Delta + V^L(x-s,y)\chi(x) + V^R(x-t,y)(1-\chi(x)), \quad \text{acting on } L^2(\mathbb{R}^2),$$

where V^L and V^R are two $\mathbb{Z}^2$–periodic potentials. The cut-off function χ only depends on the x variable, and is chosen as before: on the left-side (resp. right-side), we observe the potential V^L (resp. V^R). In practice, we are only interested in a specific value of s and t, but since V^L and V^R are not related *a priori*, it makes sense to study the whole family $(s,t) \mapsto H_{s,t}$. For simplicity, we set $s=0$ and drop the s notation. So we study $t \mapsto H_t := H_{s=0,t}$.

After a Bloch transform in the y-direction, we obtain a family $(k,t) \mapsto H_{k,t}$ of operators acting on the tube Ω. Our goal is to study the spectrum of H_t for a fixed $t \in \mathbb{T}^1$. It is the union of the spectra of $k \mapsto H_{k,t}$. However, in order to understand the structure of this spectrum, it is easier to first fix $k_0 \in \mathbb{T}^1$, and consider the family $t \mapsto H_{k_0,t}$. So we study

$$H_{k_0,t} := -\Delta + V^L(x,y)\chi(x) + V^R(x-t,y)(1-\chi(x)), \quad \text{acting on } L^2(\Omega),$$

and with domain reflecting the k_0-quasi periodic boundary conditions in the y–direction. The map $t \mapsto H_{k_0,t}$ is of the junction form studied previously. We proved that the essential spectrum of $H_{k_0,t}$ is independent of t. Actually, with obvious notation,

$$\sigma_{\text{ess}}\left(\{H_{k_0,t}\}\right) = \sigma_{\text{ess}}\left(H^L_{k_0}\right) \cup \sigma_{\text{ess}}\left(H^R_{k_0}\right).$$

In addition, a spectral flow of eigenvalues appears in each essential gap. More specifically, if E is in such a gap, the spectral flow equals the number of Bloch bands of the right operator H^R below E (it is independent of the left operator H^L).

We now fix $t_0 \in \mathbb{T}^1$, and let λ_0 be an eigenvalue of H_{k_0,t_0}. By continuity in k and t, this eigenvalue can be continued in a map $(k,t) \mapsto \lambda(k,t)$. In particular, each flow of eigenvalue $t \mapsto \lambda_{k_0}(t)$ becomes, when taking the union in $k \in \mathbb{T}^1$, a flow of essential spectrum $t \mapsto \bigcup_{k\in\mathbb{T}^1} \lambda(k,t)$ for H_t. In addition, each corresponding eigenvector of $\lambda(k,t)$ is localized near the cut.

This proves the following structure for the spectrum of H_t. First, it is purely essential by Bloch theory (the operator is periodic in the y–direction), and contains $\sigma_{\text{ess}}(H^R)$ and $\sigma_{\text{ess}}(H^L)$ for all $t \in \mathbb{T}^1$. In addition, we have

$$\sigma(H_t) = \sigma_{\text{ess}}(H^L) \cup \sigma_{\text{ess}}(H^R) \cup \sigma_{\text{edge}}(H_t),$$

where we defined $\sigma_{\text{edge}}(H_t)$ the *edge spectrum* of H_t by

$$\sigma_{\text{edge}}(H_t) := \bigcup_{k\in\mathbb{T}^1} \sigma_{\text{edge}}(H_{k,t}), \qquad \sigma_{\text{edge}}(H_{k,t}) := \sigma(H_{k,t}) \setminus \sigma_{\text{ess}}(H_{k,t}).$$

This edge spectrum is purely essential. The previous discussion shows that, as t runs through $\mathbb{T}^1$, this edge spectrum exhibits a "*spectral flow of essential spectrum*". These flows are linked to the number of Bloch bands of the right operator H^R below the considered energy. These flows can overlap, can overlap with $\sigma_{\text{ess}}(H^{L/R})$, and may fill the bulk gaps for some values of t.

To sum up, we proved the following. When $t = t_0$ is fixed, we obtain an operator H_{t_0} with three types of essential spectrum. First, we have $\sigma_{\text{ess}}(H^L)$, describing modes propagating to the left and $\sigma_{\text{ess}}(H^R)$ describing modes propagating to the right. These spectra are independent of t_0. In addition, we have the edge spectrum $\sigma_{\text{edge}}(H_{t_0})$ describing modes that propagates along the cut. This edge spectrum depends on t, and exhibit a spectral flow motion as t moves from 0 to 1. A schematic figure is shown in Figure 2.3.

Since only the right potential depends on t, this spectral flow only depends on the number of bands for the right operator below E (the blue bands in Figure 2.3). By contrast, let us restore the s variable, and study the family $s \mapsto H_{s,t=0}$. In the case, only the left potential depends on s, and cells are now deleted as s moves from 0 to 1. Using our Grand Hilbert Hotel analogy, one would deduce that, for the family $s \mapsto H_s$, there is a spectral flow of eigenvalues going upwards, and depending only on the number of red bands below E.

2.3.4 Application: two–dimensional materials cut with a general angle

In this section, we describe the results in [G7].

We now consider the case of a two-dimensional periodic material which is cut at any angle. Our goal is to study the edge operator

$$H^\sharp(\theta) := -\Delta + V_\theta \quad \text{acting on } L^2(\mathbb{R}^2_+) \text{ with domain } H^2(\mathbb{R}^2_+) \cap H^1_0(\mathbb{R}^2_+).$$

This operator acts on the half space $\mathbb{R}^2_+ := \mathbb{R}_+ \times \mathbb{R}$, and has Dirichlet boundary conditions at the cut $\{0\} \times \mathbb{R}$. The potential V_θ is a θ-rotated version of a $\mathbb{Z}^2$–periodic potential V, that is

$$V_\theta(\mathbf{x}) := V(R_\theta^{-1}\mathbf{x}), \qquad R_\theta := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

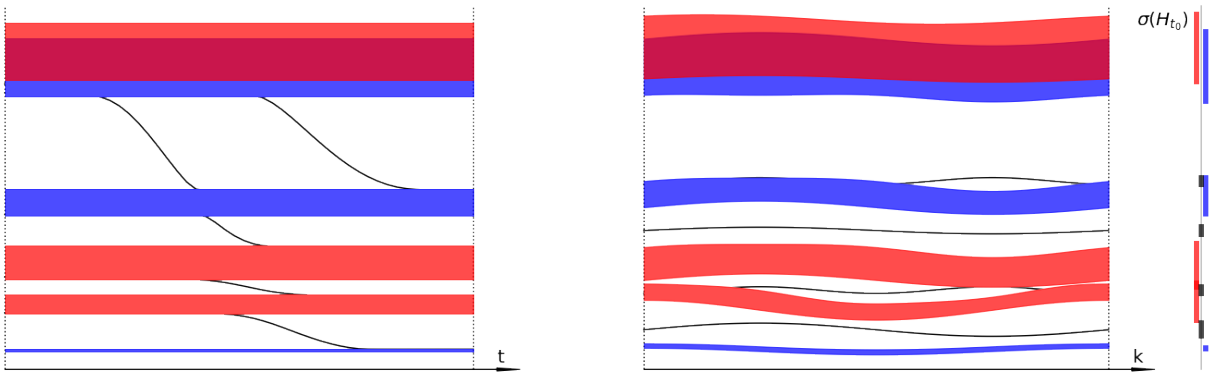


Figure 2.3: Schematic view of the spectrum of $t \mapsto H_{k_0,t}$ (left) and $k \mapsto H_{k,t_0}$ (right). The red (resp. blue) part is the essential spectrum of the left (resp. right) operator, and the black lines correspond to the edge spectrum. The spectral flow of $t \mapsto H_{k_0,t}$ equals the number of blue bands below it. The spectrum of H_{t_0} is the union of spectra of the right figure, and is the union of the left/right essential spectra, and of the edge essential spectrum.

The bulk operator $H_\theta := -\Delta + V_\theta$ acting on $L^2(\mathbb{R}^2)$ with domain $H^2(\mathbb{R}^2)$ is a rotated version of $H_{\theta=0}$, hence has a spectrum independent of θ. Its spectrum σ_{bulk} has the band-gap structure by Bloch theory.

When $\tan(\theta)$ is a rational number, of the form $\tan(\theta) = \frac{p}{q}$, the operator $H^\sharp(\theta)$ is still L–periodic in the x_2–direction, with $L := \sqrt{p^2+q^2}$, so its spectrum also has the band-gap structure. Our main result concerns the case $\tan(\theta) \notin \mathbb{Q}$.

Theorem 2.7 ([G7] Theorem 1)**.** *If* $\tan(\theta) \notin \mathbb{Q}$*, then* $\sigma(H^\sharp(\theta)) = [\Sigma, \infty)$*, where* $\Sigma := \inf \sigma(H_0)$*.*

In other words, all bulk gaps are filled with edge spectrum. Unfortunately, our proof does not provide any insight on the nature of this edge spectrum. When $\tan(\theta)$ is rational, it is absolutely continuous by Bloch theory, but when $\tan(\theta)$ is not rational, everything could happen *a priori*. For instance, one could expect pure point edge spectrum, describing Anderson localization for edge modes.

Proof. Let us give some ideas of the proof. We consider the family of bulk operators

$$H_t(\theta) := -\Delta + V_\theta(x_1 - t, x_2).$$

Assume first that $\tan(\theta) = \frac{p}{q}$ is rational, and set $L := \sqrt{p^2+q^2}$. As we already mentioned, these operators, and their cut versions, are periodic in the x_2–direction, and can be Bloch diagonalized in this direction. We obtain a family of Hamiltonians $H_{k,t}(\theta)$ acting on the channel $\mathbb{R} \times (0, L)$, with domains reflecting the k quasi–periodic boundary conditions. In what follows, we study the operator with $k = 0$, which has periodic boundary conditions. Specifically,

$$H_{0,t}(\theta) = -\Delta + V_\theta(x_1 - t, x_2) \quad \text{acting on } L^2\left(\mathbb{R} \times (L\mathbb{T}^1)\right) \text{ with domain } H^2\left(\mathbb{R} \times (L\mathbb{T}^1)\right).$$

One can also check that the map $t \mapsto H_{0,t}$ is L–periodic. Using our Grand Hilbert Hotel analogy, we see that L^2 new cells are created in the channel as t moves from 0 to L (we create a big square of area L^2, filled with unit cells of area 1). Since there are $\mathcal{N}(E)$ electrons per unit cell, we deduce as in Theorem 2.6 that

$$\mathrm{Sf}\left(H^\sharp_{0,t}, E, L\mathbb{T}^1\right) = N^2\mathcal{N}(E).$$

Simple geometrical considerations shows that the operator $H^\sharp_{0,t+\frac{1}{L}}$ is a translated version of $H^\sharp_{0,t}$ in the x_2–direction. So the spectrum of $t \mapsto H^\sharp_{0,t}$ is actually $\frac{1}{L}$–periodic, and we can write, with a slight

abuse of notation,

$$\mathrm{Sf}\Big(H^\sharp_{0,t}, E, (1/L)\mathbb{T}^1\Big) = \mathcal{N}(E). \tag{2.5}$$

This already implies that there must be numerous eigenvalues in each bulk gap.

Lemma 2.8 ([G7] Lemma 6)**.** *Assume V is ν–Lipschitz. Then, for any $t \in \mathbb{R}$, any $k \in \mathbb{R}$, and any gap (a,b) of σ_{bulk}, there are at least*

$$\left\lfloor \frac{(b-a)L}{\nu} \right\rfloor \mathcal{N}(E)$$

eigenvalues of $H^\sharp_{k,t}(\theta)$ in the essential gap (a,b).

The idea of the proof is that, since V is ν-Lipschitz, all branches of eigenvalues $t \mapsto \lambda(t)$ are ν-Lipschitz as well, hence cannot cross the gap very fast. But since these branches must produce a non-null spectral flow, they must «wind» several times, see Fig. 2.4.

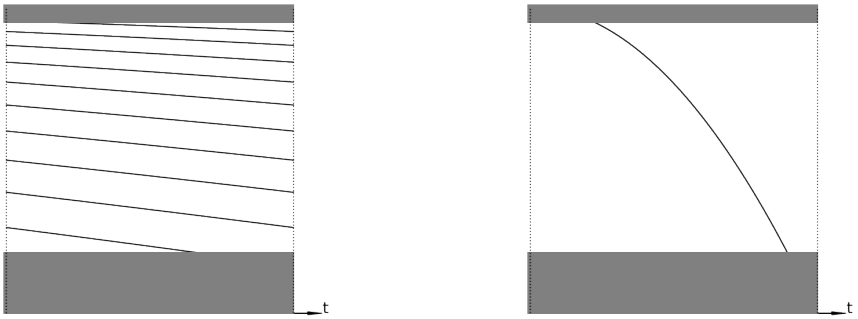


Figure 2.4: These two situations have a spectral flow of 1, but in left figure, the branch of eigenvalues have a smaller Lipschitz constant.

This result suggests that as $L \to \infty$, the edge spectrum will fill the whole gap. To make this precise, we make the following construction.

We fix $\theta \in \mathbb{R}$ so that $\tan(\theta) \notin \mathbb{Q}$, and consider a sequence $\theta_n \to \theta$ with $\tan(\theta_n) \in \mathbb{Q}$. We set $L_n := \sqrt{p_n^2 + q_n^2}$, which is a sequence diverging to ∞. Using (2.5), we deduce that for all n, there is $t_n \in [0,1]$ and ψ_n so that

$$\begin{cases} (-\Delta + V_{\theta_n,t_n})\psi_n = E\psi \quad \text{in the distributional sense } \mathcal{D}'(\mathbb{R}^2_+) \\ \psi_n \in H^2(\Omega_n^+) \cap H^1_0(\Omega_n^+), \qquad \Omega_n^+ := \mathbb{R}^+ \times (0, L_n) \\ \psi_n(x_1, x_2 + L_n) = \psi_n(x_1, x_2) \\ \|\psi_n\|_{L^\infty(\mathbb{R}^2_+)} = 1. \end{cases}$$

Note that we normalized our functions ψ_n in $L^\infty(\mathbb{R}^2_+)$ instead of the usual $L^2(\Omega_n^+)$. This is because the space $L^\infty(\mathbb{R}^2_+)$ is independent of n. Taking weak limits as $n \to \infty$ proves the existence of t_* and $\psi_* \in L^\infty(\mathbb{R}^2_+)$ satisfying

$$(-\Delta + V_{\theta,t^*})\psi_* = E\psi_*.$$

At this point, it may happen that $\psi_* = 0$, in which case one cannot conclude. In order to have a non-null limit, we need to control the mass of the sequence (ψ_n).

By general elliptic estimates and the fact that $\|\psi_n\|_\infty = 1$, there is $\delta > 0$ independent of n and a sequence $(\mathbf{x}_n)$ in $\mathbb{R}^2_+$ so that

$$\forall \mathbf{x} \in \mathcal{B}(\mathbf{x}_n, \delta), \qquad |\psi_n|(\mathbf{x}) \geq 1/2. \tag{2.6}$$

We claim that the sequence $(\mathbf{x}_n)$ can be chosen to be bounded.

First, upon shifting the whole system in the x_2–direction (which amounts to changing t_n), we may assume $x_{n,2} = 0$. Then, using Combes-Thomas estimates, we can prove that ψ_n is exponentially decaying away from the cut. This implies that there is $X > 0$ independent of n so that $0 \leq x_{n,1} \leq X$. This proves our claim. Now, the sequence $(\mathbf{x}_n)$ lies in the compact $[0, X] \times \{0\}$. Up to a subsequence, it converges to some $\mathbf{x}_* \in \mathbb{R}^2_+$. Together with (2.6), it implies that ψ_* is non null in a neighbourhood of $\mathbf{x}_*$, and in particular that $\psi_* \neq 0$. We refer to [G7] for the details. We deduce that $E \in \sigma(-\Delta + V_{\theta,t^*})$. Finally, by ergodicity in the irrational case, the spectrum of $H^\sharp_t(\theta)$ is independent of t, so $E \in \sigma(H^\sharp(\theta))$. Since this holds whenever $\mathcal{N}(E) \geq 1$, we obtain that all bulk gaps are filled with edge spectrum, as wanted. Finally, the fact that $\Sigma = \inf \sigma(H_0)$, *i.e.* the bottom of the spectrum of $H^\sharp(\theta)$ equals the one of the bulk operator H_0, comes from the fact that the core domain of $H^\sharp(\theta)$, which is $C^\infty_0(\mathbb{R}^+ \times \mathbb{R})$ is strictly contained in the one of H_θ, which is $C^\infty(\mathbb{R}^2)$. □

2.4 Perspectives

2.4.1 Soft wall models

A natural extension of these results concerns the tight-binding setting. In this case, the Hilbert space is $\ell^2(\mathbb{Z}^2)$ for the bulk–operator. It is however unclear how to define the Hilbert space for the edge operator, in particular in the context of dislocations. One would like to consider a Hilbert space of the form

$$\mathcal{H}_t := \ell^2\left((\mathbb{Z} \cap [t, \infty)) \times \mathbb{Z}\right),$$

but then the Hilbert space changes with t and we can no longer define spectral flows. One promising idea is to replace the hard truncation (encoded with Dirichlet boundary condition) by the addition of a wall W, which is a continuous function $W : \mathbb{R} \to \mathbb{R}$ with

$$\lim_{x \to -\infty} W(x) = +\infty \quad \text{and} \quad \lim_{x \to \infty} W(x) = 0.$$

We can then consider the family of edge operators

$$H^\sharp_t := -\Delta + V(x, y) + W(x + t).$$

Since $W \to \infty$ on the left, the particles can no longer propagate to the left–side, so the presence of W confines the particles on the right-side, hence the denomination *soft wall*. If V is 1–periodic in the x–direction, then $t \mapsto H^\sharp_t$ is (quasi)–periodic in t, and we can define the spectral flows of $t \mapsto H^\sharp_t$.

This soft-wall model can be extended to the tight-binding setting. In particular, the Hilbert space for the edge operator will always be $\ell^2(\mathbb{Z}^2)$, independent of t.

We started to study this soft wall model with Hanne Van Den Bosch and Camilo Gómez Araya.

2.4.2 Lagrangian setting

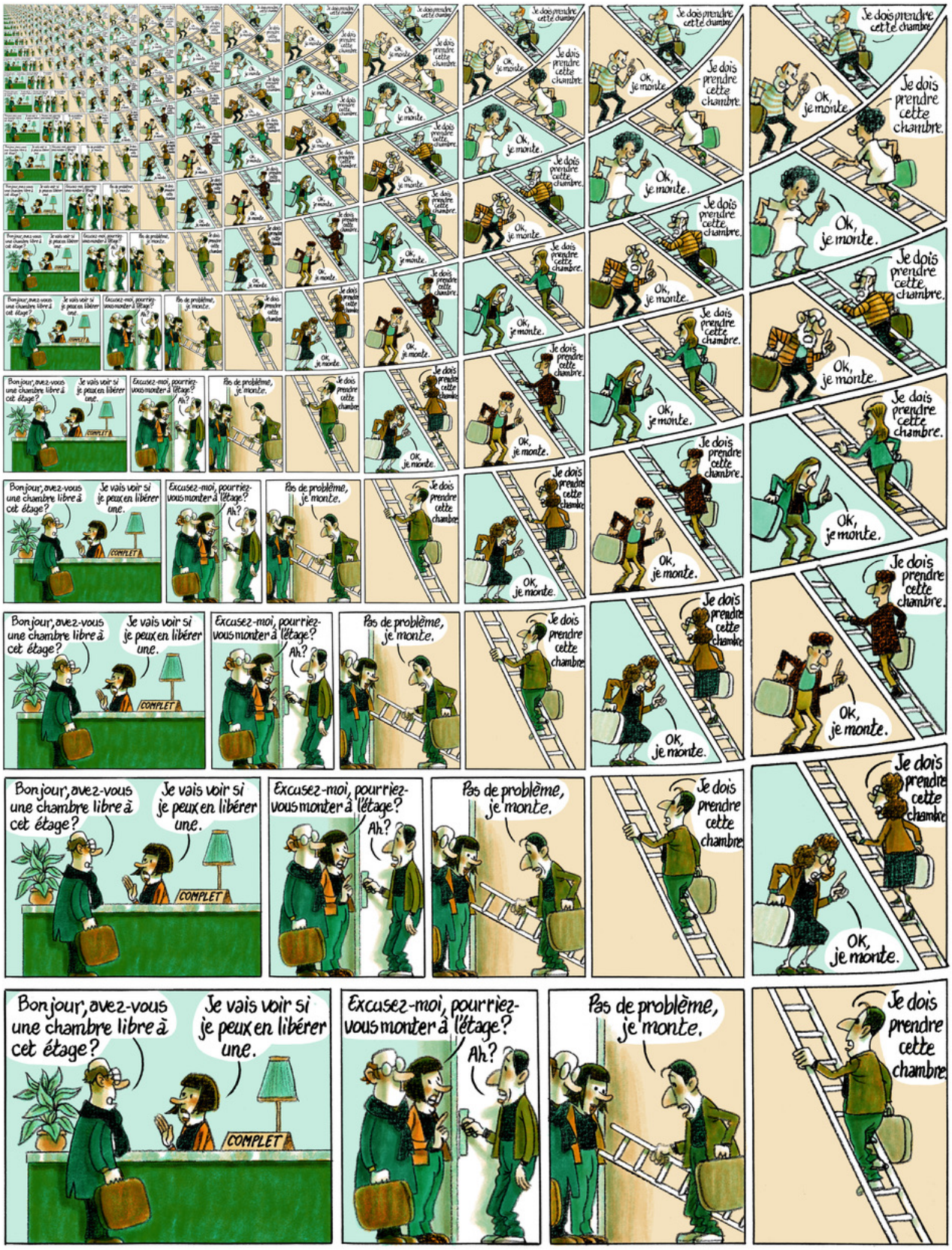
In [G6], we made the following observation that

$$\mathrm{Ker}(H^\sharp - E) = \mathrm{Ker}(H^\sharp_{\max} - E) \cap \mathcal{D}, \tag{2.7}$$

where $H^\sharp_{\max}$ is the maximal operator $H^\sharp_{\max} := -\Delta + V$ with domain $H^2(\mathbb{R}^2_+)$ (no boundary conditions), and $\mathcal{D}$ is the set of wave–vectors satisfying Dirichlet boundary conditions. The first set $\mathrm{Ker}(H^\sharp_{\max} - E)$ is a bulk quantity, since it does not depend on the specific boundary conditions at the edge, and the second set $\mathcal{D}$ is an edge quantity, since it does not depend on the potential V. So we can detect edge spectrum as the crossing between a bulk set and an edge set.

In addition, we proved that a wave vector ψ in any of these two sets can be encoded by its boundary value at the cut (say $(\psi(0), \psi'(0))$ in the one-dimensional setting). As a consequence, we are able to reformulate (2.7) as the crossing of two *Lagrangian planes* in a symplectic boundary space. In the case where $H^\sharp$ depends periodically on t, the spectral flow of $t \mapsto H_t^\sharp$ becomes a Maslov index for the corresponding Lagrangian planes.

We believe that one can interpret the Kitaev table (see Table 1.1) as the classification of Lagrangian planes satisfying some symmetries. This would eventually prove the common belief that *edge modes always appear in the junction of two 2d materials having different bulk indices*. Note that in the previous Section 1.3.3, we study homotopies within the class of bulk operators (or rather bulk projectors). The classification we obtained does not provide any information about non-periodic systems, such as a junction between two different periodic materials.



The Grand Hilbert Hôtel (by Étienne Lécroart).

CHAPTER 3

THE HARTREE–FOCK GAS

3.1 Introduction

In this chapter, we study the electron gas, also called quantum Fermi sea, or jellium. The goal is to describe the gas of interacting fermions with average density $\rho \geq 0$, and at a temperature $T \geq 0$, put into an homogeneous positive and neutralizing background. This model was first introduced by Wigner [Wig38], and is one of the main object behind Density Functional Theory [KS65; LLS19]. In the Coulomb case, the existence of the thermodynamic limit, for all reasonable physical quantities, was proved by Lieb and Narnhofer in [LN75], and precise Quantum Monte–Carlo computations have been performed by Ceperley and Alder in [CA80] (more recent and more precise numerical simulations are available in [CBC04; HM20]).

Here is what is expected (but not proved) for this jellium (see also Figure 3.2 below). At low temperature and low density, the gas is crystallized: due to the low–density of the electrons, they somehow keep their particle identity, and arrange themselves on some lattice, called the Wigner lattice. As the density increases, there is a phase transition from the Wigner crystal to the fluid phase, where the jellium is invariant by translations. This fluid is first ferromagnetic, but becomes smoothly paramagnetic at high density (Stoner transition). In addition, we expect the jellium to always be a paramagnetic fluid when the temperature is high enough, or when the density is high enough. The lowest critical temperature temperature T_c for which the gas is a paramagnetic fluid for all densities is called the Curie temperature.

In this chapter, we focus on a simpler model to describe the Fermi sea, namely the *Hartree–Fock* gas. In this situation, the electronic state is described by a one-body density matrix γ, which is a self adjoint operator satisfying the Pauli principle $0 \leq \gamma \leq 1$. This model can be obtained from the general quantum setting by restricting all minimizations to Slater determinants. To define this model, we use a standard thermodynamic limit.

For $L > 0$, we denote by $L\mathbb{T}^d := [-\frac{1}{2}L, \frac{1}{2}L]^d$ a box of size L, seen as a torus (that is with periodic boundary conditions). The set of supercell one-body density matrices with average density $\rho > 0$ is defined by

$$\mathcal{P}_L(\rho) := \left\{ \gamma \in \mathcal{S}(L^2(L\mathbb{T}^d, \mathbb{C}^2)),\ 0 \leq \gamma \leq 1,\ \mathrm{Tr}(\gamma) = \rho L^d \right\}.$$

Here, $\mathbb{C}^2$ stands for the spin, and we often write $\gamma = \begin{pmatrix} \gamma^{\uparrow\uparrow} & \gamma^{\uparrow\downarrow} \\ \gamma^{\downarrow\uparrow} & \gamma^{\downarrow\downarrow} \end{pmatrix}$. For such an operator, the electronic density of γ is $\rho_\gamma(\mathbf{x}) = \mathrm{tr}_{\mathbb{C}^2}\gamma(\mathbf{x}, \mathbf{x})$, where $\gamma(\mathbf{x}, \mathbf{y})$ is the kernel of the operator γ. The equality $\mathrm{Tr}(\gamma) = \rho L^d$ states that γ represents an electronic state with ρL^d particles in the box $L\mathbb{T}^d$, hence with an average density ρ. For $\gamma \in \mathcal{P}_L(\rho)$, we consider the (supercell) *Hartree–Fock free energy per unit*

volume

$$\mathcal{F}^{\mathrm{HF}}_{\rho,T,L}(\gamma) := \frac{1}{L^d}\Big[\frac{1}{2}\mathrm{Tr}_L\left(-\Delta_L\gamma\right) + \frac{1}{2}\iint_{(L\mathbb{T}^d)^2}(\rho_\gamma-\rho)(\mathbf{x})(\rho_\gamma-\rho)(\mathbf{y})w_L(\mathbf{x}-\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} \\ - \frac{1}{2}\iint_{(L\mathbb{T}^d)^2}\mathrm{tr}_{\mathbb{C}^2}|\gamma(\mathbf{x},\mathbf{y})|^2 w_L(\mathbf{x}-\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y} + T\,\mathrm{Tr}_L\,S(\gamma)\Big]. \tag{3.1}$$

The first term is the kinetic energy of the electrons, the second term is the Hartree energy, which represents a mean-field interaction between the electrons and the constant positive background of constant density ρ. The third term is the Fock term, and amounts to a fermionic correction of the Hartree term. The last term is the entropy, with the usual (convex) fermionic entropy function

$$S(x) := x\log(x) + (1-x)\log(1-x).$$

When $T = 0$, the entropy term is absent, and we obtain the *Hartree–Fock energy per unit volume*. In what follows, we will mainly focus on potentials w_L which are periodized Riesz interactions. In this case, w_L is defined as the inverse Fourier transform of

$$\forall \mathbf{k} \in \left(\tfrac{2\pi}{L}\right)\mathbb{Z}^d \setminus \{\mathbf{0}\}, \quad \widehat{w}_L(\mathbf{k}) := \frac{c_{d,s}}{|\mathbf{k}|^{d-s}}, \qquad \widehat{w_L}(\mathbf{0}) = c_L, \qquad \text{with} \quad c_{d,s} := \frac{2^{\frac{d-s}{2}}\Gamma\left(\frac{d-s}{2}\right)}{2^{\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)}. \tag{3.2}$$

The null Fourier coefficient $\widehat{w}_L(\mathbf{0}) = c_L$, which corresponds to the average of w_L, is chosen so that $\min w_L = 0$. Note that $\widehat{w}_L(\mathbf{k}) = \widehat{w}(\mathbf{k})$ is independent of L, and that we have the pointwise convergence $w_L(\mathbf{x}) \to w(\mathbf{x})$, where

$$\widehat{w}(\mathbf{k}) = \frac{c_{d,s}}{|\mathbf{k}|^{d-s}} \quad \text{is the Fourier transform of} \quad w(\mathbf{x}) := \frac{1}{|\mathbf{x}|^s}.$$

The usual Coulomb case corresponds to $d = 3$ and $s = d-2 = 1$. For $\rho, T \geq 0$, we define the Hartree–Fock (free) energy per unit cell as the limit

$$F^{\mathrm{HF}}(\rho,T) := \liminf_{L\to\infty} F_L(\rho,T), \qquad \text{with} \quad F^{\mathrm{HF}}_L(\rho,T) := \inf\left\{\mathcal{F}^{\mathrm{HF}}_{\rho,T,L}(\gamma),\ \gamma \in \mathcal{P}_L(\rho)\right\}.$$

There are two parameters in the model, namely the density $\rho > 0$ and the temperature $T \geq 0$, and our goal is to describe phase transitions of the Hartree–Fock energy, in the (ρ, T) plane.

3.2 Phase diagram in the translation–invariant setting

In this section, we describe the results of [G8]. This is joint work with Mathieu Lewin.

We first focus on a simpler case where we restrict γ to commute with all translations. In terms of kernel, this means that $\gamma(\mathbf{x},\mathbf{y}) = \gamma(\mathbf{x}-\mathbf{y})$. In this case, the density of γ is constant: $\rho(\mathbf{x}) = \gamma(\mathbf{0})$. In particular, we have $\rho(\mathbf{x}) = \rho$, and the Hartree term vanishes in (3.1). In addition, γ is a Fourier multiplier. In what follows, we denote the corresponding Fourier multiplier by

$$G(\mathbf{k}) := (2\pi)^{d/2}\widehat{\gamma}(\mathbf{k}) = \int_{\mathbb{R}^d}\gamma(\mathbf{x})\mathrm{e}^{-\mathrm{i}\mathbf{k}\cdot\mathbf{x}}\mathrm{d}\mathbf{k} = \begin{pmatrix} g^{\uparrow\uparrow}(\mathbf{k}) & g^{\uparrow\downarrow}(\mathbf{k}) \\ g^{\downarrow\uparrow}(\mathbf{k}) & g^{\downarrow\downarrow}(\mathbf{k})\end{pmatrix}.$$

The operator inequality $0 \leq \gamma \leq 1$ translates into the pointwise inequality $0 \leq G(\mathbf{k}) \leq \mathbb{I}_2$, and the density condition $\rho = \gamma(\mathbf{0})$ becomes $(2\pi)^{-d}\int \mathrm{tr}_{\mathbb{C}^2}G = \rho$. Taking the limit $L \to \infty$ in the free energy gives the translation invariant limit (we use the *tilde* notation for translation invariant quantities)

$$\widetilde{\mathcal{F}}^{\mathrm{HF}}_T(G) = \frac{1}{2(2\pi)^d}\int_{\mathbb{R}^d}|\mathbf{k}|^2\mathrm{tr}_{\mathbb{C}^2}\left[G(\mathbf{k})\right]\mathrm{d}\mathbf{k} - \frac{1}{2(2\pi)^d}\int_{\mathbb{R}^d}\mathrm{tr}_{\mathbb{C}^2}\left[G(\mathbf{k})G(\mathbf{k}')\right]W(\mathbf{k}-\mathbf{k}')\mathrm{d}\mathbf{k}\mathrm{d}\mathbf{k}' \\ + \frac{T}{(2\pi)^d}\int_{\mathbb{R}^d}\mathrm{tr}_{\mathbb{C}^2}S(G(\mathbf{k}))\mathrm{d}\mathbf{k}.$$

Note that the Fock term $\int |\gamma(\mathbf{x}-\mathbf{y})|^2 w(\mathbf{x}-\mathbf{y})\mathrm{d}\mathbf{x}$ (independent of $\mathbf{y}$) has a convolution form in Fourier space. Here, we set $W(\mathbf{k}) = (2\pi)^{-d/2}\widehat{w}(\mathbf{k})$. It is the Fourier transform of the interaction, up to a $(2\pi)^{-d/2}$ factor. In the case of Riesz potentials for instance, we have $W(\mathbf{k}) = \dfrac{\widetilde{c}_{d,s}}{|\mathbf{k}|^{d-s}}$, with $\widetilde{c}_{d,s} = (2\pi)^{-d/2}c_{d,s}$. The argument G is now a 2×2 hermitian matrix valued function (instead of an operator). We obtain the minimization problem

$$\boxed{\widetilde{F}^{\mathrm{HF}}(\rho,T) := \inf\left\{\widetilde{\mathcal{F}}^{\mathrm{HF}}_T(G), \quad 0\le G(\mathbf{k}) = G(\mathbf{k})^* \le \mathbb{I}_2, \ \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} \mathrm{tr}_{\mathbb{C}^2} G = \rho\right\}.}$$

3.2.1 Reduction to the no-spin case

We say that G is *diagonal* if it is of the form

$$G(\mathbf{k}) = U\begin{pmatrix} g^{\uparrow}(\mathbf{k}) & 0 \\ 0 & g^{\downarrow}(\mathbf{k})\end{pmatrix}U^*, \quad \text{for some } U\in \mathrm{SU}(2) \text{ independent of } \mathbf{k}\in\mathbb{R}^d. \tag{3.3}$$

When $g^{\uparrow} = g^{\downarrow}$, G describes a *paramagnetic state*, when $g^{\downarrow} = 0$ or $g^{\uparrow} = 0$, it describes a *ferromagnetic state*, and G describes a *general ferromagnetic state* otherwise.

For G a diagonal state, its energy simplifies into

$$\widetilde{\mathcal{F}}^{\mathrm{HF}}_T(G) = \widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}(g^{\uparrow}) + \widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}(g^{\downarrow}),$$

with the no–spin free energy

$$\widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}(g) := \frac{1}{2(2\pi)^d}\int_{\mathbb{R}^d}|\mathbf{k}|^2 g(\mathbf{k})\mathrm{d}\mathbf{k} - \frac{1}{2(2\pi)^d}\int_{\mathbb{R}^d} g(\mathbf{k})g(\mathbf{k}')W(\mathbf{k}-\mathbf{k}')\mathrm{d}\mathbf{k}\mathrm{d}\mathbf{k}' + \frac{T}{(2\pi)^d}\int_{\mathbb{R}^d} S(g(\mathbf{k}))\mathrm{d}\mathbf{k}.$$

The corresponding no–spin minimization problem is

$$\widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho,T) := \inf\left\{\widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}(g), \quad 0\le g(\mathbf{k})\le 1, \ \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} g = \rho\right\}.$$

Lemma 3.1 ([G8] Lemma 2.1 and Theorem 2.2)**.**
• *Assume $W\in L^1(\mathbb{R}^d)+L^\infty(\mathbb{R}^d)$. For all $\rho>0$ and $T\ge 0$, the minimization problems defining $\widetilde{F}^{\mathrm{HF}}(\rho,T)$ and $\widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho,T)$ are well-posed, and have minimizers.*
• *Assume in addition that $W>0$. Then any minimizer of $\widetilde{F}^{\mathrm{HF}}(\rho,T)$ is diagonal, of the form (3.3). In particular,*

$$\widetilde{F}^{\mathrm{HF}}(\rho,T) = \inf_{0\le t\le 1}\left\{\widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(t\rho,T) + \widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}((1-t)\rho,T)\right\}. \tag{3.4}$$

• *Any minimizer g of $\widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho,T)$ satisfies the Euler-Lagrange equation*

$$g(\mathbf{k}) = \begin{cases}\left(1+\mathrm{e}^{\frac{1}{T}\left(\frac{1}{2}|\mathbf{k}|^2 - g*W(\mathbf{k})-\mu\right)}\right)^{-1} & \text{if } T>0\\ \mathbb{1}\left(\frac{1}{2}|\mathbf{k}|^2 - g*W(\mathbf{k})<\mu\right) + \widetilde{g}(\mathbf{k}) & \text{if } T=0,\end{cases} \tag{3.5}$$

where μ is the Fermi level, chosen so that $(2\pi)^{-d}\int g = \rho$. In the latter case $T=0$, we have $\widetilde{g}(\mathbf{k})\neq 0$ only for $\mathbf{k}$ so that $\left(\frac{1}{2}|\mathbf{k}|^2 - g*W(\mathbf{k}) = \mu\right)$. We have similar Euler-Lagrange equations for minimizers of $\widetilde{F}^{\mathrm{HF}}$.
• *If W is radial decreasing, then so is any minimizer of $\widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho,T)$.*

The optimal t in (3.4) is sometimes called the *polarization* of the gas. If $t=\frac{1}{2}$, the gas is paramagnetic, and if $t=0$ (or $t=1$) it is ferromagnetic.

Proof. The first and third points are standard. Note that from the conditions $0 \leq G \leq 1$ and $(2\pi)^{-d} \int \mathrm{tr}_{\mathbb{C}^2} G = \rho$, G is bounded in $L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d)$, so the convolution $G * W$ is well–defined whenever $W \in L^1(\mathbb{R}^d) + L^\infty(\mathbb{R})$. For the last point, if W is radial decreasing, then $\widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}(g^*) \leq \widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}(g)$, where g^* is the symmetric decreasing re-arrangement of g. Finally, the second point comes from the following inequality: if $D_1 = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \mu_1 \end{pmatrix}$ and $D_2 = \begin{pmatrix} \lambda_2 & 0 \\ 0 & \mu_2 \end{pmatrix}$ are two diagonal matrices with $\lambda_i \geq \mu_i$, then, for any $U \in \mathrm{SU}(2)$, we have

$$\mathrm{tr}_{\mathbb{C}^2}\left[D_1 D_2\right] - \mathrm{tr}_{\mathbb{C}^2}\left[D_1 U D_2 U^*\right] = (\lambda_1 - \mu_1)(\lambda_2 - \mu_2)[1 - |U_{11}|^2] \geq 0.$$

In particular, writing the diagonalization of the hermitian matrix $G(\mathbf{k})$ as $G(\mathbf{k}) = U(\mathbf{k})\widetilde{G}(\mathbf{k})U^*(\mathbf{k})$, where $\widetilde{G}(\mathbf{k})$ is of the form (3.3) with $g^\uparrow \geq g^\downarrow$, we obtain

$$\widetilde{\mathcal{F}}^{\mathrm{HF}}_T(G) - \widetilde{\mathcal{F}}^{\mathrm{HF}}_T(\widetilde{G}) = \frac{1}{2(2\pi)^d} \times \int_{\mathbb{R}^d} \left\{ \mathrm{tr}_{\mathbb{C}^2}\left[\widetilde{G}(\mathbf{k})\widetilde{G}(\mathbf{k}')\right] - \mathrm{tr}_{\mathbb{C}^2}\left[\widetilde{G}(\mathbf{k})U^*(\mathbf{k})U(\mathbf{k}')\widetilde{G}(\mathbf{k}')U^*(\mathbf{k}')U(\mathbf{k})\right] \right\} W(\mathbf{k}-\mathbf{k}')\mathrm{d}\mathbf{k}\mathrm{d}\mathbf{k}' \geq 0.$$

□

Equation (3.4) shows that one can focus on the no–spin functional $\widetilde{\mathcal{F}}^{\mathrm{HF}}_{T,\mathrm{no-spin}}$. We make the important remark that if G is a minimizer of the spin–polarized free energy $\widetilde{\mathcal{F}}^{\mathrm{HF}}_T$ of the diagonal form (3.3), then the Euler–Lagrange equation for G takes the form

$$g^\uparrow = \left(1 + \mathrm{e}^{\frac{1}{T}\left(\frac{1}{2}|\mathbf{k}|^2 - g^\uparrow * W(\mathbf{k}) - \mu\right)}\right)^{-1} \quad \text{and} \quad g^\downarrow = \left(1 + \mathrm{e}^{\frac{1}{T}\left(\frac{1}{2}|\mathbf{k}|^2 - g^\downarrow * W(\mathbf{k}) - \mu\right)}\right)^{-1}.$$

So both $g^\uparrow$ and $g^\downarrow$ satisfies (3.5) for the same Fermi level μ. In particular, if $g^\uparrow \neq g^\downarrow$, then the Euler–Lagrange equations (3.5) must have at least two solutions for this μ. Conversely, if μ is such that the Euler–Lagrange equations (3.5) have a unique solution, then $g^\uparrow = g^\downarrow$, and G is a paramagnetic state.

3.2.2 The free Fermi gas at null temperature

We first focus on the case $T = 0$. Our main result in the no–spin setting is the following.

Theorem 3.2 ([G8] Theorem 2.6)**.** *Assume $W \in L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$ is radial decreasing positive. Then at $T = 0$, the no–spin problem has a unique minimizer given by*

$$g_0(\mathbf{k}) := \mathbb{1}\left(|\mathbf{k}| \leq k_F\right), \qquad \textit{with} \quad k_F := 2\pi \left(\frac{d}{|\mathbb{S}^{d-1}|}\right)^{1/d} \rho^{1/d}.$$

We call this state the **free Fermi gas** at null temperature.

Proof. Let us give a short proof of this fact. Since W is radial decreasing, so is any minimizer g. This implies that $\mathbf{k} \mapsto \frac{|\mathbf{k}|^2}{2} - g * W(\mathbf{k})$ is a radial increasing function. In particular, the Euler–Lagrange equation (3.5) at $T = 0$ shows that g is also of the form $g = \mathbb{1}\left(|\mathbf{k}| \leq k_F\right) + \widetilde{g}(\mathbf{k})$. The level lines of $|\mathbf{k}|$ are spheres, hence of null Lebesgue measure, so $\widetilde{g} = 0$ a.e. Finally, the value of k_F is found with the constraint that $(2\pi)^{-d} \int g = \rho$.

Let us give another proof in the case where $\widehat{W}$ is also positive (this is the case for Riesz potentials). Then, the energy $g \mapsto \widetilde{\mathcal{F}}^{\mathrm{HF}}_{T=0,\mathrm{no-spin}}(g)$ is concave in g. So g must saturates the constraint $0 \leq g \leq 1$, that is $g(\mathbf{k}) \in \{0, 1\}$. Since g is radial decreasing, g can only be the characteristic function of a ball. □

Theorem 3.2 states that the free Fermi gas is always the minimizer of the translational invariant Hartree–Fock gas, for any *repulsive* interaction w, that is for any w so that $\widehat{w} = (2\pi)^{d/2}W$ is positive radial decreasing. This theorem allows to compute exactly the function $\rho \mapsto \widetilde{F}^{\rm HF}_{\rm no-spin}(\rho)$ (hence also $\widetilde{F}^{\rm HF}$ using (3.4)). In the Riesz case where $W(\mathbf{k}) = \frac{\widetilde{c_{d,s}}}{|\mathbf{k}|^{d-s}}$, we find that

$$\widetilde{F}^{\rm HF}_{\rm no-spin}(\rho, T=0) = c_{\rm TF}^{\rm no-spin}(d)\rho^{1+\frac{2}{d}} + \lambda(d,s)\rho^{1+\frac{s}{d}},$$

for some (undisplayed) positive constants $c_{\rm TF}^{\rm no-spin}(d)$ and $\lambda(d,s)$. Note that for large ρ, the kinetic energy $\rho^{1+\frac{2}{d}}$ dominates only in the case $s < 2$. When the spin is taken into account, we can prove the following at $T = 0$ (see [G8, Theorem 3.1]):

- If $0 < s < \min(2, d)$, there is a sharp phase transition at some $\rho_c > 0$, where the system is ferromagnetic for $\rho < \rho_c$ and paramagnetic for $\rho > \rho_c$;
- if $2 < s < d$, there is a smooth phase transition. There is $0 < \rho_{c,1} < \rho_{c,2} < \infty$ so that the system is ferromagnetic for $\rho \leq \rho_{c,1}$, general ferromagnetic if $\rho_{c,1} < \rho < \rho_{c,2}$, and paramagnetic for $\rho \geq \rho_{c,2}$.

In particular, without assuming further constraints on the potential W, paramagnetism is not always the best scenario at high densities. Also, when the potential is more complex, several spin phase transitions may appear. We display in [G8] such a situation using an interaction of the form $W(\mathbf{k}) = \frac{\kappa_1}{|\mathbf{k}|^{d-s_1}} + \frac{\kappa_2}{|\mathbf{k}|^{d-s_2}}$ with $\kappa_1, \kappa_2 \geq 0$ and $s_1 < 2 < s_2$, that is for a combination of two Riesz potentials.

3.2.3 Phase transitions with temperature

When the temperature is positive, we observe phase transitions between (generalized) ferromagnetic states to paramagnetic states. We focus only on the case $d \geq 2$ and $1 \leq s < 2$ in what follows. The case $1 < s < 2$ is sometimes called the short range case, while the case $s = 1$ is critical and called the long-range case. It corresponds to the important Coulomb case in dimension $d = 3$. We start with the no–spin situation.

Theorem 3.3 (No spin free energy, [G8] Theorem 2.9)**.**
Short range case. Assume $d \geq 2$, and that $W \in L^1(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$ is radial decreasing, and satisfies

$$0 < W(\mathbf{k}) \leq \frac{\kappa_1}{|\mathbf{k}|^{d-s_1}} + \frac{\kappa_2}{|\mathbf{k}|^{d-s_2}}, \qquad \text{for some} \quad \kappa_1, \kappa_2 \geq 0, \quad 1 < s_1 \leq s_2 < 2.$$

Then there is $C \geq 0$ and $\rho_C \geq 0$ so that, for all (ρ, T) in the region

$$\Omega := \left\{(\rho, T) \in \mathbb{R}^+ \times \mathbb{R}^+, \quad T \geq C\rho^{s_1/d}, \quad \text{or} \quad \rho \geq \rho_C\right\},$$

The problem defining $F^{\rm HF}_{\rm no-spin}(\rho, T)$ has a unique minimizer.

Long range case. Assume $d \geq 2$ and that $W(\mathbf{k}) = \kappa|\mathbf{k}|^{1-d}$ with $\kappa > 0$. Then there is $C \geq 0$ and $\alpha > 0$ such that similar result holds on the region

$$\Omega := \left\{(\rho, T) \in \mathbb{R}^+ \times \mathbb{R}^+, \quad T \geq C\rho^{1/d}\mathrm{e}^{-\alpha\rho^{1/d}}\right\}.$$

Proof. Let us highlight some ideas of the proof. We focus on the Euler–Lagrange equations (3.5), which we recall are equations on the two variables (g, μ), and take the form

$$g(\mathbf{k}) = \left[1 + \mathrm{e}^{\frac{1}{T}\left(\frac{|\mathbf{k}|^2}{2} - g*W(\mathbf{k}) - \mu\right)}\right]^{-1}, \qquad \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} g = \rho.$$

The Fermi level μ is chosen to ensure the normalization condition $(2\pi)^{-d}\int g = \rho$.

Step 1. Let us first fix $T > 0$. For our analysis, it is interesting to fix μ and to relax the normalization condition, so we study the first equation with a fixed $\mu \in \mathbb{R}$. It is of the form $g = \mathcal{G}_{\mu,T}(g)$, with $\mathcal{G}_{\mu,T}$ an increasing function, in the sense that

$$0 \leq g_1 \leq g_2 \leq 1 \quad \Longrightarrow \quad 0 < \mathcal{G}_{\mu,T}(g_1) \leq \mathcal{G}_{\mu,T}(g_2) < 1.$$

Consider the sequence $g_0 = 0$, $g_1 = \mathcal{G}_{\mu,T}(g_0) > 0$, and $g_{n+1} = \mathcal{G}_{\mu,T}(g_n)$. This sequence is increasing (since $g_1 > g_0$), and is bounded above by 1, hence converges pointwise to some $g_{\min}$. In addition, $g_{\min}$ is a fixed point of $\mathcal{G}_{\mu,T}$.

Similarly, consider the sequence $\widetilde{g}_0 = 1$, $\widetilde{g}_1 = \mathcal{G}_{\mu,T}(\widetilde{g}_0) < 1$, and $\widetilde{g}_{n+1} = \mathcal{G}_{\mu,T}(\widetilde{g}_n)$ (this is slightly incorrect, since $\mathcal{G}_{\mu,T}(1)$ may be ill-defined. We refer to [G8] for the correct way to initialize $\widetilde{g}_0$). The sequence is decreasing and is bounded from below by $g_{\min}$, hence converges pointwise to some $g_{\max} \geq g_{\min}$, and $g_{\max}$ is a fixed point of $\mathcal{G}_{\mu,T}$. Any other fixed point g of $\mathcal{G}_{\mu,T}$ must satisfy the point-wise inequality $0 \leq g \leq 1$, hence, by iterating the map $\mathcal{G}_{\mu,T}$,

$$g_{\min} \leq g \leq g_{\max}.$$

The functional $\mathcal{G}_{\mu,T}$ has a unique fixed point iff $g_{\min} = g_{\max}$.

Note that the construction of $g_{\min}$ and $g_{\max}$ can be easily implemented numerically. In addition, for $\mu_1 < \mu_2$, we have $\mathcal{G}_{\mu_1,T}(\cdot) < \mathcal{G}_{\mu_2,T}(\cdot)$, hence the maps $\mu \mapsto g_{\min}(\mu)$ and $\mu \mapsto g_{\max}(\mu)$ are pointwise increasing. With obvious notation, we deduce that $\rho_{\min}(\mu) \leq \rho_{\max}(\mu)$ and that the maps $\mu \mapsto \rho_{\min/\max}(\mu)$ are strictly increasing. If μ is so that $\rho_{\min}(\mu) = \rho_{\max}(\mu)$, then $\mathcal{G}_{\mu,T}$ has a unique fixed point.

Step 2. At this point, we do not know anything about the smoothness of the maps $\mu \mapsto \rho_{\min/\max}(\mu)$, and indeed, we observe numerically that these maps can be discontinuous. However, we claim that there is a region with μ small enough or μ large enough in which these two maps are analytic. More specifically, we prove the following.

Lemma 3.4. *If g is a fixed point of $\mathcal{G}_{\mu,T}$ whose density ρ is such that (ρ, T) belongs to the region Ω of Theorem 3.3, then the implicit function theorem applies, and g belongs to a local analytic branch $(\mu, T) \mapsto g_{\mu,T}$ of fixed points for the $L^\infty(\mathbb{R}^d)$ topology.*

Proof. We consider $g \mapsto g - \mathcal{G}_{\mu,T}(g)$ as a map from $L^\infty(\mathbb{R}^d)$ to itself. The linearization of this map around the fixed point g can be written as

$$\mathcal{K}_g = 1 - A_g, \qquad \text{with} \quad A_g : v \mapsto \frac{1}{T} W * [g(1-g)v]\,.$$

Since $0 \leq g \leq 1$ pointwise and since W is positive, the kernel of the operator A_g is positive. In particular, since $f \leq |f| \leq \|f\|_\infty 1$, we have $|A(f)| \leq A(|f|) \leq \|f\|_\infty A(1)$, hence

$$\|A_g\|_{L^\infty \to L^\infty} = \|A(1)\|_\infty = \frac{1}{T}\|W * [g(1-g)]\,\|_\infty.$$

We claim that under the condition $(\rho, T) \in \Omega$, we have $\|A_g\|_{L^\infty \to L^\infty} < 1$, hence $\mathcal{K}_g$ is invertible, and the implicit function theorem implies. This would prove our Lemma.

For instance, in the case $W \leq \frac{\kappa}{|\mathbf{k}|^{d-s}}$, we have, using $0 \leq g \leq 1$ and $(2\pi)^{-d}\int g = \rho$, that

$$\begin{aligned}\|A_g\|_{L^\infty \to L^\infty} &\leq \frac{1}{T}\|W * g\|_\infty \leq \frac{\kappa}{T}\left\||\mathbf{k}|^{s-d} * g\right\|_\infty \\ &\leq \frac{\kappa}{T}\sup\left\{\left\||\mathbf{k}|^{s-d} * g\right\|_\infty,\ 0 \leq g \leq 1,\ \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d} g = \rho\right\}.\end{aligned}$$

By rearrangement, the last supremum is obtained for g radial decreasing, and the bathtub principle shows that it is maximized for $g(\mathbf{k}) = \mathbb{1}(|\mathbf{k}| < k_F)$ with $k_F = c\rho^{1/d}$. This gives

$$\|A_g\|_{L^\infty \to L^\infty} \leq \frac{\kappa}{T} c\rho^{s/d}$$

which is strictly smaller than 1 if $T > c\kappa\rho^{s/d}$. This is one of the condition in the definition of Ω. We refer to [G8] for the proof in the remaining of Ω. □

Step 3. We denote by

$$\Omega^\mu_{\min/\max} := \left\{(\mu, T) \in \mathbb{R} \times \mathbb{R}^+, \qquad \text{so that } (\rho_{\min/\max}(T, \mu), T) \in \Omega\right\}.$$

The previous lemma shows that the maps $(\mu, T) \mapsto \rho_{\min/\max}(\mu, T)$ are analytic on $\Omega^\mu_{\min/\max}$. In addition, since $\mu \mapsto \rho_{\min/\max}(\mu)$ is strictly increasing, this map is invertible with analytic inverse. Since Ω is connected, we deduce that $\Omega^\mu_{\min/\max}$ are also connected sets.

Next, we can prove that there is $(\mu, T) \in \Omega^\mu_{\min/\max}$ for which $\mathcal{G}_{\mu,T}$ is a contraction, hence has a unique fixed point. So the two analytic branches of functions $\rho_{\min}$ and $\rho_{\max}$ must coincide on the whole set Ω. In particular, $\Omega^\mu_{\min} = \Omega^\mu_{\max} =: \Omega^\mu$, and for all $(\mu, T) \in \Omega^\mu$, the function $\mathcal{G}_{\mu,T}$ has a unique fixed point. By monotonicity and analyticity of the map $\mu \mapsto \rho(\mu)$, the map $(\mu, T) \mapsto (\rho_{\min/\max}(T, \mu), T)$ is a bijection from Ω^μ to Ω.

Step 4. Finally, let us fix $(\rho, T) \in \Omega$. Let μ be the Fermi level so that $(\mu, T) \in \Omega^\mu$ with $\rho = \rho_{\min}(T, \mu) = \rho_{\max}(T, \mu)$, and let g be the unique fixed point of $\mathcal{G}_{\mu,T}$. We claim that g is the unique minimizer of $F^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho, T)$.

Let g' by any minimizer of $F^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho, T)$. Then this minimizer is solution to the Euler–Lagrange equations (3.5) for some Fermi level μ'. In particular, g' is a fixed point of $\mathcal{G}_{\mu',T}$. Assume first that $\mu < \mu'$. Since $\mu \mapsto \mathcal{G}_{\mu,T}(\cdot)$ is increasing, we have $g' \geq g_{\min}(\mu') > g_{\min}(\mu) = g$, hence $\rho' > \rho$, a contradiction. The case $\mu > \mu'$ is similar. This proves that $\mu = \mu'$, but since $\mathcal{G}_{\mu,T}$ has a unique fixed point, we must have $g' = g$. □

Finally, we conclude with the case where the spin is included.

Theorem 3.5 ([G8] Theorem 2.10)**.** *With the same assumptions on the potential W as in Theorem 3.3, there is $C \geq 0$ and $\alpha > 0$ so that, with*

$$\widetilde{\Omega} := \begin{cases} \left\{(\rho, T) \in \mathbb{R}^+ \times \mathbb{R}^+, \qquad T \geq C\rho^{s/d}\, \mathrm{e}^{-\alpha\rho^{(2-s)/d}}\right\} & \textit{(short range case } 1 < s < 2\textit{)} \\ \left\{(\rho, T) \in \mathbb{R}^+ \times \mathbb{R}^+, \qquad T \geq C\rho^{1/d}\, \mathrm{e}^{-\alpha\rho^{1/(2d)}}\right\} & \textit{(long range case } s = 1\textit{)}, \end{cases}.$$

the spin-polarized minimization problem $\widetilde{\mathcal{F}}^{\mathrm{HF}}$ has a unique minimizer, which is paramagnetic, and given by $G = g_{\rho/2,T}\mathbb{I}_2$, where $g_{\rho/2,T}$ is the (unique) minimizer of the no–spin problem with density $\rho/2$.

Proof. Let us give the main idea of the proof in the sub-region

$$\widetilde{\Omega_1} := \left\{T \geq C\rho^{s/d}\right\} \subset \widetilde{\Omega}.$$

(the proof in the other region will be explained in the next section). Recall that the map $\mu \mapsto \rho$ is increasing, hence so is the map $\rho \mapsto \mu$. In particular, since the Fermi level μ is also $\mu = \partial_\rho \widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho, T)$, the map $\rho \mapsto \widetilde{F}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho, T)$ is convex in all segments $[\rho_1, \rho_2]$ where $[\rho_1, \rho_2] \times \{T\} \in \Omega$. In particular it is convex in $\widetilde{\Omega_1}$. This gives

$$\frac{1}{2}\widetilde{\mathcal{F}}^{\mathrm{HF}}_{\mathrm{no-spin}}(t\rho) + \frac{1}{2}\widetilde{\mathcal{F}}^{\mathrm{HF}}_{\mathrm{no-spin}}((1-t)\rho) \geq \widetilde{\mathcal{F}}^{\mathrm{HF}}_{\mathrm{no-spin}}\left(\frac{1}{2}t\rho + \frac{1}{2}(1-t)\rho\right) = \widetilde{\mathcal{F}}^{\mathrm{HF}}_{\mathrm{no-spin}}(\rho/2).$$

So the minimum in (3.4) is attained for $t = \frac{1}{2}$, that is for the paramagnetic state. □

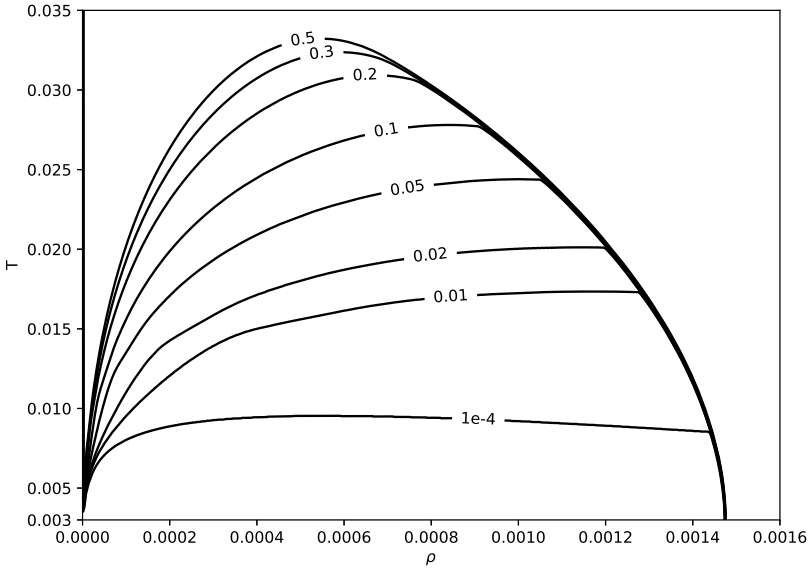


Figure 3.1: Numerical computation of the phase diagram in the translation–invariant model, for Coulomb interaction ($s = 1$) in dimension $d = 3$. We plot the level lines of the polarisation t ($t = 1/2$ is paramagnetic, and $t \approx 0$ is ferromagnetic).

In particular, in the short range case, there is T large enough so that the Hartree–Fock translation–invariant gas is always paramagnetic, for any density $\rho \geq 0$. The smallest T having this property is sometimes called the Curie temperature. We display a numerical simulation of the phase diagram in the physical case $d = 3$ and $s = 1$ in Figure 3.1. At low density, the gas is paramagnetic. As the density increases, it becomes slowly ferromagnetic (smooth phase transition), and then suddenly paramagnetic again (sharp phase transition).

3.3 The Hartree–Fock gas versus the free Fermi gas

In this section, we present the results of [G9]. This is joint work with Christian Hainzl and Mathieu Lewin.

We now go back to the Hartree–Fock gas, without assuming translation invariance. Our goal is to compare F^{HF} with its translation–invariant counterpart $\widetilde{F}^{\mathrm{HF}}$. Of course, we always have $F^{\mathrm{HF}} \leq \widetilde{F}^{\mathrm{HF}}$, and we would like to quantify the error $\left|\widetilde{F}^{\mathrm{HF}} - F^{\mathrm{HF}}\right|$.

At zero temperature, and for the three-dimensional Coulomb case ($d = 3$ and $s = 1$), it is known since the works of Overhauser [Ove60; Ove62; Ove68] that we always have the strict inequality $F^{\mathrm{HF}}(\rho, T = 0) < \widetilde{F}^{\mathrm{HF}}(\rho, T = 0)$ for all $\rho \geq 0$. Overhauser indeed proved that one can always perturb the free paramagnetic Fermi gas $\gamma_{\mathrm{FG}} := \mathbb{1}(|\mathbf{k}| \leq k_F)\mathbb{I}_2$ in order to lower the Hartree–Fock energy per unit volume. In other words, the free Fermi gas is unstable under the formation of spin and charge density waves. In [Del+15], the authors exhibit a perturbation which lowers the energy by an explicit quantity, namely

$$F^{\mathrm{HF}}(\rho, T = 0) - \widetilde{F}^{\mathrm{HF}}(\rho, T = 0) \leq -C\rho^{2/3}\mathrm{e}^{-\alpha\rho^{1/6}} \qquad (< 0), \tag{3.6}$$

for some explicit (undisplayed) constants $C > 0$ and $\alpha > 0$. In addition, detailed numerical simulations performed in [ZC08; Bag+13; Bag+14] suggest that the Hartree–Fock gas is always crystallized.

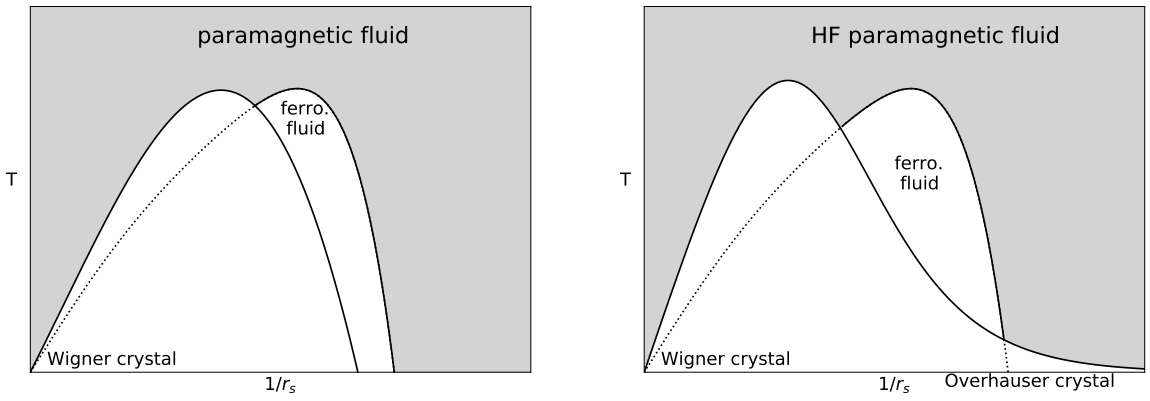


Figure 3.2: (Left) Expected form of the true jellium phase diagram. (Right) Expected form of the Hartree-Fock gas phase diagram, according to our work. At $T = 0$, the system is believed to be crystallized at all densities. This "Overhauser phase" shrinks exponentially fast to the horizontal axis at large densities. The corresponding (free) energy gain is also exponentially small.

In our work [G9], we proved a lower bound for this quantity, which has a form similar to (3.6). We only focused on the case $s = 1$ and $d = 3$, but, as we noticed in [G8], the result can be generalized. We state the first part only in the high density case.

Theorem 3.6 ([G9] for $s = 1$ and [G8] Theorem 2.9 for $1 < s < 2$)**.** *Under the same hypotheses on W as in Theorem 3.3,*

- *At $T = 0$, for all $\rho_* > 0$, there are explicit constants $C \geq 0$ and $\alpha > 0$ so that*

$$F^{\mathrm{HF}}(\rho, T = 0) - \widetilde{F}^{\mathrm{HF}}(\rho, T = 0) \geq \begin{cases} -C\mathrm{e}^{-\alpha\rho^{(2-s)/d}} & \text{(short range case } 1 < s < 2\text{)} \\ -C\mathrm{e}^{-\alpha\rho^{1/(2d)}} & \text{(long range case } s = 1\text{).} \end{cases}$$

- *There is $T_c = T_c(\rho)$ so that, for all $T > T_c$, the optimizer of $F^{\mathrm{HF}}(\rho, T)$ is invariant by translations (hence $F^{\mathrm{HF}}(\rho, T) = \widetilde{F}^{\mathrm{HF}}(\rho, T)$ in this region). For all $\rho_* > 0$, there is $C \geq 0$ and $\alpha > 0$ so that, for all $\rho > \rho_*$ the lowest critical temperature having this property satisfies*

$$T_c(\rho) \leq \begin{cases} -C\mathrm{e}^{-\alpha\rho^{(2-s)/d}} & \text{(short range case } 1 < s < 2\text{)} \\ -C\mathrm{e}^{-\alpha\rho^{1/(2d)}} & \text{(long range case } s = 1\text{).} \end{cases}$$

We display in Figure 3.2 what is expected for the phase diagrams of the jellium and for the Hartree–Fock gas respectively. At $T = 0$, it is believed, based on numerical simulations, that the true jellium is fluid at high density. We emphasize however that it could be crystallized, and that the gain of energy due to crystallization could be exponentially small (as in the HF gas). If this is the case, this phenomenon would be quite difficult to capture numerically.

3.3.1 A lower bound involving a degenerate operator

Let us explain the key ingredient for the proof of the first point of Theorem 3.6. Our goal is to prove that, at $T = 0$ the gain of energy due to the Overhauser instability is exponentially small in the density.

The proof given in [G9] is short and compact, and only focuses on the case $d = 3$ and $s = 1$. In this section, we give some extra details of the proof. We consider a general situation where the interaction w is positive and such that its Fourier transform $W = (2\pi)^{d/2}\widehat{w}$ is positive radial decreasing (this is the case for the Riesz interaction $w(\mathbf{x}) = |\mathbf{x}|^{-s}$). We denote by w_L the corresponding L–periodic potential, defined as in (3.2), and by $W_L := (2\pi)^{d/2}\widehat{w_L}$.

Recall that the Hartree–Fock energy in a supercell of size L was defined in (3.1). We introduce the supercell free Fermi gas

$$\gamma_{\mathrm{FG}}^L = G_{\mathrm{FG}}(-\mathrm{i}\nabla_L), \qquad \text{with} \quad \forall \mathbf{k} \in \tfrac{2\pi}{L}\mathbb{Z}^d, \quad G_{\mathrm{FG}}(\mathbf{k}) := \mathbb{1}(|\mathbf{k}| \leq k_F)\mathbb{I}_2,$$

and the linearized operator

$$H_{\mathrm{FG}}^L := \varepsilon_L(-\mathrm{i}\nabla_L) \quad \text{with} \quad \forall \mathbf{k} \in \tfrac{2\pi}{L}\mathbb{Z}^d, \quad \varepsilon_L(\mathbf{k}) := \frac{1}{2}|\mathbf{k}|^2 - (G_{\mathrm{FG}} *_L W_L)(\mathbf{k}).$$

Since the Hartree term in $\mathcal{F}_L^{\mathrm{HF}}$ is always positive, we can drop this term for a lower bound, and get

$$\mathcal{F}_L^{\mathrm{HF}}(\gamma, T=0) \geq Q_L(\gamma) := \frac{1}{2L^d}\left[\mathrm{Tr}_L\left(-\Delta_L\gamma\right) - \iint_{(L\mathbb{T}^d)^2} \mathrm{tr}_{\mathbb{C}^2}|\gamma(\mathbf{x},\mathbf{y})|^2 w_L(\mathbf{x}-\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\right].$$

Actually, since the Hartree term vanishes for the free Fermi gas, we also have

$$\mathcal{F}_L^{\mathrm{HF}}(\gamma_{\mathrm{FG}}^L, T=0) = Q_L(\gamma_{\mathrm{FG}}^L).$$

Let γ^L be the minimizer of the Q_L functional, with $0 \leq \gamma^L \leq 1$ and $\mathrm{Tr}_L(\gamma^L) = \mathrm{Tr}_L(\gamma_{\mathrm{FG}}^L)$. Recall that $w_L \geq 0$ is positive. In particular, $\gamma \mapsto Q_L(\gamma)$ is concave, and γ^L is a projector. A computation shows that

$$\begin{aligned} &Q_L(\gamma^L) - Q_L(\gamma_{\mathrm{FG}}^L) = \\ &\quad = \frac{1}{L^d}\left\{\mathrm{Tr}\left(H_{\mathrm{FG}}^L(\gamma^L - \gamma_{\mathrm{FG}}^L)\right) - \frac{1}{2}\iint_{(L\mathbb{T}^d)^2} \mathrm{tr}_{\mathbb{C}^2}|\gamma^L(\mathbf{x},\mathbf{y}) - \gamma_{\mathrm{FG}}^L(\mathbf{x}-\mathbf{y})|^2 w_L(\mathbf{x}-\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\right\}. \end{aligned}$$

For the first term, we use the fact that γ^L and γ_{FG}^L are projectors, hence

$$(1-\gamma_{\mathrm{FG}}^L)(\gamma^L - \gamma_{\mathrm{FG}}^L)(1-\gamma_{\mathrm{FG}}^L) - \gamma_{\mathrm{FG}}^L(\gamma^L - \gamma_{\mathrm{FG}}^L)\gamma_{\mathrm{FG}}^L = \gamma^L - \gamma^L\gamma_{\mathrm{FG}}^L - \gamma_{\mathrm{FG}}^L\gamma^L + \gamma_{\mathrm{FG}}^L = (\gamma^L - \gamma_{\mathrm{FG}}^L)^2.$$

Together with the fact that γ_{FG}^L is of the form $\gamma_{\mathrm{FG}}^L = \mathbb{1}\left(H_{\mathrm{FG}}^L \leq \mu^L\right)$ with $\mu^L := \varepsilon_L(k_F)$ (see (3.5)), and $\mathrm{Tr}_L(\gamma^L) = \mathrm{Tr}_L(\gamma_{\mathrm{FG}}^L)$, we obtain

$$\mathrm{Tr}_L\left(H_{\mathrm{FG}}^L(\gamma^L - \gamma_{\mathrm{FG}}^L)\right) = \mathrm{Tr}_L\left((H_{\mathrm{FG}}^L - \mu^L)(\gamma^L - \gamma_{\mathrm{FG}}^L)\right) = \mathrm{Tr}_L\left[|H_{\mathrm{FG}}^L - \mu^L|\left(\gamma^L - \gamma_{\mathrm{FG}}^L\right)^2\right].$$

We deduce the key equality

$$\begin{aligned} &Q_L(\gamma^L) - Q_L(\gamma_{\mathrm{FG}}^L) \\ &\quad = \frac{1}{L^d}\left\{\mathrm{Tr}\left[|H_{\mathrm{FG}}^L - \mu^L|(\gamma^L - \gamma_{\mathrm{FG}}^L)^2\right] - \frac{1}{2}\iint_{(L\mathbb{T}^d)^2} \mathrm{tr}_{\mathbb{C}^2}|\gamma^L(\mathbf{x},\mathbf{y}) - \gamma_{\mathrm{FG}}^L(\mathbf{x}-\mathbf{y})|^2 w_L(\mathbf{x}-\mathbf{y})\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\right\}. \end{aligned}$$

The left-hand side is now quadratic in the difference $\Psi^L(\mathbf{x},\mathbf{y}) := \gamma^L(\mathbf{x},\mathbf{y}) - \gamma_{\mathrm{FG}}^L(\mathbf{x}-\mathbf{y})$. We can simplify the first term as follows. Let us introduce the Fourier coefficients

$$\forall \mathbf{k}, \mathbf{k}' \in \frac{2\pi}{L}\mathbb{Z}^d, \qquad \widehat{\Psi^L}(\mathbf{k},\mathbf{k}') := \frac{1}{L^d}\iint_{(L\mathbb{T}^d)\times(L\mathbb{T}^d)} \Psi^L(\mathbf{x},\mathbf{y})\mathrm{e}^{-\mathrm{i}(\mathbf{k}\cdot\mathbf{x}+\mathbf{k}'\cdot\mathbf{y})},$$

so that $\mathbf{k}$ (resp. $\mathbf{k}'$) is the dual variable of $\mathbf{x}$ (resp. $\mathbf{y}$). In this Fourier basis, the operator H_{FG}^L has kernel $H_{\mathrm{FG}}^L(\mathbf{k},\mathbf{k}') = \varepsilon_L(\mathbf{k})\delta_{\mathbf{k}=\mathbf{k}'}$. We deduce that the first term can be expressed in the Fourier basis as

$$\mathrm{Tr}_L\left[|H_{\mathrm{FG}}^L - \mu^L|\left(\Psi^L\right)^2\right] = \sum_{\mathbf{k},\mathbf{k}' \in \frac{2\pi}{L}\mathbb{Z}^d} \left|\varepsilon_L(\mathbf{k}) - \mu^L\right| \mathrm{tr}_{\mathbb{C}^2}\left|\widehat{\Psi^L}(\mathbf{k},\mathbf{k}')\right|^2. \tag{3.7}$$

Performing the sum in $\mathbf{k}'$ first, and using Parseval, we obtain

$$\mathrm{Tr}_L\left[|H_{\mathrm{FG}}^L-\mu^L|\left(\Psi^L\right)^2\right]=\int_{L\mathbb{T}^d}\mathrm{d}\mathbf{y}\left\langle\Psi_{\mathbf{y}}^L,\left|\varepsilon_L(-\mathrm{i}\nabla_{L,\mathbf{x}})-\mu^L\right|\Psi_{\mathbf{y}}^L\right\rangle$$

where we set $\Psi_{\mathbf{y}}^L(\mathbf{x}):=\Psi^L(\mathbf{x},\mathbf{y})$. We deduce that

$$\begin{aligned}Q_L(\gamma^L)-Q_L(\gamma_{\mathrm{FG}}^L)&=\frac{1}{L^d}\int_{L\mathbb{T}^d}\mathrm{d}\mathbf{y}\left\langle\Psi_{\mathbf{y}}^L,\left|\varepsilon_L(-\mathrm{i}\nabla_{L,\mathbf{x}})-\mu^L\right|-\frac{1}{2}w_L(\mathbf{x}-\mathbf{y}),\Psi_{\mathbf{y}}^L\right\rangle_{L^2(L\mathbb{T}^d)}\\&\geq\frac{1}{L^d}\lambda_1\left(\left|\varepsilon_L(-\mathrm{i}\nabla_L)-\mu^L\right|-\frac{1}{2}w_L(\mathbf{x})\right)\int_{L\mathbb{T}^d}\|\Psi_{\mathbf{y}}^L\|_{L^2(L\mathbb{T}^d)}^2\mathrm{d}y,\end{aligned}$$

where $\lambda_1(A)$ denotes the lowest eigenvalue of an operator A. We will prove below that $\lambda_1\leq 0$. To evaluate the last integral, we use the fact that $(a+b)^2\leq 2a^2+2b^2$ and that γ^L and γ_{FG}^L are projectors, to obtain

$$\int_{L\mathbb{T}^d}\|\Psi_{\mathbf{y}}^L\|_{L^2(L\mathbb{T}^d)}^2\mathrm{d}\mathbf{y}=\iint_{(L\mathbb{T}^d)^2}\left(\gamma^L(\mathbf{x},\mathbf{y})-\gamma_{\mathrm{FG}}^L(\mathbf{x}-\mathbf{y})\right)^2\mathrm{d}\mathbf{x}\mathrm{d}\mathbf{y}\leq 2\mathrm{Tr}\left((\gamma^L)^2\right)+2\mathrm{Tr}\left((\gamma_{\mathrm{FG}}^L)^2\right)=4\rho L^d.$$

This gives the bound

$$Q_L(\gamma^L)-Q_L(\gamma_{\mathrm{FG}}^L)\geq 4\rho\lambda_1\left(|\varepsilon_L(-\mathrm{i}\nabla_L)-\mu|-\frac{1}{2}w_L(\mathbf{x})\right).$$

Finally, since both functions $\mathbf{k}\mapsto\frac{1}{2}|\mathbf{k}|^2$ and $\mathbf{k}\mapsto -G_{\mathrm{FG}}*_L W_L(\mathbf{k})$ are radial increasing, we have

$$|\varepsilon_L(\mathbf{k})-\mu|=|\varepsilon_L(\mathbf{k})-\varepsilon_L(k_F)|=\frac{1}{2}||\mathbf{k}|^2-k_F^2|+|G_{\mathrm{FG}}*W_L(\mathbf{k})-G_{\mathrm{FG}}*_L W_L(k_F)|\geq\frac{1}{2}||\mathbf{k}|^2-k_F^2|.$$

In particular, we have the operator inequality

$$|\varepsilon_L(-\mathrm{i}\nabla_L)-\mu|\geq\frac{1}{2}|-\Delta_L-k_F^2|.$$

Letting $L\to\infty$ gives the final inequality

$$\boxed{F^{\mathrm{HF}}(\rho,T=0)-\widetilde{F}^{\mathrm{HF}}(\rho,T=0)\geq 2\rho\lambda_1\left(|\Delta+k_F^2|-w(\mathbf{x})\right).}\tag{3.8}$$

It remains to study the first eigenvalue of the operator on the left-hand side. Note that this one has a degenerate kinetic part, which vanishes on the sphere $|\mathbf{k}|=k_F$.

3.3.2 Study of the degenerate operator

Recall that we assumed $w\geq 0$. In order to study the first eigenvalue of this operator, we use the Birman-Schwinger principle. We have (we write $-E$ for the eigenvalue, so that in what follows, E is a positive quantity)

$$-E\in\sigma\left(|\Delta+k_F^2|-w(\mathbf{x})\right)\quad\text{iff}\qquad 1\in\sigma\left(K_E\right),\quad\text{with}\quad K_E:=\sqrt{w}\frac{1}{|\Delta+k_F^2|+E}\sqrt{w}.$$

We prove below that K_E is a compact positive operator. In addition, $E\mapsto K_E$ is operator decreasing. Hence the lowest eigenvalue $-E_1$ is the unique energy for which $\|K_{E_1}\|_{\mathrm{op}}=1$. Let us find an upper bound for $\|K_E\|_{\mathrm{op}}$.

In order to use the surface singularity on the sphere $|\mathbf{k}|=k_F$, we follow the approach in [LSW02; HS10]. We write that

$$K_E=\frac{1}{(2\pi)^d}\int_0^\infty\frac{r^{d-1}\mathrm{d}r}{|r^2-k_F^2|+E}A_r,\quad\text{hence}\quad\|K_E\|_{\mathrm{op}}\leq\frac{1}{(2\pi)^d}\int_0^\infty\frac{r^{d-1}\mathrm{d}r}{|r^2-k_F^2|+E}\|A_r\|_{\mathrm{op}},$$

where A_r is the operator acting on $L^2(\mathbb{R}^d)$ with kernel

$$A_r(\mathbf{x},\mathbf{y}) := \sqrt{w}(\mathbf{x}) \left(\int_{\mathbb{S}^{d-1}} \mathrm{e}^{\mathrm{i}r\omega\cdot(\mathbf{x}-\mathbf{y})} \mathrm{d}\omega \right) \sqrt{w}(\mathbf{y}).$$

The operator is of the form $A_r = B_r^* B_r$ where B_r is the operator from $L^2(\mathbb{R}^d)$ to $L^2(\mathbb{S}^{d-1})$ with kernel

$$B_r(\omega,\mathbf{y}) := \mathrm{e}^{-\mathrm{i}r\omega\cdot\mathbf{y}} \sqrt{w}(\mathbf{y}).$$

This gives $\|A_r\|_{\mathrm{op}} = \|B_r^* B_r\|_{\mathrm{op}} = \|B_r B_r^*\|_{\mathrm{op}} =: \|C_r\|_{\mathrm{op}}$, where C_r is the operator from $L^2(\mathbb{S}^{d-1})$ to $L^2(\mathbb{S}^{d-1})$, with kernel

$$C_r(\omega,\omega') = \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}r(\omega-\omega')\cdot\mathbf{y}} w(\mathbf{y}) \mathrm{d}\mathbf{y} = (2\pi)^{d/2} \widehat{w}(r(\omega-\omega')).$$

Since $\widehat{w} \geq 0$, the operator C_r has a positive kernel. In addition, if $f := \mathbb{1}$ denotes the constant function on the sphere, we have

$$C_r f = (2\pi)^{d/2} \int_{\mathbb{S}^{d-1}} \widehat{w}(r(\omega-\omega')) \mathrm{d}\omega' = \left[(2\pi)^{d/2} \int_{\mathbb{S}^{d-1}} \widehat{w}(r(\omega-\mathbf{e}_1)) \mathrm{d}\omega \right] f. \tag{3.9}$$

Since $f > 0$, the Krein-Rutman theorem implies that the term in bracket is the highest eigenvalue and the operator norm of C_r. It is an eigenvalue of multiplicity 1 whose corresponding (positive) eigenvector is the constant function $f = \mathbb{1}$. Altogether, we obtain the bound

$$\|K_E\|_{\mathrm{op}} \leq \frac{1}{(2\pi)^{d/2}} \int_0^\infty \frac{r^{d-1}}{|r^2 - k_F^2| + E} \left[\int_{\mathbb{S}^{d-1}} \widehat{w}(r(\omega-\mathbf{e}_1)) \mathrm{d}\omega \right] \mathrm{d}r.$$

Short range case $1 < s < 2$. We now compute the last integral in the case of Riesz potentials $w(\mathbf{x}) = |\mathbf{x}|^{-s}$, so that $\widehat{w}(\mathbf{k}) = c_{d,s}|\mathbf{k}|^{s-d}$. We obtain directly

$$\int_{\mathbb{S}^{d-1}} \widehat{w}(r(\omega-\mathbf{e}_1)) \mathrm{d}\omega = r^{s-d} \int_{\mathbb{S}^{d-1}} \widehat{w}(\omega-\mathbf{e}_1) \mathrm{d}\omega = c_{d,s} r^{s-d} \int_{\mathbb{S}^{d-1}} \frac{1}{|\omega-\mathbf{e}_1|^{d-s}} \mathrm{d}\omega.$$

The last integral has a singularity $|\omega-\mathbf{e}_1|^{d-s}$ on the hypersurface of dimension $d-1$, but since we assumed $s > 1$, it is integrable. So our upper bound is indeed finite. This gives, in the high density limit $k_F \gg 1$,

$$\|K_E\|_{\mathrm{op}} \leq C \int_0^\infty \frac{r^{s-1}\mathrm{d}r}{|r^2-k_F^2|+E} = C k_F^{s-2} \int_0^\infty \frac{r^{s-1}\mathrm{d}r}{|r^2-1|+(E/k_F^2)} \leq C' k_F^{s-2} \left| \ln\left(\frac{k_F^2}{E}\right) \right|.$$

For the last bound, we used that the integral is finite, and has a singularity at $r = 1$ when $E/k_F \to 0$. At $E = E_1$, we have $\|K_{E_1}\|_{\mathrm{op}} = 1$. Let $E_2 > 0$ be the energy for which the right–hand side equals 1, so $E_2 = k_F \exp(-\alpha k_F^{2-s})$ and $\|K_{E_2}\|_{\mathrm{op}} \leq 1$. Since $E \mapsto \|K_E\|_{\mathrm{op}}$ is decreasing, and since $\|K_{E_1}\|_{\mathrm{op}} = 1$, we deduce that $E_1 < E_2$, that is

$$E_1 \leq k_F \exp\left(-\alpha k_F^{2-s}\right) = C\rho^{1/d} \exp\left(-\widetilde{\alpha}\rho^{\frac{2-s}{d}}\right).$$

Inserting this inequality in (3.8) proves our Theorem 3.6 in the sort range case. Note that for large ρ, the polynomial prefactor can be absorbed in the exponential part.

Long range case $s = 1$. In the case $s = 1$, the term in bracket in (3.9) is no longer a convergent integral. In particular the operators C_r, hence A_r, are no longer bounded. In order to handle this difficulty, we consider a parameter $a > 0$ (that we optimize at the end), and the function

$$w_a(\mathbf{x}) := \frac{\mathrm{e}^{-a|\mathbf{x}|}}{|\mathbf{x}|}.$$

Lemma 3.7.

- *We have the inequalities*
$$w_a(\mathbf{x}) \leq \frac{1}{|\mathbf{x}|} \leq w_a(\mathbf{x}) + a$$
- *The Fourier transform of w_a is*
$$\widehat{w}_a(\mathbf{k}) = \frac{\kappa(d)}{(|\mathbf{k}|^2 + a^2)^{\frac{d-1}{2}}}, \quad \text{with} \quad \kappa(d) := \frac{|\mathbb{S}^{d-1}|}{(2\pi)^{d/2}}\Gamma(d-1). \tag{3.10}$$

Proof. The first part comes from the inequality $1 - t \leq \mathrm{e}^{-t} \leq 1$ valid for all $t > 0$. To prove the second part, we compute the inverse Fourier transform of the left-hand side of (3.10). We set $f(\mathbf{k}) := (|\mathbf{k}|^2 + a^2)^{-\frac{d-1}{2}}$. Using that
$$\frac{1}{\lambda^p} = \frac{1}{\Gamma(p)} \int_0^\infty \mathrm{e}^{-\lambda t} t^{p-1} \mathrm{d}t,$$
we obtain
$$f(\mathbf{k}) := \frac{1}{(|\mathbf{k}|^2 + a^2)^{\frac{d-1}{2}}} = \frac{1}{\Gamma(\frac{d-1}{2})} \int_0^\infty \mathrm{e}^{-|\mathbf{k}|^2 t} \mathrm{e}^{-a^2 t} t^{\frac{d-1}{2}-1} \mathrm{d}t.$$
Using Fubini and the fact that the Fourier transform of the Gaussian is a Gaussian[1], we obtain that the Fourier transform of this function is
$$\widehat{f}(\mathbf{x}) = \frac{1}{\Gamma(\frac{d-1}{2})} \int_0^\infty \frac{1}{(2t)^{d/2}} \mathrm{e}^{-\frac{|\mathbf{x}|^2}{4t}} \mathrm{e}^{-a^2 t} t^{\frac{d-1}{2}-1} \mathrm{d}t = \frac{1}{\sqrt{2}^d \Gamma(\frac{d-1}{2})} \int_0^\infty \mathrm{e}^{-\frac{|\mathbf{x}|^2}{4t}} \mathrm{e}^{-a^2 t} t^{\frac{-3}{2}} \mathrm{d}t.$$
The last integral no longer depends on the dimension d. But for $d = 3$, we recover the usual Yukawa potential. This proves that $\widehat{f}(\mathbf{x})$ is $w_a(\mathbf{x})$, up to a multiplicative constant. We write $\widehat{w}_a(\mathbf{k}) = \kappa(d) f(\mathbf{k})$. To find the value of $\kappa(d)$, we evaluate at $\mathbf{k} = \mathbf{0}$. We find
$$\kappa(d) = \frac{\widehat{w}_a(\mathbf{0})}{f(\mathbf{0})} = a^{d-1} \widehat{w}_a(\mathbf{0}) = \frac{a^{d-1}}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} \frac{\mathrm{e}^{-a|\mathbf{x}|}}{|\mathbf{x}|} \mathrm{d}\mathbf{x} = \frac{|\mathbb{S}^{d-1}|}{(2\pi)^{d/2}} \int_0^\infty r^{d-2} \mathrm{e}^{-r} \mathrm{d}r = \frac{|\mathbb{S}^{d-1}|}{(2\pi)^{d/2}} \Gamma(d-1).$$
For $d = 3$, we recover the usual Yukawa prefactor $\kappa(3) = \sqrt{2}/\sqrt{\pi}$. □

We deduce the operator inequality
$$|\Delta + k_F^2| - w_a(\mathbf{x}) - a \leq |\Delta + k_F^2| - \frac{1}{|\mathbf{x}|} \leq |\Delta + k_F^2| - w_a(\mathbf{x}).$$
In particular, the lowest eigenvalue $-E$ of $|\Delta + k_F^2| - \frac{1}{|\mathbf{x}|}$ satisfies
$$-E_a - a \leq -E \leq -E_a, \quad \text{or equivalently} \quad E_a \leq E \leq E_a + a,$$
where $-E_a$ is the lowest eigenvalue of $|\Delta + k_F^2| - w_a(\mathbf{x})$. This one can be estimated as before. If $K_{a,E}$ is the corresponding Birman-Schwinger operator, we have
$$\|K_{a,E}\|_{\mathrm{op}} \leq \frac{1}{(2\pi)^{d/2}} \int_0^\infty \frac{r^{d-1}}{|r^2 - k_F^2| + E} \left[\int_{\mathbb{S}^{d-1}} \frac{\mathrm{d}\omega}{(r^2|\omega - \mathbf{e}_1|^2 + a^2)^{\frac{d-1}{2}}} \right] \mathrm{d}r.$$
We focus on the high density regime $k_F \gg 1$. For the term in bracket, we have, in the limit $a \to 0$ (we will choose $a \approx \mathrm{e}^{-\sqrt{k_F}}$ below, which goes to 0 as $k_F \to \infty$),
$$\int_{\mathbb{S}^{d-1}} \frac{\mathrm{d}\omega}{(r^2|\omega - \mathbf{e}_1|^2 + a^2)^{\frac{d-1}{2}}} = \frac{1}{r^{d-1}} \int_{\mathbb{S}^{d-1}} \frac{\mathrm{d}\omega}{(|\omega - \mathbf{e}_1|^2 + (a/r)^2)^{\frac{d-1}{2}}} \lesssim \frac{1}{r^{d-1}} \left| \log\left(\frac{r}{a}\right) \right|.$$

[1] We have $\mathcal{F}_d(G_a) = G_{1/a}$ where $G_a(\mathbf{x}) := a^{d/4} \mathrm{e}^{-\frac{a}{2}|\mathbf{x}|^2}$ is the d–dimensional Gaussian.

This gives, in the regime where $E \ll k_F^{-2}$ and $a \ll k_F^{-1}$

$$\begin{aligned}\|K_{a,E}\|_{\mathrm{op}} &\lesssim \int_0^\infty \frac{1}{|r^2-k_F^2|+E}\left|\log\left(\frac{r}{a}\right)\right|\mathrm{d}r = \frac{1}{k_F}\int_0^\infty \frac{1}{|u^2-1|+(E/k_F^2)}\left|\log\left(\frac{k_F u}{a}\right)\right|\mathrm{d}u \\ &\lesssim \frac{C}{k_F}\left|\log\left(\frac{k_F^2}{E}\right)\right|\cdot\left|\log\left(\frac{k_F}{a}\right)\right| \approx \frac{C}{k_F}\left|\log\left(E\right)\right|\cdot\left|\log\left(a\right)\right|.\end{aligned}$$

Making the specific choice $a = \mathrm{e}^{-\sqrt{k_F}}$ gives, for this a, $\|K_{a,E}\|_{\mathrm{op}} \leq \frac{C}{\sqrt{k_F}}|\log(E)|$. Reasoning as before, we obtain that there is $C \geq 0$ and $\alpha > 0$ so that

$$E_a \leq C\mathrm{e}^{-\alpha\sqrt{k_F}}, \quad \text{hence} \quad C\mathrm{e}^{-\alpha\sqrt{k_F}} \leq E \leq C\mathrm{e}^{-\alpha\sqrt{k_F}} + \mathrm{e}^{-\sqrt{k_F}},$$

and the result follows.

3.4 Phase diagram in the Peierls model for polyacetylene

In this section, we describe the results in [G10]. This is joint work with Éric SÉRÉ and Adéchola KOUANDÉ.

In the last section of this chapter, we would like to describe some results on the SSH/Peierls model for polyacetylene [Pei96; Frö54]. Although this model does not describe the Fermi gas, it has features which are similar to the Hartree–Fock gas. It is a model with two parameters μ and T, where μ is related to the rigidity of the chain, and T is the temperature. Although the model is 1–periodic, the minimizer is always 2–periodic at $T = 0$ (Peierls dimerization, see Figure 3.3), but becomes 1–periodic as we increase the temperature.

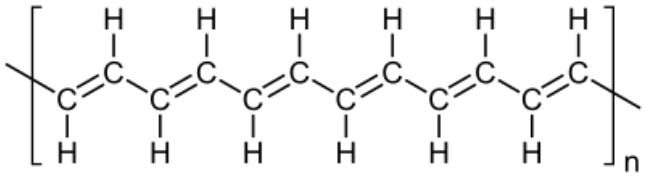


Figure 3.3: The polyacetylene is a periodic chain of carbon atom, but the covalent bonds arrange themselves in a 2–periodic fashion.

We consider a linear periodic even chain of $L = 2N$ atoms (or rather a $C-H$ motif for polyacetylene), and denote by d_i the distance between the i-th and $(i+1)$-th atom, with the convention that indices are taken modulo L. We add electrons to this system, encoded by a one-body density matrix $0 \leq \gamma \leq 1$, and each electron feels a tight–binding Hamiltonian $H_{\mathbf{t}}$ acting on $\ell_2(\mathbb{C}^L)$, where

$$H_{\mathbf{t}} := \begin{pmatrix} 0 & t_1 & 0 & 0 & \cdots & t_L \\ t_1 & 0 & t_2 & \cdots & 0 & 0 \\ 0 & t_2 & 0 & t_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & t_{L-2} & 0 & t_{L-1} \\ t_L & 0 & \cdots & 0 & t_{L-1} & 0 \end{pmatrix}. \tag{3.11}$$

It is a nearest neighbour model in which the hopping parameters t_i depend on the distances between the atoms. For an atomic configuration $\mathbf{t} = (t_1, \cdots, t_L)$ and electronic configuration γ, the Peierls energy is

$$\mathcal{E}^L(\mathbf{t}, \gamma) := \frac{1}{2}g\sum_{i=1}^{L}(d_i - d_\sharp)^2 + 2\mathrm{Tr}\left(H_{\mathbf{t}}\gamma\right) + 2T\,\mathrm{Tr}\left(S(\gamma)\right).$$

The first term is the distortion energy of the atoms. We model the atom/atom interaction as an effective spring with strength g and distance at rest $d_\sharp$. The second term is the electronic energy. The 2 factor stands for the spin of the electrons. The last term is the entropy of the electrons (we neglect the entropy of the atoms). Here, $T > 0$ is the temperature, and $S(x) = x\log(x) + (1-x)\log(1-x)$ is the already introduced fermionic entropy.

In the simple and original SSH/Peierls model, it is assumed that the hopping amplitudes t_i are linear functions of the distances d_i, of the form $t_i - t_\sharp = -\alpha(d_i - d_\sharp)$, where $t_\sharp$ is the hopping energy at rest. The minus sign in front of $\alpha > 0$ states that electrons needs more energy to jump to an atom which is further. After the normalization $\mathcal{E} \to \frac{\mathcal{E}}{t^\sharp}$, $t_i \to \frac{t_i}{t^\sharp}$, we end up with an energy of the form

$$\mathcal{E}^L_{\mu,T}(\mathbf{t},\gamma) = \frac{1}{2}\mu\sum_{i=1}^{L}(t_i-1)^2 + 2\mathrm{Tr}\left(H_{\mathbf{t}}\gamma\right) + 2T\,\mathrm{Tr}\left(S(\gamma)\right).$$

There are only two parameters in the model, namely the temperature $T \geq 0$, and the parameter $\mu := t_\sharp g\alpha^{-2}$, which we interpret as a normalized rigidity of the chain. Our goal is to compute

$$E^L(\mu,T) := \inf\left\{\mathcal{E}^L_{\mu,T}(\mathbf{t},\gamma),\ \mathbf{t}\in\mathbb{R}^L_+,\ \gamma\in\mathcal{S}\left(\ell_2(\mathbb{C}^L)\right),\ 0\leq\gamma\leq 1\right\},$$

Upon rescaling, we may assume $b = 1$. We are left with two parameters in the model, namely $\mu > 0$ and $T > 0$. Our goal is to compute

$$E^L(\mu,T) := \inf\left\{\mathcal{E}^L_{\mu,T}(\mathbf{t},\gamma),\ \mathbf{t}\in\mathbb{R}^L_+,\ \gamma\in\mathcal{S}\left(\ell_2(\mathbb{C}^L)\right),\ 0\leq\gamma\leq 1\right\},$$

and to describe the corresponding minimizers. One can perform the minimization in γ first. We get

$$\min_{0\leq\gamma\leq1} 2\left\{\mathrm{Tr}(H\gamma) + T\,\mathrm{Tr}(S(\gamma))\right\} = -\mathrm{Tr}\left(h_T(H^2)\right), \quad \text{with} \quad h_T(x) := \begin{cases} 2T\ln\left(2\cosh\left(\frac{\sqrt{x}}{2T}\right)\right) & \text{if } T>0;\\ \sqrt{x} & \text{if } T=0.\end{cases}$$

We skip the proof of this equality, and note that the minimizer is given by the Fermi–Dirac expression $\gamma_* = (1+\mathrm{e}^{H/T})^{-1}$. This gives the reduced energy, depending solely on $\mathbf{t}\in\mathbb{R}^L_+$, and defined by

$$\mathcal{F}^L(\mathbf{t}) := \frac{1}{2}\mu\sum_{i=1}^{L}(t_i-1)^2 - \mathrm{Tr}\left(h_T(H^2_{\mathbf{t}})\right).$$

It is a classical result by Kennedy and Lieb [KL04] that, in the case where $L = 2N$ is even, the minimizers of $\mathcal{F}^L$ are 2–periodic, of the form

$$t_i = W + (-1)^i\delta \qquad \text{or} \qquad t_i = W - (-1)^i\delta, \qquad \text{with} \quad W>0,\ \delta\geq 0.$$

The quantity $W > 0$ is the average distance between 2 atoms, in the presence of the electrons (one can show that $W < 1$: the atoms are attracted by the presence of the electrons), and $\delta \geq 0$ is the *distorsion*. Actually, Kennedy and Lieb only considered the case $T = 0$, although their proof applies similarly to the case of positive temperature $T \geq 0$, as it only relies on the concavity of the h_T function.

When $\delta = 0$, the solution is 1-periodic, and when $\delta > 0$, it is 2-periodic. In the latter case, we say that the system is dimerized (Peierls distorsion). The Kennedy–Lieb result allows to reduce the problem as a minimization over 2 variables only, namely (W,δ). The corresponding energy is found to be

$$\mathcal{F}^L(W,\delta) = L\frac{\mu}{2}\left[(W-1)^2+\delta^2\right] - \sum_{k=1}^{L} h_T\left(4W^2\cos^2\left(\frac{2k\pi}{L}\right) + 4\delta^2\sin^2\left(\frac{2k\pi}{L}\right)\right).$$

Dividing by L and taking the limit $L = 2N \to \infty$ gives the thermodynamic energy per unit atom, defined by (we use the notation $⨍_\Omega$ for $|\Omega|^{-1}\int_\Omega$)

$$\boxed{g_{\mu,T}(W,\delta) := \frac{\mu}{2}\left[(W-1)^2+\delta^2\right] - ⨍_0^{2\pi} h_T\left(4W^2\cos^2(s)+4\delta^2\sin^2(s)\right)\mathrm{d}s.}$$

We define the minimum energy for all dimerized configurations, and the one for all 1–periodic configurations, respectively by

$$E^{(2)}(\mu,T) := \inf\left\{g_{\mu,T}(W,\delta),\ W\geq 0, \delta\geq 0\right\}, \qquad \text{and} \qquad E^{(1)}(\mu,T) := \inf\left\{g_{\mu,T}(W,0),\ W\geq 0\right\}.$$

We have $E^{(2)}(\mu,T) \leq E^{(1)}(\mu,T)$, with equality if the optimal dimerized configuration is, actually, 1–periodic.

Theorem 3.8 ([G10] Theorems 1.4 and A.1)**.**

- *At $T = 0$, for all $\mu_* > 0$, there is $C \geq 0$ so that*

$$\forall \mu \geq \mu_*, \qquad E^{(1)}(\mu,0) - C\mathrm{e}^{-\frac{\pi}{2}\mu} \leq E^{(2)}(\mu,0) < E^{(1)}(\mu,0). \tag{3.12}$$

- *There is $T_c = T_c(\mu)$ so that, for all $T < T_c$, the optimal configurations are dimerized, while for $T \geq T_c$, the optimal configuration is 1–periodic. In addition, we have, as $\mu \to \infty$,*

$$T_c(\mu) \sim C\mathrm{e}^{-\frac{\pi}{4}\mu+o(1)}, \qquad \text{with} \qquad C \approx 1.6686.$$

Comparing this result with Theorem 3.6, we see that the Peierls model exhibits features similar to the Hartree–Fock model for the electrons gas. At $T = 0$, the minimizer never shares the periodicity of the model, but the gain of energy due the breaking of this symmetry is exponentially small (in ρ or μ). In addition, when increasing the temperature by an exponentially small amount, the minimizer becomes periodic.

Proof. We only give some ideas of the proof. Let us drop the T parameter, and call h a general concave function. We denote by $W_0 = W_0(\mu)$ the optimizer for the $E^{(1)}$ problem. This one solves the Euler–Lagrange equation

$$\mu(W_0-1) - ⨍_0^{2\pi} h'(4W_0^2\cos^2(s))\left[8W_0\cos^2(s)\right]\mathrm{d}s = 0. \tag{3.13}$$

Now, for (W,δ) any other trial state, we write $W = W_0+\varepsilon$, and get

$$\begin{aligned} g(W_0+\varepsilon,\delta) - g(W_0,0) = \frac{\mu}{2}\left[\varepsilon^2+\delta^2\right] + \mu(W_0-1)\varepsilon \\ - ⨍_0^{2\pi}\left[h\left(4(W_0+\varepsilon)^2\cos^2(s)+4\delta^2\sin^2(s)\right) - h(4W_0^2\cos^2(s))\right]\mathrm{d}s. \end{aligned}$$

Since h is concave, we have $h(a) - h(b) \leq h'(b)(a-b)$. With $b = 4W_0^2\cos^2(s)$ and $a = 4(W_0+\varepsilon)^2\cos^2(s)+4\delta^2\sin^2(s)$, and noting that the term linear in ε cancels due the previous Euler–Lagrange equations, we get

$$\begin{aligned} g(W_0+\varepsilon,\delta) - g(W_0,0) &\geq \frac{\mu}{2}\left[\varepsilon^2+\delta^2\right] - 4⨍_0^{2\pi} h'(4W_0^2\cos^2(s))\left[\varepsilon^2\cos^2(s)+\delta^2\sin^2(s)\right]\mathrm{d}s \\ &= \varepsilon^2\left(\frac{\mu}{2} - 4⨍_0^{2\pi} h'(4W_0^2\cos^2(s))\cos^2(s)\mathrm{d}s\right) + \delta^2\left(\frac{\mu}{2} - 4⨍_0^{2\pi} h'(4W_0^2\cos^2(s))\sin^2(s)\mathrm{d}s\right). \end{aligned}$$

Under the condition

$$\mu \geq \max\left\{ 8 \fint_0^{2\pi} h'(4W_0^2\cos^2(s))\cos^2(s)\mathrm{d}s \; ; \; 8 \fint_0^{2\pi} h'(4W_0^2\cos^2(s))\sin^2(s)\mathrm{d}s \right\}, \tag{3.14}$$

the right-hand side is always positive, which proves that the minimum of g is attained for $\varepsilon = 0$ and $\delta = 0$. In the case $T > 0$, both integrals appearing in (3.14) converge. This already proves that for all $T > 0$ and all μ larger than some $\mu(T)$, the minimizer is 1-periodic. Detailed computations using Euler–Lagrange equations, and that we do not detail here, allow to estimate the critical temperature $T_c = T_c(\mu)$.

In the case $T = 0$ and $h(x) = \sqrt{x}$, the derivative $h'(x) = \frac{1}{2\sqrt{x}}$ has a singularity at $x = 0$, and the function $h'(4W_0^2\cos^2)\sin^2$ appearing in (3.14) is no longer integrable, due to the singularities at $s = \pi/2$ and $s = 3\pi/2$, so the previous lower bound is meaningless. In this case however, one can perform exact computations. First, we compute W_0 in (3.13). We find that

$$\fint_0^{2\pi} h'(4W_0^2\cos^2(s))\left[8W_0\cos^2(s)\right]\mathrm{d}s = \fint_0^{2\pi} 2|\cos(s)|\mathrm{d}s = \frac{4}{\pi},$$

hence $W_0(\mu) = 1 + \frac{4}{\pi\mu}$. On the other hand, we have, with E denoting the complete elliptic integral of the second kind,

$$\fint_0^{2\pi} \left[\sqrt{4(W_0+\varepsilon)^2\cos^2(s) + 4\delta^2\sin^2(s)} - \sqrt{4W_0^2\cos^2(s)}\right]\mathrm{d}s = \frac{4(W_0+\varepsilon)}{\pi} E\left(1 - \frac{\delta^2}{(W_0+\varepsilon)^2}\right) - \frac{4W_0}{\pi}.$$

This gives the exact energy difference

$$g(W_0+\varepsilon,\delta) - g(W_0,0) = \frac{\mu}{2}\left[\varepsilon^2+\delta^2\right] - \frac{4}{\pi}(W_0+\varepsilon)\left[E\left(1-\frac{\delta^2}{(W_0+\varepsilon)^2}\right) - 1\right]$$

It remains to compute the minimum of the right-hand side, as a function of ε and δ. We denote by $\varepsilon_* = \varepsilon_*(\mu)$ and $\delta_* = \delta_*(\mu)$ the minimiser. First, using the bound $E(1-a) - 1 \leq \sqrt{a}$, we get

$$0 \geq g(W_0+\varepsilon_*,\delta_*) - g(W_0,0) \geq \frac{\mu}{2}\left[\varepsilon_*^2+\delta_*^2\right] - \frac{4}{\pi}\delta_*.$$

We deduce that $|\varepsilon_*| \leq \frac{4}{\pi\mu}$ and $\delta_* < \frac{8}{\pi\mu}$. In particular, $W_0 + \varepsilon_* \geq 1$. Together with the fact that the map $x \mapsto x\left[E(1-\delta x^{-2}) - 1\right]$ is decreasing, we obtain the lower bound

$$g(W_0+\varepsilon_*,\delta_*) - g(W_0,0) \geq \frac{\mu}{2}\left[\varepsilon_*^2+\delta^2\right] - \frac{4}{\pi}\left[E\left(1-\delta_*^2\right) - 1\right] \geq \frac{\mu}{2}\delta_*^2 - \frac{4}{\pi}\left[E\left(1-\delta_*^2\right) - 1\right].$$

Using that $\mu \mapsto \delta_*(\mu)$ goes to 0 as $\mu \to \infty$ and that $E(1-\delta^2) - 1 \sim \delta^2|\log(\delta)|/2$ as $\delta \to 0$, the left-hand side is equivalent to

$$\delta_*^2\left(\frac{\mu}{2} - \frac{2}{\pi}|\log(\delta_*)|\right).$$

The minimum of this last expression is obtained for $\delta = \mathrm{e}^{-\frac{\pi}{4}\mu+\frac{1}{2}}$, giving the lower bound (3.12). □

3.5 Perspectives

Concerning the Peierls model, together with Éric SÉRÉ and Adéchola KOUANDÉ, we are studying *kink* modes, which are general solutions to the Euler–Lagrange equations. We prove that they are exponentially close to dimerized solution on the right and on the left (not necessarily the same dimerized configuration), and that they exist an infinity of such modes, up to translations.

The phase diagram for polyacetylene can be studied in other contexts. It was proved in [FL11] that *graphene* could have lattice distortions, but that these distortions must have a periodic structure. We would like to apply the method developed in [G10] to graphene, and prove that (a) at $T = 0$, the optimal configuration is not the 1–periodic one (usually considered in the literature), and (b) that it becomes 1–periodic when increasing the temperature. We started to look at this problem with Éric Séré and Thaddeus Roussigné.

CHAPTER 4

LIEB-THIRRING (AND RELATED) INEQUALITIES

4.1 Introduction

Lieb-Thirring inequalities play an important role in the study of large fermionic quantum systems. They were introduced by Lieb and Thirring in [LT75; LT76] to provide a short proof of the stability of matter. They read as follows. Let κ (usually γ in the literature, but γ is our one-body reduced density matrix...) be so that

$$\kappa \begin{cases} \geq \frac{1}{2} & \text{for} \quad d = 1, \\ > 0 & \text{for} \quad d = 2, \\ \geq 0 & \text{for} \quad d \geq 3. \end{cases} \tag{4.1}$$

Then, there exists an optimal (smallest) constant $L_{\kappa,d}$ so that for all positive valued potential $V \in L^{\kappa+\frac{d}{2}}(\mathbb{R}^d, \mathbb{R}_+)$, we have

$$\boxed{\sum_{j=1}^{\infty} |\lambda_j(-\Delta - V)|^{\kappa} \leq L_{\kappa,d} \int_{\mathbb{R}^d} V^{\kappa+\frac{d}{2}},} \tag{4.2}$$

where $\lambda_j(-\Delta - V)$ denotes the j-th lowest negative eigenvalue of $(-\Delta - V)$ if exists, and 0 if $-\Delta - V$ has less that j strictly negative eigenvalues. The importance of this inequality comes from the fact that the right-hand side is **extensive** in the following sense. Assume that V is the sum of two potentials V_L and V_R, with V_L located on the far left, and V_R located on the far right. Then, we can expect that the spectrum of $-\Delta - V_L - V_R$ is close to

$$\sigma(-\Delta - V_L - V_R) \approx \sigma(-\Delta - V_L) \bigcup \sigma(-\Delta - V_R).$$

The Lieb-Thirring inequality is one way to quantify this intuitive picture, in the form of an inequality.

In the original articles [LT75; LT76], this inequality was proved in the non-critical case $\kappa > \max\{0, 1 - \frac{d}{2}\}$. The equality case $\kappa = \frac{1}{2}$ in $d = 1$ was proved by Weidl [Wei96], and the case $\kappa = 0$ for $d \geq 3$ was proved independently by Cwikel, Lieb and Rozenblum in [Cwi77; Lie76; Roz72], and is known as the CLR inequality. In this case, the left-hand side simply counts the number of negative eigenvalues.

In this chapter, we study the optimal (smallest) constants $L_{\kappa,d}$ for the original Lieb-Thirring inequality, and for another one that we called the *fermionic non-linear Schrödinger inequality*. In what follows, we restrict our attention to **positive** potentials V, and write our Schrödinger operator with a minus sign: $H_V := -\Delta - V$. Recall that if V is not positive, we can write $V = V_+ - V_-$, and we have the operator inequality $-\Delta - V \geq -\Delta - V_+$, hence $|\lambda_j(-\Delta - V)| \leq |\lambda_j(-\Delta - V_+)|$. In

particular, (4.2) implies that

$$\sum_{j=1}^{\infty} |\lambda_j(-\Delta - V)|^\kappa \leq \sum_{j=1}^{\infty} |\lambda_j(-\Delta - V_+)|^\kappa \leq L_{\kappa,d} \int_{\mathbb{R}^d} V_+^{\kappa+\frac{d}{2}}.$$

4.2 Lieb-Thirring inequality

In this section, we describe results in [G12; G13; G14], obtained in collaboration with Mathieu LEWIN and Rupert FRANK.

4.2.1 The finite rank Lieb-Thirring inequality

Several conjectures concern the value of the optimal (smallest) constants $L_{\kappa,d}$. We will review what is currently known and conjectured in Section 4.2.3, and refer to the recent book [FLW22] for more details. In [G12; G14], we asked whether this optimal value could be attained for a potential V generating only a finite number of negative eigenvalues. For a given $N \in \mathbb{N}$, we study the optimal (smallest) constant $L_{\kappa,d}^{(N)}$ in the *finite rank Lieb–Thirring inequality*

$$\boxed{V \in L^{\kappa+\frac{d}{2}}(\mathbb{R}^d, \mathbb{R}_+), \quad \sum_{j=1}^{N} |\lambda_j(-\Delta - V)|^\kappa \leq L_{\kappa,d}^{(N)} \int_{\mathbb{R}^d} V^{\kappa+\frac{d}{2}}.} \tag{4.3}$$

In the case $N = 1$, this is sometimes called the Keller problem [Kel61]. It is not difficult to see that

$$L_{\kappa,d}^{(1)} \leq L_{\kappa,d}^{(2)} \leq \cdots \qquad \text{and that} \qquad L_{\kappa,d} = \lim_{N\to\infty} L_{\kappa,d}^{(N)} = \sup_{N\in\mathbb{N}} L_{\kappa,d}^{(N)}. \tag{4.4}$$

Our main result in [G12; G14] is the following. We state it for the non-critical case $\kappa > \max\{0, 1 - \frac{d}{2}\}$ (that is when the inequality is strict in (4.1)).

Theorem 4.1 ([G12] Theorem 1, and [G14] Theorem 1)**.** *For all $\kappa > \max\{0, 1 - \frac{d}{2}\}$ and all $N \in \mathbb{N}$, there is an optimal potential $V \in L^{\kappa+\frac{d}{2}}(\mathbb{R}^d, \mathbb{R}_+)$ for (4.3). In addition, under the extra condition $\kappa > \max\{0, 2 - \frac{d}{2}\}$, we have*

$$L_{\kappa,d}^{(2N)} > L_{\kappa,d}^{(N)}. \tag{4.5}$$

In particular, the optimal Lieb-Thirring constant $L_{\kappa,d}$ satisfies $L_{\kappa,d} > L_{\kappa,d}^{(N)}$ for all $N \in \mathbb{N}$.

In the case $\kappa > \max\{0, 2 - \frac{d}{2}\}$, this implies that if the problem defining $L_{\kappa,d}$ has an optimizer $V \in L^{\kappa+\frac{d}{2}}(\mathbb{R}^d, \mathbb{R}_+)$ (this is still an open problem, and we conjecture that there is none, see Section 4.2.3 below), then this potential must generate an infinity of negative eigenvalues.

Proof. We only give few elements of the proof, as it is quite lengthy. First, we can prove the existence of minimizers for the $L_{\kappa,d}^{(N)}$ problem using the classical «bubble decomposition» or «profile decomposition» in nonlinear analysis. Using the Euler-Lagrange equation for an minimizer V_N shows that it is of the form

$$V_N = C \left(\sum_{j=1}^{N} |\lambda_j|^{\kappa-1} |u_j|^2 \right)^{\frac{1}{\kappa+\frac{d}{2}-1}},$$

where $(\lambda_j, u_j)_{1\leq j\leq N}$ are the first eigenpairs of the operator $-\Delta - V_N$, that is

$$(-\Delta - V_N)\, u_j = \lambda_j u_j, \qquad \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_j \leq 0.$$

For simplicity, let us assume that $\lambda_N < 0$. Otherwise, the potential V_N only generates $M < N$ strictly negative eigenvalues, and is also an optimizer for $L^{(M)}_{\kappa,d}$. Since V_N goes to 0 at infinity, we obtain the decay at infinity

$$u_j(\mathbf{x}) \sim C\left(\frac{\mathrm{e}^{-\sqrt{|\lambda_j|}|\mathbf{x}|}}{1+|\mathbf{x}|^{d-1}}\right), \quad \text{hence} \quad V_N(\mathbf{x}) \sim C\left(\frac{\mathrm{e}^{-\sqrt{|\lambda_N|}|\mathbf{x}|}}{1+|\mathbf{x}|^{d-1}}\right)^{\frac{2}{\kappa+\frac{d}{2}-1}}. \tag{4.6}$$

The main idea to prove the inequality (4.5) is to construct a trial potential V_R for the $L^{(2N)}_{\kappa,d}$ problem by putting two copies of V_N far apart, separated by a distance $R > 0$. More specifically, we set

$$V_R(\mathbf{x}) = \left(V_N^{\frac{1}{p-1}}(\mathbf{x}-\tfrac{R}{2}) + V_N^{\frac{1}{p-1}}(\mathbf{x}+\tfrac{R}{2})\right)^{p-1}, \qquad \text{with} \quad p := \frac{\kappa+\frac{d}{2}}{\kappa+\frac{d}{2}-1}. \tag{4.7}$$

The quantity $\rho := V^{\frac{1}{p-1}} = V^{\kappa+\frac{d}{2}-1} = C'\sum_{j=1}^N |\lambda_j|^{\kappa-1}|u_j|^2$ can be interpreted as an electronic density. Note that p is dual exponent of $\kappa+\frac{d}{2} > 0$ (the density ρ is the dual variable of the potential V). Equation (4.7) states that the trial potential V_R is constructed by adding the densities instead of the potentials. Detailed computations, that we will not display here, then show that

$$L^{(2N)}_{\kappa,d} \geq \frac{\sum_{j=1}^{2N}|\lambda_j\left(-\Delta - V_R\right)|^\kappa}{\int_{\mathbb{R}^d} V_R^{\kappa+\frac{d}{2}}} \geq L^{(N)}_{\kappa,d}\left(1+\frac{1}{2}\left(\kappa+\frac{d}{2}-1\right)A_R\right) + O(A_R^2+e_R^2), \tag{4.8}$$

where

$$A_R := \int_{\mathbb{R}^d}\left(V_N^{\frac{1}{p-1}}(\cdot-\tfrac{R}{2}) + V_N^{\frac{1}{p-1}}(\cdot+\tfrac{R}{2})\right)^{p-1} - V_N(\cdot-\tfrac{R}{2}) - V_N(\cdot+\tfrac{R}{2}).$$

and

$$e_R \leq C\left(R^{\frac{3-d}{2}} + R^{3-d}\right)\exp\left(-2\sqrt{|\lambda_N|}R\right).$$

The error term e_R comes from the orthonormalization procedure that we use to evaluate $|\lambda_j(-\Delta - V_R)|$. To obtain a lower bound for A_R, we evaluate the integrand around $\mathbf{x} = \mathbf{0}$ using the explicit decay of V_N in (4.6) and find that

$$A_R \geq \frac{C}{R^{2p(d-1)}}\exp\left(-p\sqrt{|\lambda_N|}R\right).$$

Inequality (4.8) is interesting only in the case where $e_R^2 = o(A_R)$. In view of the exponents appearing in the exponentials, this happens when $p < 2$. Since p is the dual exponent of $\kappa+\frac{d}{2}$, this is also $\kappa > 2 - \frac{d}{2}$, which is our condition in the Theorem. In this case, we gain a positive exponentially small amount as $R \to \infty$ in (4.8), which implies $L^{(2N)}_{\kappa,d} > L^{(N)}_{\kappa,d}$. □

Interestingly, we do not know whether the strict inequality $L^{(N+1)}_{\kappa,d} > L^{(N)}_{\kappa,d}$ holds. We only know that it is the case for infinitely many N, see below.

4.2.2 The periodic Lieb-Thirring inequality

Theorem 4.1 shows that one can always improve the Lieb-Thirring constant by adding eigenvalues (or particles). Actually, numerical simulations done in other contexts (see Figs. 4.4 and 4.5 below) suggest that sequences of optimizers $(V_N)_{N\in\mathbb{N}}$ converge, up to translation and rotation, to a periodic configuration V_{per}. This limit potential is not a trial state for the original Lieb-Thirring inequality, since it does not belong to any $L^{\kappa+\frac{d}{2}}(\mathbb{R}^d)$ space.

In [G13], we introduce a periodic version of the Lieb–Thirring inequality.

Lemma 4.2 ([G13] Theorem 1.1)**.** *Let $\mathcal{R}$ be any lattice in $\mathbb{R}^d$. For any $\kappa > \max\{0, 1 - \frac{d}{2}\}$ and any positive potential $V \in L^{\kappa+\frac{d}{2}}_{\mathrm{loc}}(\mathbb{R}^d)$ which is $\mathcal{R}$-periodic potential, we have*

$$\underline{\mathrm{Tr}}\,(-\Delta - V)_-^\kappa \leq L_{\kappa,d} \fint V^{\kappa+\frac{d}{2}},$$

with the same optimal Lieb-Thirring constant as in (4.2). Here $\underline{\mathrm{Tr}}$ denotes the trace per unit volume, and $\fint$ the integral per unit volume, respectively defined by

$$\underline{\mathrm{Tr}}\,(-\Delta - V)_-^\kappa := \lim_{n\to\infty} \frac{1}{|\Omega_n|} \mathrm{Tr}\,(-\Delta - V\mathbb{1}_{\Omega_n})_-^\kappa \quad \text{and} \quad \fint f := \frac{1}{|\Omega_n|}\int_{\mathbb{R}^d} f\mathbb{1}_{\Omega_n},$$

with $\Omega_n := n\Omega$ and $\Omega := \mathbb{R}^d/\mathcal{R}$.

We refer to [G13] for the proof. An important trial state for this inequality is the constant potential $V = \mu > 0$. This potential is $\mathcal{R}$–periodic for any lattice $\mathcal{R}$. This gives the lower bound

$$\begin{aligned} L_{\kappa,d} &\geq \frac{\underline{\mathrm{Tr}}\,(-\Delta - \mu)_-^\kappa}{\fint \mu^{\kappa+\frac{d}{2}}} = \frac{1}{\mu^{\kappa+\frac{d}{2}}}\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left(k^2 - \mu\right)_-^\kappa \mathrm{d}k \\ &= \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left(k^2 - 1\right)_-^\kappa \mathrm{d}k = \frac{\Gamma(\kappa+1)}{2^d\pi^{d/2}\Gamma(\kappa+\frac{d}{2}+1)} =: L^{\mathrm{sc}}_{\kappa,d}. \end{aligned} \tag{4.9}$$

The constant $L^{\mathrm{sc}}_{\kappa,d}$ is known as the semi-classical Lieb–Thirring constant. In a regime where $L_{\kappa,d} = L^{\mathrm{sc}}_{\kappa,d}$, we say that (one of) the optimal scenario is the *semi-classical one*. Note that the operator $-\Delta - \mu$ describes a **fluid phase**, invariant by translation. The operator $\gamma := \mathbb{1}(-\Delta < \mu)$ is the free Fermi-gas at null temperature described in Section 3.2.2.

4.2.3 What is known and conjectured about Lieb–Thirring «best scenarii»

Let us review what is known and conjectured about the «best» potential V in the Lieb-Thirring inequality, depending on the parameter κ and the dimension d.

- (Lower bound) From (4.4) and (4.9), we obtain the inequality

$$L_{\kappa,d} \geq \max\left\{L^{(1)}_{\kappa,d}, L^{\mathrm{sc}}_{\kappa,d}\right\}. \tag{4.10}$$

  A famous conjecture by Lieb and Thirring in the original article [LT76] states that there is always equality in (4.10). This conjecture is now known to be incorrect in some cases, and correct in other cases, as we explain below.

- (monotonicity) The maps

$$\forall N \in \mathbb{N}, \qquad \kappa \mapsto \frac{L^{(N)}_{\kappa,d}}{L^{\mathrm{sc}}_{\kappa,d}} \qquad \text{and} \qquad \kappa \mapsto \frac{L_{\kappa,d}}{L^{\mathrm{sc}}_{\kappa,d}} \qquad \text{are non–increasing.}$$

  This was proved in [AL78] for the ratio $L_{\kappa,d}/L^{\mathrm{sc}}_{\kappa,d}$, but the proof works similarly for all $N \in \mathbb{N}$. In [G12], we proved that when $N = 1$, the map $\kappa \mapsto L^{(1)}_{\kappa,d}/L^{\mathrm{sc}}_{\kappa,d}$ is strictly decreasing. By (4.10), the map $\kappa \mapsto L_{\kappa,d}/L^{\mathrm{sc}}_{\kappa,d}$ is lower bounded by 1, so it stays equal to 1 whenever it touches 1.

At this point, it is interesting to introduce the crossing points

$$\forall N \in \mathbb{N}, \quad \kappa_{N\cap\mathrm{sc}}(d) := \inf\left\{\kappa > \max\{0, 1 - \tfrac{d}{2}\}, \quad L^{(N)}_{\kappa,d} \leq L^{\mathrm{sc}}_{\kappa,d}\right\}.$$

d	1	2	3	4	5	6	7	$d \geq 8$
$\kappa_{1\cap\mathrm{sc}}(d)$	$= \frac{3}{2}$	1.1654	0.8627	0.5973	0.3740	0.1970	0.0683	0 (no crossing)

Figure 4.1: Numerical computation of $\kappa_{1\cap\mathrm{sc}}(d)$.

It corresponds to the minimum value of κ where the semi-classical scenario becomes better than the N-particle one. Since $L^{(N+1)}_{\kappa,d} \geq L^{(N)}_{\kappa,d}$, we have $\kappa_{(N+1)\cap\mathrm{sc}} \geq \kappa_{N\cap\mathrm{sc}}$.The constants $L^{(1)}_{\kappa,d}$ and $L^{\mathrm{sc}}_{\kappa,d}$ can be evaluated numerically with high precision, so we can compute the values $\kappa_{1\cap\mathrm{sc}}$ (see Table 4.1). The fact that $\kappa_{1\cap\mathrm{sc}}(1) = 3/2$ will be shown later.

We also introduce the semi-classical crossing point

$$\kappa_{\mathrm{sc}}(d) = \lim_{N\to\infty} \kappa_{N\cap\mathrm{sc}}(d) = \sup_N \kappa_{N\cap\mathrm{sc}}(d) = \inf\left\{\kappa > \max\{0, 1-\tfrac{d}{2}\}, \quad L_{\kappa,d} = L^{\mathrm{sc}}_{\kappa,d}\right\}.$$

For $\kappa \geq \kappa_{\mathrm{sc}}(d)$, we have $L_{\kappa,d} = L^{\mathrm{sc}}_{\kappa,d}$, and (one of) the best scenario is the semi-classical one.

- ($\kappa = \frac{3}{2}$ in dimension $d = 1$). We have, for all $N \in \mathbb{N}$,

$$\kappa_{1\cap\mathrm{sc}}(1) = \kappa_{N\cap\mathrm{sc}}(1) = \kappa_{\mathrm{sc}}(1) = \frac{3}{2}, \quad \text{and, in particular} \quad L^{(1)}_{\frac{3}{2},1} = L^{(N)}_{\frac{3}{2},1} = L^{\mathrm{sc}}_{\frac{3}{2},1} = L_{\frac{3}{2},1}.$$

  As already noticed in [LT76], this particular case is linked to the integrable Korteweg-de-Vries (KdV) equation. Actually, any N-solitons for KdV is a minimizer for $L^{(N)}_{\kappa,d}$. In [G13, Theorem 2.1], we also proved that, for all $k > 0$, the periodic Lamé potentials V_k are all optimizers for the periodic Lieb-Thirring equation, where

$$V_k(\mathbf{x}) := 1 + k^2 - 2k^2 \mathrm{sn}\,(\mathbf{x}|k)^2\,, \qquad \text{with minimal period} \qquad 2K(k) > 0.$$

  Here $\mathrm{sn}\,(\cdot|k)$ is a Jacobi elliptic function and $K(k)$ is the complete elliptic integral of the first kind. Actually, $V_k(\cdot) \to 2\cosh^{-2}(\cdot)$ (the 1-soliton) as $k \to \infty$, and $V_k(\cdot) \to \mu := 1$ as $k \to 0$, so the family $k \mapsto V_k$ somehow interpolates between the $N = 1$ case and the semi-classical case.

- ($\kappa \geq \frac{3}{2}$). For all $\kappa \geq \frac{3}{2}$, and in all dimension $d \geq 1$, we have the semi-classical regime

$$\forall d \geq 1,\ \forall \kappa \geq \frac{3}{2}, \qquad L_{\kappa,d} = L^{\mathrm{sc}}_{\kappa,d}, \quad \text{or equivalently} \quad \kappa_{\mathrm{sc}}(d) \leq \frac{3}{2}.$$

  The equality $L_{\frac{3}{2},d} = L^{\mathrm{sc}}_{\frac{3}{2},d}$ for all d was proved in [LW00] (see also [BL99]). The case $\kappa > \frac{3}{2}$ comes from the monotonicity of the ratio $L_{\kappa,d}/L^{\mathrm{sc}}_{\kappa,d}$.

- ($\kappa = 1/2$ in dimension $d = 1$). We have

$$L_{1/2,1} = L^{(1)}_{1/2,1},$$

  and the corresponding minimization problems have no minimizer. In some sense, the minimizer is the Dirac distribution $V = \delta_0$, see [HLT98].

- ($\kappa < 1$ is not semi-classical). For all $d \geq 1$ and all $\kappa < 1$, we have

$$L_{\kappa,d} > L^{\mathrm{sc}}_{\kappa,d}, \quad \text{or equivalently} \quad \kappa_{\mathrm{sc}}(d) \geq 1.$$

  This was proved in [HR90] by studying the next order in $\hbar$ in the semi-classical limit. Equivalently, one can study the stability of the constant function $V = \mu$ in the periodic Lieb-Thirring inequality.

- ($\kappa_{(2N)\cap\mathrm{sc}} > \kappa_{N\cap\mathrm{sc}}$) Our Theorem 4.1 can be recast as follows.

  If $\kappa_{N\cap\mathrm{sc}} > \max\{0, 2-\frac{d}{2}\}$, then $\kappa_{(2N)\cap\mathrm{sc}}(d) > \kappa_{N\cap\mathrm{sc}}(d)$. Hence $\forall N \in \mathbb{N},\ \kappa_{\mathrm{sc}}(d) > \kappa_{N\cap\mathrm{sc}}(d)$.

  This is particularly interesting in dimension $d = 2$, since $\kappa_{1\cap\mathrm{sc}} \approx 1.1654 > 1 = \max\{0, 2-\frac{d=2}{2}\}$. This implies that $\kappa_{\mathrm{sc}}(2) > \kappa_{N\cap\mathrm{sc}}(2)$ for all $N \in \mathbb{N}$. We display in Figure 4.3 a periodic potential V which «beats» both the $L^{(1)}_{\kappa,d}$ and the $L^{\mathrm{sc}}_{\kappa,d}$ scenarios, for some $\kappa_{1\cap\mathrm{sc}}(2) < \kappa < \kappa_{\mathrm{sc}}(2)$.

In view of all these considerations, we made the following conjecture in [G13].

Conjecture 4.3. *For all κ satisfying (4.1), we have*

- **either** *there is $N \in \mathbb{N}$ so that $L_{\kappa,d} = L^{(N)}_{\kappa,d}$. In particular, any optimizer $V_N \in L^{\kappa+\frac{d}{2}}(\mathbb{R}^d)$ for $L^{(N)}_{\kappa,d}$ is an optimizer for $L_{\kappa,d}$, and the operator $-\Delta - V_N$ has less than N negative eigenvalues.*
- **or** *we have $L_{\kappa,d} > L^{(N)}_{\kappa,d}$ for all N, in which case the periodic Lieb-Thirring inequality has a periodic optimizer V_*.*

By definition of $\kappa_{\mathrm{sc}}(d)$, we know that for $\kappa > \kappa_{\mathrm{sc}}(d)$, the constant functions are optimizers of the periodic Lieb-Thirring inequality (liquid phase). On the other hand, if $\kappa < \kappa_{\mathrm{sc}}(d)$, then the liquid phase is not optimal. But for $\kappa > \max\{0, 2-\frac{d}{2}\}$, we also proved that the N-particle scenario is never optimal. So our conjecture would imply the following other one.

Conjecture 4.4. *For all $\max\{0, 2-\frac{d}{2}\} < \kappa < \kappa_{\mathrm{sc}}(d)$, the periodic Lieb-Thirring inequality has a periodic optimal optimizer which is not a constant potential (crystallized phase).*

We sum up what is known in dimensions $d \in \{1, 2, 3\}$ in Figure 4.2. The conjecture that $L_{\kappa,1} = L^{(1)}_{\kappa,1}$ for $\frac{1}{2} < \kappa < 3/2$ in dimension $d = 1$ comes from numerical simulations in [Lev14], and the conjecture that $\kappa_{\mathrm{sc}}(d = 3) = 1$ is the original Lieb-Thirring conjecture.

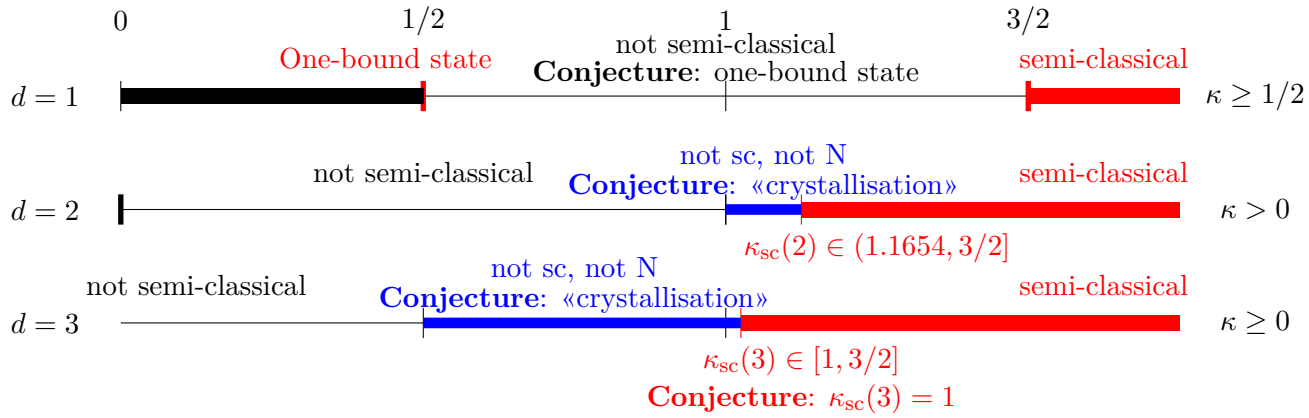


Figure 4.2: What is known and conjectured about Lieb-Thirring best scenarii for $d \in \{1, 2, 3\}$.

We conclude this section by exhibiting a two–dimensional periodic potential V_* for $\kappa = 1.165400$, which beats both the semi-classical and the 1-particle scenario. Our numerical simulations performed in [G13] suggest that $\kappa_{\mathrm{sc}}(2) \approx 1.165417$. Recall that $\kappa_{1\cap\mathrm{sc}}(2) = 1.165378$. The fact that these two values are extremely close is probably a manifestation of the exponential gain of energy that we found in the proof of Theorem 4.1.

It turns out that the corresponding operator $-\Delta - V_*$ describes an insulator (there is a gap at the Fermi level $\varepsilon_F = 0$). This allowed us to compute the quantity $\underline{\mathrm{Tr}}(-\Delta - V_*)$ with high precision (see Section 1.5.1), which was necessary to capture the exponentially small attraction.

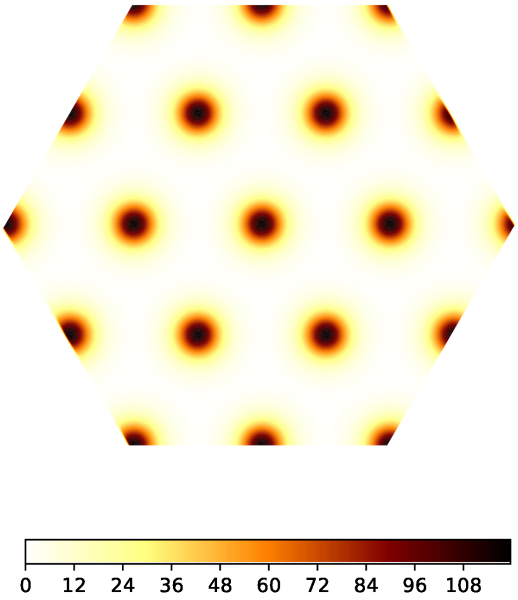


Figure 4.3: Numerical simulation of the best optimal potential (among the ones having the triangular symmetry) for the periodic Lieb-Thirring inequality, for $\kappa = 1.165400 > \kappa_{1\cap \mathrm{sc}}(2)$.

4.3 Fermionic non-linear Schrödinger

4.3.1 Fermionic minimization problems

In this section, we describe the results obtained in [G11]. This is joint work with Mathieu Lewin and Faizan Nazar.

Let us now present another inequality, namely the fermionic non-linear Schrödinger (NLS) equation, which exhibits similar features. We explain the connection between the two problems in the next section. For $0 \leq \gamma = \gamma^* $ a positive compact self-adjoint operator on $L^2(\mathbb{R}^d)$, we denote by $\rho_\gamma(x) := \gamma(x,x)$ its density, and by

$$\mathrm{Tr}(-\Delta\gamma) := \sum_{j=1}^{d} \mathrm{Tr}(P_j \gamma P_j) \quad \text{with} \quad P_j := -\mathrm{i}\partial_j.$$

its kinetic energy. For an exponent

$$1 < p < 1 + \frac{2}{d}, \tag{4.11}$$

we define the NLS minimization problem

$$\boxed{J^{\mathrm{NLS}}_{p,d}(N) := \inf\left\{ \mathrm{Tr}(-\Delta\gamma) - \frac{1}{p}\int_{\mathbb{R}^d} \rho_\gamma^p, \quad \gamma \in \mathcal{S}\left(L^2(\mathbb{R}^d)\right), \quad 0 \leq \gamma \leq 1, \quad \mathrm{Tr}(\gamma) = N \right\}.}$$

The denomination *fermionic* comes from the Pauli constraint $0 \leq \gamma \leq 1$. The condition (4.11) together with the Lieb-Thirring inequality implies that $J^{\mathrm{NLS}}(N)$ is always bounded from below. Note that the problem is concave in γ, hence, if a minimizer exists, it must be a projector of rank N (and in particular $\mathrm{Rank}(\gamma) = \mathrm{Tr}(\gamma) = N$).

It is unclear that the problem $J^{\mathrm{NLS}}(N)$ is always well-posed (*i.e.* admits minimizers). It will be the case if the following strong binding condition holds (see [Lew11])

$$\forall 1 \leq k \leq N-1, \quad J^{\mathrm{NLS}}_{p,d}(N) < J^{\mathrm{NLS}}_{p,d}(N-k) + J^{\mathrm{NLS}}_{p,d}(k).$$

Note that the weak binding inequality

$$\forall 1 \leq k \leq N-1, \quad J^{\mathrm{NLS}}_{p,d}(N) \leq J^{\mathrm{NLS}}_{p,d}(N-k) + J^{\mathrm{NLS}}_{p,d}(k). \tag{4.12}$$

always hold, as can be seen by putting trial states for $(N-k)$ and k–particle far from each other. The main result of [G11] is the following.

Theorem 4.5 ([G11] Theorem 4)**.** *For all p satisfying (4.11), the problem $J^{\mathrm{NLS}}_{p,d}(N=1)$ is well-posed. Under the extra condition $1<p<2$, we have*

$$\forall N \in \mathbb{N}, \qquad J^{\mathrm{NLS}}_{p,d}(2N) < 2J^{\mathrm{NLS}}_{p,d}(N). \tag{4.13}$$

In particular, for $1<p<2$, there is an infinite sequence of integers $1=N_1<N_2<N_3<\cdots$ so that the problem $J^{\mathrm{NLS}}_{p,d}(N_k)$ is well-posed.

If γ is such an optimizer, it is a projector, and the corresponding Euler–Lagrange equations take the form

$$\gamma = \mathbb{1}\left(-\Delta - V < \varepsilon_F\right) + \widetilde{\gamma}, \qquad \text{with} \quad V := \rho_\gamma^{p-1}, \tag{4.14}$$

for some Fermi level $\varepsilon_F \leq 0$, and $0 \leq \widetilde{\gamma} \leq \mathbb{1}(-\Delta - V = \varepsilon_F)$. Equation (4.14) can be rewritten as follows: there is an orthonormal family $(u_1, \cdots, u_N)$ in $L^2(\mathbb{R}^d)$ so that $\gamma = \sum_{j=1}^N |u_j\rangle\langle u_j|$, and which solves the *fermionic non–linear Schrödinger equation*

$$\forall 1 \leq j \leq N, \quad \left(-\Delta - \rho^{p-1}\right) u_j = \mu_j u_j, \qquad \text{with} \quad \rho = \sum_{j=1}^N |u_j|^2,$$

where $\mu_1 < \mu_2 \leq \cdots \leq \mu_N \leq \varepsilon_F$ are the N lowest eigenvalues of $-\Delta - \rho^{p-1}$.

Proof of Theorem 4.5. The first part of the proof follows the lines of Theorem 4.1. Actually, we found this result first, and this is how we realized that one should add the densities instead of the potentials in (4.7).

For the second part of the proof, we use a pigeon–hole argument. Let $\mathcal{N} \subset \mathbb{N}$ be the subset of integers N so that $J^{\mathrm{NLS}}_{p,d}(N)$ has a minimizer. Assume $\mathcal{N}$ is finite, and let M be the sum of all elements in $\mathcal{N}$. Consider the minimization problem for $M+1$ particles. Since $M+1 \notin \mathcal{N}$, the strong binding inequality must fail for some $1 \leq k \leq M-1$. However, the weak-binding inequality always holds. Iterating this reasoning, we deduce that we must have a splitting of the form

$$J^{\mathrm{NLS}}_{p,d}(M+1) = \sum_k J^{\mathrm{NLS}}_{p,d}(N_k), \quad \text{with} \quad M+1 = \sum_k N_k \quad \text{and} \quad N_k \in \mathcal{N}.$$

In other words, any minimization sequence for $\mathcal{J}^{\mathrm{NLS}}_{p,d}(M+1)$ splits into stable ones. Our hypothesis for M shows that at least two bubbles must have the same number of particles, say $N_1 = N_2 \in \mathcal{N}$. Inequality (4.13) implies $J^{\mathrm{NLS}}_{p,d}(N_1+N_2) < J^{\mathrm{NLS}}_{p,d}(N_1) + J^{\mathrm{NLS}}_{p,d}(N_2)$, and the weak binding inequality gives the contradiction

$$J^{\mathrm{NLS}}_{p,d}(M+1) = \sum_k J^{\mathrm{NLS}}_{p,d}(N_k) > J^{\mathrm{NLS}}_{p,d}(N_1+N_2) + \sum_{k\geq 3} J^{\mathrm{NLS}}_{p,d}(N_k) \geq J^{\mathrm{NLS}}_{p,d}(M+1).$$

So $\mathcal{N}$ is infinite. □

We display in Figures 4.4 and 4.5 numerical computations for the density of the minimizer of $J^{\mathrm{NLS}}_{p,d}$ in the one– and two–dimensional case. In both these figures, we observe a crystallization phenomenon, where well-localized bubbles are bound by an exponentially small attraction. It is tempting to interpret these bubbles as the location of the fermionic quantum particles.

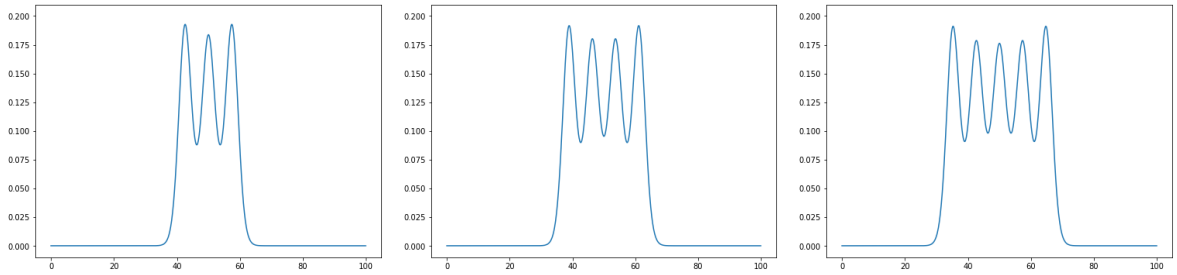


Figure 4.4: Optimal density for $J^{\mathrm{NLS}}_{p,d}$ in the case $d = 1$, $p = 1.3$ and $N = 3, 4, 5$.

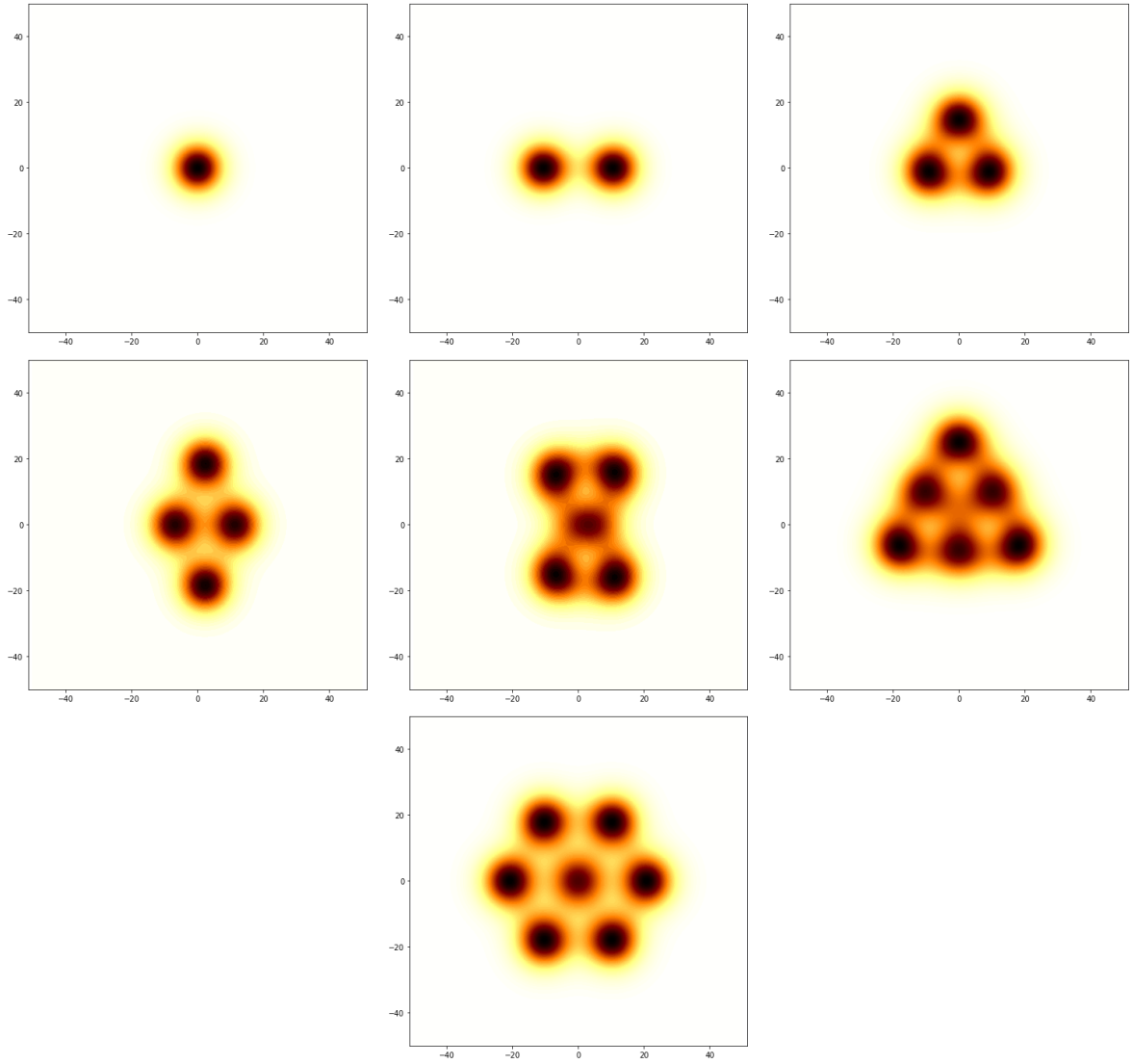


Figure 4.5: Optimal density for $J^{\mathrm{NLS}}_{p,d}$ in the case $d = 2$, $p = 1.5$ and for N from 1 to 7.

4.3.2 Dual formulation of Lieb-Thirring and NLS inequalities

Finally, we relate the Lieb-Thirring problem and NLS problem with their dual formulations.

Dual Lieb-Thirring

When $\kappa \geq 1$, the Lieb-Thirring inequality has the following dual formulation. Recall that, for a compact operator A, we denote by $\|A\|_{\mathfrak{S}_q} := \mathrm{Tr}(|A|^q)^{1/q}$ the Schatten norm of A. The dual version of the Lieb-Thirring inequalities takes the following form (see [G12]).

Lemma 4.6. *Let $\kappa \geq 1$ and let*

$$p := \frac{\kappa + \frac{d}{2}}{\kappa + \frac{d}{2} - 1}, \quad \text{and} \quad q := \frac{\kappa}{\kappa - 1}$$

be the dual exponents of $\kappa + \frac{d}{2}$ and κ respectively. Then there is an optimal (largest) constant $K_{p,d}^{(N)}$ so that, for all γ of rank N, we have

$$\boxed{K_{p,d}^{(N)} \|\rho_\gamma\|_{L^p(\mathbb{R}^d)}^{\frac{2p}{d(p-1)}} \leq \|\gamma\|_{\mathfrak{S}_q}^{\frac{p(2-d)+d}{d(p-1)}} \mathrm{Tr}(-\Delta\gamma).} \tag{4.15}$$

In addition, the constants $K_{p,d}^{(N)}$ and $L_{\kappa,d}^{(N)}$ satisfy the relation

$$K_{p,d}^{(N)} \left(L_{\kappa,d}^{(N)}\right)^{\frac{2}{d}} = \left(\frac{\kappa}{\kappa + \frac{d}{2}}\right)^{\frac{2\kappa}{d}} \left(\frac{d}{2\kappa + d}\right). \tag{4.16}$$

When $N = 1$, an operator γ of rank 1 is of the form $\gamma = |u\rangle\langle u|$ for some (un-normalized) function u. In this case, $\rho_\gamma = |u|^2$. We find that $K_{p,d}^{(1)}$ is the optimal constant in the inequality

$$\forall u \in H^1(\mathbb{R}^d), \qquad K_{p,d}^{(1)} \|u\|_{L^{2p}}^{\frac{4p}{d(p-1)}} \leq \|u\|_{L^2}^{\frac{2p(2-d)+2d}{d(p-1)}} \int_{\mathbb{R}^d} |\nabla u|^2.$$

We recognize the Gagliardo–Niremberg inequalities, so $K_{p,d}^{(1)}$ is, actually, the Gagliardo–Niremberg constant $K_{p,d}^{(1)} = K_{p,d}^{\mathrm{GN}}$. The condition $\kappa \geq 1$ translates into $1 \leq p \leq 1 + \frac{2}{d}$.

On the other hand, taking the critical value $\kappa = 1$, which implies $p = 1 + \frac{2}{d}$ and $q = \infty$, we recover the usual kinetic Lieb–Thirring inequality

$$K_{\frac{d+2}{d},d}^{(N)} \left(\int_{\mathbb{R}^d} \rho_\gamma^{1+\frac{2}{d}}\right) \leq \mathrm{Tr}(-\Delta\gamma),$$

valid for all one-body density operator γ of rank N, and satisfying $0 \leq \gamma \leq 1$.

Going back to our Theorem 4.1, which states that $L_{\kappa,d}^{(2N)} > L_{\kappa,d}^{(N)}$ whenever $\kappa > \max\{0, 2 - \frac{d}{2}\}$ and $\kappa \geq 1$, together with (4.16), we obtain that for all $1 \leq p \leq 1 + \frac{d}{2}$ satisfying the extra condition $p < 2$, we have

$$K_{p,d}^{(2N)} > K_{p,d}^{(N)}.$$

Dual NLS

Similarly, we have a dual formulation of the NLS problem.

Lemma 4.7. *Let $1 < p < 1 + \frac{2}{d}$. Then there is an optimal (largest) constant $\widetilde{K}_{p,d}^{(N)}$, so that, for all γ with $\mathrm{Tr}(\gamma) = N$, we have*

$$\widetilde{K}_{p,d}^{(N)} \|\rho\|_{L^p(\mathbb{R}^d)}^{\frac{2p}{2(p-1)}} \leq N^{\frac{p(2-d)+d}{d(p-1)}\frac{1}{q}} \|\gamma\|^{\frac{p(2-d)+d}{d(p-1)}} \mathrm{Tr}(-\Delta\gamma), \tag{4.17}$$

In addition, the constants $\widetilde{K}_{p,d}^{(N)}$ and $J_{p,d}^{\mathrm{NLS}}(N)$ satisfy the relation

$$\widetilde{K}_{p,d}^{(N)} = \left[\left(\frac{J_{p,d}^{\mathrm{NLS}}(N)}{N}\right)\left(1 + \frac{2}{d} - p\right)\right]^{-\frac{d+2-pd}{d(p-1)}} \frac{1}{p-1}\left(\frac{d}{2p}\right)^{\frac{2}{d(p-1)}}.$$

In particular, Theorem 4.5 implies that if p satisfies $1 < p \leq 1 + \frac{2}{d}$ and the extra condition $p < 2$, then

$$\widetilde{K}_{p,d}^{(2N)} < \widetilde{K}_{p,d}^{(N)}.$$

This leaves open the case $d = 1$, and $p \in [2, 3]$. The relation between the two constants $K_{p,d}^{(N)}$ and $\widetilde{K}_{p,d}^{(N)}$ are given by the following Lemma.

Lemma 4.8. *We have*

$$K_{p,d}^{(N)} \leq \widetilde{K}_{p,d}^{(N)} \leq \widetilde{K}_{p,d}^{(1)} = K_{p,d}^{(1)}.$$

In addition, for $d = 1$ and $p = 2$, we have

$$K_{2,1}^{(N)} = \widetilde{K}_{2,1}^{(N)} = \widetilde{K}_{2,1}^{(1)} = K_{2,1}^{(1)}.$$

Proof. The first part, we use the simple bound $\|\gamma\|_{\mathfrak{S}_q} \leq N^{\frac{1}{q}}\|\gamma\|$, valid for all positive γ of rank less than N. This bound shows that (4.15) implies (4.17) with the constant $K_{p,d}^{(N)}$. But since $\widetilde{K}_{p,d}^{(N)}$ is the largest constant for which (4.17) holds, we deduce that $K_{p,d}^{(N)} \leq \widetilde{K}_{p,d}^{(N)}$. The fact that $N \mapsto \widetilde{K}_{p,d}^{(N)}$ is decreasing is standard. Finally, for $N = 1$, we have the equality $\|\gamma\|_{\mathfrak{S}_q} = \|\gamma\|$ for all positive γ of rank 1. So the dual Lieb–Thirring and dual NLS problems are equivalent at $N = 1$.

The second part comes from the equality $L_{3/2,1}^{(N)} = L_{3/2,1}^{(1)}$, which implies $K_{2,1}^{(N)} = K_{2,1}^{(1)}$. □

4.4 Perspectives

The exponential attraction between *bubbles* is a phenomenon that we expect to find in other fermionic models. Together with Salma LAHBABI and Simona ROTA NODARI, we would like to see whether it can be applied to mean–field models for nucleons. It might explain the non–spherical shapes of nuclei that we find experimentally.

In the recent work [G24], together with Jean DOLBEAULT, Fabio PIZZICHILLO and Hanne VAN DEN BOSCH, we derived Keller and Lieb–Thirring inequalities for (massive) Dirac operators, of the form $\not{\mathcal{D}}_m - V$, with positive potentials $V \in L^p(\mathbb{R}^d, \mathbb{R}_+)$. Here, $\not{\mathcal{D}}_m := \sum_j (-\mathrm{i}\partial_j)\alpha_j + m\beta$ is the usual free Dirac operator with mass $m > 0$, acting on $L^2(\mathbb{R}^d, \mathbb{C}^N)$. We would like to extend this work to the case $\not{\mathcal{D}}_m - V\beta$.

BIBLIOGRAPHY

Contributions of the author

[G1] H. D. Cornean, D. Gontier, A. Levitt, and D. Monaco. "Localised Wannier Functions in Metallic Systems". In: *Annales Henri Poincaré* 20.4 (2019), pp. 1367–1391.

[G2] D. Gontier, A. Levitt, and S. Siraj-dine. "Numerical construction of Wannier functions through homotopy". In: *Journal of Mathematical Physics* 60.3 (2019), p. 031901.

[G3] É. Cancès, V. Ehrlacher, D. Gontier, A. Levitt, and D. Lombardi. "Numerical quadrature in the Brillouin zone for periodic Schrödinger operators". In: *Numerische Mathematik* 144.3 (2020), pp. 479–526.

[G4] D. Gontier, D. Monaco, and S. Perrin-Roussel. "Symmetric Fermi projections and Kitaev's table: Topological phases of matter in low dimensions". In: *Journal of Mathematical Physics* 63.4 (2022), p. 041902.

[G5] D. Gontier. "Edge states in ordinary differential equations for dislocations". In: *Journal of Mathematical Physics* 61.4 (2020), p. 043507.

[G6] D. Gontier. "Edge states for second order elliptic operators in a channel". In: *Journal of Spectral Theory* 12.3 (2023), pp. 1155–1202.

[G7] D. Gontier. "Spectral properties of periodic systems cut at an angle". In: *Comptes Rendus Mathématique* 359.8 (2021), pp. 949–958.

[G8] D. Gontier and M. Lewin. "Spin Symmetry Breaking in the Translation-Invariant Hartree–Fock Electron Gas". In: *SIAM Journal on Mathematical Analysis* 51.4 (2019), pp. 3388–3423.

[G9] D. Gontier, C. Hainzl, and M. Lewin. "Lower bound on the Hartree-Fock energy of the electron gas". In: *Physical Review A* 99.5 (2019).

[G10] D. Gontier, A. E. K. Kouande, and É. Séré. "Phase Transition in the Peierls Model for Polyacetylene". In: *Annales Henri Poincaré* 24.11 (2023), pp. 3945–3966.

[G11] D. Gontier, M. Lewin, and F. Q. Nazar. "The nonlinear Schrödinger Equation for Orthonormal Functions: Existence of Ground States". In: *Archive for Rational Mechanics and Analysis* 240.3 (2021), pp. 1203–1254.

[G12] R. L. Frank, D. Gontier, and M. Lewin. "The Nonlinear Schrödinger Equation for Orthonormal Functions II: Application to Lieb–Thirring Inequalities". In: *Communications in Mathematical Physics* 384.3 (2021), pp. 1783–1828.

[G13] R. L. Frank, D. Gontier, and M. Lewin. "The periodic Lieb-Thirring inequality". In: *Partial Differential Equations, Spectral Theory, and Mathematical Physics. The Ari Laptev Anniversary Volume*. Ed. by P. Exner, R. L. Frank, F. Gesztesy, H. Holden, and T. Weidl. Vol. 18. EMS Publishing House, 2021.

[G14] R. L. Frank, D. Gontier, and M. Lewin. "Optimizers for the finite-rank Lieb-Thirring inequality". 2023.

[G15] H. Ammari, B. Fitzpatrick, D. Gontier, H. Lee, and H. Zhang. "A Mathematical and Numerical Framework for Bubble Meta-Screens". In: *SIAM Journal on Applied Mathematics* 77.5 (2017), pp. 1827–1850.

[G16] H. Ammari, B. Fitzpatrick, D. Gontier, H. Lee, and H. Zhang. "Sub-wavelength focusing of acoustic waves in bubbly media". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 473.2208 (2017), p. 20170469.

[G17] H. Ammari, B. Fitzpatrick, D. Gontier, H. Lee, and H. Zhang. "Minnaert resonances for acoustic waves in bubbly media". In: *Annales de l'Institut Henri Poincaré C, Analyse non linéaire* 35.7 (2018), pp. 1975–1998.

[G18] D. Gontier and S. Lahbabi. "The reduced Hartree-Fock model with self-generated magnetic fields". In: *Journal of Mathematical Physics* 60.8 (2019), p. 081902.

[G19] A. Bakhta, V. Ehrlacher, and D. Gontier. "Numerical reconstruction of the first band(s) in an inverse Hill's problem". In: *ESAIM: Control, Optimisation and Calculus of Variations* 26 (2020), p. 59.

[G20] D. Gontier, S. Lahbabi, and A. Maichine. "Density Functional Theory for Two-Dimensional Homogeneous Materials". In: *Communications in Mathematical Physics* 388.3 (2021), pp. 1475–1505.

[G21] D. Gontier, S. Lahbabi, and A. Maichine. "Density functional theory for two-dimensional homogeneous materials with magnetic fields". In: *Journal of Functional Analysis* 285.9 (2023), p. 110100.

[G22] É. Cancès, L. Garrigue, and D. Gontier. "Simple derivation of moiré-scale continuous models for twisted bilayer graphene". In: *Physical Review B* 107.15 (2023), p. 155403.

[G23] É. Cancès, L. Garrigue, and D. Gontier. "Second-Order Homogenization of Periodic Schrödinger Operators with Highly Oscillating Potentials". In: *SIAM Journal on Mathematical Analysis* 55.3 (2023), pp. 2288–2323.

[G24] J. Dolbeault, D. Gontier, F. Pizzichillo, and H. Van Den Bosch. "Keller and Lieb-Thirring estimates of the eigenvalues in the gap of Dirac operators". In: *Revista Matemática Iberoamericana* (2023).

Other References

[AL78] M. Aizenman and E. H. Lieb. "On semiclassical bounds for eigenvalues of Schrödinger operators". In: *Physical Letters A* 66.6 (1978), pp. 427–429.

[AZ97] A. Altland and M. R. Zirnbauer. "Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures". In: *Physical Review B* 55 (2 1997), pp. 1142–1161.

[APS76] M. Atiyah, V. Patodi, and I. Singer. "Spectral asymmetry and Riemannian geometry. III". In: *Math. Proc. Camb. Philos. Soc.* Vol. 79. 1. Cambridge University Press. 1976, pp. 71–99.

[ASV13] J. Avila, H. Schulz-Baldes, and C. Villegas-Blas. "Topological invariants of edge states for periodic two-dimensional models". In: *Mathematical Physics, Analysis and Geometry* 16.2 (2013), pp. 137–170.

[Bag+14] L. Baguet, F. Delyon, B. Bernu, and M. Holzmann. "Properties of Hartree-Fock solutions of the three-dimensional electron gas". In: *Physical Review B* 90.16 (2014), p. 165131.

[Bag+13] L. Baguet, F. Delyon, B. Bernu, and M. Holzmann. "Hartree-Fock ground state phase diagram of jellium". In: *Physical Review Letters* 111.16 (2013), p. 166402.

[BL99] R. D. Benguria and M. Loss. "A simple proof of a theorem of Laptev and Weidl". In: *Mathematical Research Letters* 7 (1999), pp. 195–203.

[BJA94] P. Blöchl, O. Jepsen, and O. Andersen. "Improved tetrahedron method for Brillouin-zone integrations". In: *Physical Review B* 49.23 (1994), pp. 16223–16233.

[BGO10] G. Bräunlich, G. Graf, and G. Ortelli. "Equivalence of topological and scattering approaches to quantum pumping". In: *Communications in Mathematical Physics* 295.1 (2010), pp. 243–259.

[Bro+07] C. Brouder, G. Panati, M. Calandra, C. Mourougane, and N. Marzari. "Exponential Localization of Wannier Functions in Insulators". In: *Physical Review Letters* 98.4 (2007).

[Can+17] É. Cancès, A. Levitt, G. Panati, and G. Stoltz. "Robust determination of maximally localized Wannier functions". In: *Physical Review B* 95.7 (2017), p. 075114.

[CBC04] L. Cândido, B. Bernu, and D. Ceperley. "Magnetic ordering of the three-dimensional Wigner crystal". In: *Physical Review B* 70.9 (2004), p. 094413.

[Car26] É. Cartan. "Sur une classe remarquable d'espaces de Riemann. I". fr. In: *Bulletin de la Société Mathématique de France* 54 (1926), pp. 214–264.

[Car27] É. Cartan. "Sur une classe remarquable d'espaces de Riemann. II". fr. In: *Bulletin de la Société Mathématique de France* 55 (1927), pp. 114–134.

[CA80] D. M. Ceperley and B. J. Alder. "Ground State of the Electron Gas by a Stochastic Method". In: *Physical Review Letters* 45 (7 1980), pp. 566–569.

[Clo64a] J. des Cloizeaux. "Analytical properties of n-dimensional energy bands and Wannier functions". In: *Physical Review* 135.3A (1964), A698.

[Clo64b] J. des Cloizeaux. "Energy bands and projection operators in a crystal: Analytic and asymptotic properties". In: *Physical Review* 135.3A (1964), A685.

[CT73] J.-M. Combes and L. Thomas. "Asymptotic behaviour of eigenfunctions for multiparticle Schrödinger operators". In: *Communications in Mathematical Physics* 34.4 (1973), pp. 251–270.

[CHN16] H. Cornean, I. Herbst, and G. Nenciu. "On the construction of composite Wannier functions". In: *Annales Henri Poincaré* 17.12 (2016), pp. 3361–3398.

[CMM19] H. D. Cornean, D. Monaco, and M. Moscolari. "Parseval frames of exponentially localized magnetic Wannier functions". In: *Communications in Mathematical Physics* 371.3 (2019), pp. 1179–1230.

[Cwi77] M. Cwikel. "Weak type estimates for singular values and the number of bound states of Schrodinger operators". In: *Annals of Mathematics* (1977), pp. 93–100.

[DLY17] A. Damle, L. Lin, and L. Ying. "SCDM-k: Localized orbitals for solids via selected columns of the density matrix". In: *Journal of Computational Physics* 334 (2017), pp. 1–15.

[DLL19] A. Damle, A. Levitt, and L. Lin. "Variational Formulation for Wannier Functions with Entangled Band Structure". In: *Multiscale Modeling & Simulation* 17.1 (2019), pp. 167–191.

[Del+15] F. Delyon, B. Bernu, L. Baguet, and M. Holzmann. "Upper bounds of spin-density wave energies in the homogeneous electron gas". In: *Physical Review B* 92.23 (2015), p. 235124.

[Dro21a] A. Drouot. "Microlocal Analysis of the Bulk-Edge Correspondence". In: *Communications in Mathematical Physics* 383.3 (2021), pp. 2069–2112.

[Dro21b] A. Drouot. "The bulk-edge correspondence for continuous dislocated systems". en. In: *Annales de l'Institut Fourier* 71.3 (2021), pp. 1185–1239.

[EG02] P. Elbau and G. Graf. "Equality of bulk and edge Hall conductance revisited". In: *Communications in Mathematical Physics* 229.3 (2002), pp. 415–432.

[EGS05] A. Elgart, G. Graf, and J. Schenker. "Equality of the bulk and edge Hall conductances in a mobility gap". In: *Communications in Mathememtical Physics* 259.1 (2005), pp. 185–221.

[FMP16] D. Fiorenza, D. Monaco, and G. Panati. "Construction of real-valued localized composite Wannier functions for insulators". In: *Annales Henri Poincaré* 17.1 (2016), pp. 63–97.

[FL11] R. L. Frank and E. H. Lieb. "Possible lattice distortions in the Hubbard model for graphene". In: *Physical Review Letters* 107.6 (2011), p. 066801.

[FLW22] R. L. Frank, A. Laptev, and T. Weidl. *Schrödinger Operators: Eigenvalues and Lieb–Thirring Inequalities*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2022.

[Frö54] H. Fröhlich. "On the theory of superconductivity: The one-dimensional case". In: *Proceedings of the Royal Society of London (A)* 223.1154 (1954), pp. 296–305.

[GL16] D. Gontier and S. Lahbabi. "Convergence rates of supercell calculations in the reduced Hartree-Fock model". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 50.5 (2016), pp. 1403–1424.

[GO08] G. Graf and G. Ortelli. "Comparison of quantization of charge transport in periodic and open pumps". In: *Physical Review B* 77.3 (2008), p. 033304.

[GP13] G. Graf and M. Porta. "Bulk-edge correspondence for two-dimensional topological insulators". In: *Communications in Mathematical Physics* 324.3 (2013), pp. 851–895.

[Gra07] G. M. Graf. "Aspects of the integer quantum Hall effect". In: *Proceedings of Symposia in Pure Mathematics*. Vol. 76. 1. Providence, RI; American Mathematical Society; 1998. 2007, p. 429.

[GS18] G. M. Graf and J. Shapiro. "The Bulk-Edge Correspondence for Disordered Chiral Chains". In: *Communications in Mathematical Physics* 363.3 (2018), pp. 829–846.

[HS10] C. Hainzl and R. Seiringer. "Asymptotic behavior of eigenvalues of Schrödinger type operators with degenerate kinetic energy". In: *Mathematische Nachrichten* 283.3 (2010), pp. 489–499.

[Hal82] B. Halperin. "Quantized Hall conductance, current-carrying edge states, and the existence of extended states in a two-dimensional disordered potential". In: *Physical Review B* 25.4 (1982), p. 2185.

[Hat93a] Y. Hatsugai. "Chern number and edge states in the integer quantum Hall effect". In: *Physical Review Letters* 71.22 (1993), p. 3697.

[Hat93b] Y. Hatsugai. "Edge states in the integer quantum Hall effect and the Riemann surface of the Bloch function". In: *Physical Review B* 48.16 (1993), p. 11851.

[HHZ05] P. Heinzner, A. Huckleberry, and M. Zirnbauer. "Symmetry Classes of Disordered Fermions". In: *Communications in Mathematical Physics* 257 (2005), pp. 725–771.

[HR90] B. Helffer and D. Robert. "Riesz means of bounded states and semi-classical limit connected with a Lieb-Thirring conjecture. II". In: *Annales de l'Institut Henri Poincaré. Physique Théorique* 53.2 (1990), pp. 139–147.

[HS89] B. Helffer and J. Sjöstrand. "Equation de Schrödinger avec champ magnétique et équation de Harper". In: *Lecture Notes in Physics* 345 (1989), pp. 118–197.

[HK11a] R. Hempel and M. Kohlmann. "Spectral properties of grain boundaries at small angles of rotation". In: *Journal of Spectral Theory* 1 (2 2011), pp. 197–219.

[HK12] R. Hempel and M. Kohlmann. "Dislocation problems for periodic Schrödinger operators and mathematical aspects of small angle grain boundaries". In: *Spectral Theory, Mathematical System Theory, Evolution Equations, Differential and Difference Equations*. Springer, 2012, pp. 421–432.

[Hem+15] R. Hempel, M. Kohlmann, M. Stautz, and J. Voigt. "Bound states for nano-tubes with a dislocation". In: *Journal of Mathematical Analysis and Applications* 431.1 (2015), pp. 202–227.

[HK11b] R. Hempel and M. Kohlmann. "A variational approach to dislocation problems for periodic Schrödinger operators". In: *Journal of mathematical analysis and applications* 381.1 (2011), pp. 166–178.

[HM20] M. Holzmann and S. Moroni. "Itinerant-electron magnetism: The importance of many-body correlations". In: *Physical Review Letters* 124.20 (2020), p. 206404.

[HLT98] D. Hundertmark, E. H. Lieb, and L. E. Thomas. "A sharp bound for an eigenvalue moment of the one-dimensional Schrödinger operator". In: *Advances in Theoretical and Mathematical Physics* 2.4 (1998), pp. 719–731.

[Kat13] T. Kato. *Perturbation theory for linear operators*. Vol. 132. Springer Science & Business Media, 2013.

[KRS02] J. Kellendonk, T. Richter, and H. Schulz-Baldes. "Edge current channels and Chern numbers in the integer quantum Hall effect". In: *Reviews in Mathematical Physics* 14.01 (2002), pp. 87–119.

[Kel61] J. B. Keller. "Lower Bounds and Isoperimetric Inequalities for Eigenvalues of the Schrödinger Equation". In: *Journal of Mathematical Physics* 2.2 (1961), pp. 262–266.

[KL04] T. Kennedy and E. H. Lieb. "Proof of the Peierls instability in one dimension". In: *Condensed Matter Physics and Exactly Soluble Models*. Springer, 2004, pp. 85–88.

[Koh59] W. Kohn. "Analytic properties of Bloch waves and Wannier functions". In: *Physical Review* 115.4 (1959), p. 809.

[KS65] W. Kohn and L. Sham. "Self-consistent equations including exchange and correlation effects". In: *Physical Review* 140.4A (1965), A1133–A1138.

[Kor00] E. Korotyaev. "Lattice dislocations in a 1-dimensional model". In: *Communications in Mathematical Physics* 213.2 (2000), pp. 471–489.

[Kor05] E. Korotyaev. "Schrödinger operator with a junction of two 1-dimensional periodic potentials". In: *Asymptotic Analysis* 45.1, 2 (2005), pp. 73–97.

[LSW02] A. Laptev, O. Safronov, and T. Weidl. "Bound state asymptotics for elliptic operators with strongly degenerated symbols". In: *Nonlinear problems in mathematical physics and related topics, I*. Vol. 1. Int. Math. Ser. (N. Y.) Kluwer/Plenum, New York, 2002, pp. 233–246.

[LW00] A. Laptev and T. Weidl. "Sharp Lieb-Thirring inequalities in high dimensions". In: *Acta Mathematica* 184.1 (2000), pp. 87–111.

[Lev14] A. Levitt. "Best constants in Lieb-Thirring inequalities: a numerical investigation". In: *Journal of Spectral Theory* 4.1 (June 2014), pp. 153–175.

[Lew11] M. Lewin. "Geometric methods for nonlinear many-body quantum systems". In: *Journal of Functional Analysis* 260.12 (2011), pp. 3535–3595.

[LLS19] M. Lewin, E. H. Lieb, and R. Seiringer. "The local density approximation in Density Functional Theory". In: *Pure and Applied Analysis* 2.1 (2019), pp. 35–73.

[LT75] E. Lieb and W. Thirring. "Bound on kinetic energy of fermions which proves stability of matter". In: *Physical Review Letters* 35 (1975), pp. 687–689.

[LT76] E. Lieb and W. Thirring. "Inequalities for the moments of the eigenvalues of the Schrödinger hamiltonian and their relation to Sobolev inequalities". In: Studies in Mathematical Physics. Princeton University Press, 1976, pp. 269–303.

[Lie76] E. H. Lieb. "Bounds on the eigenvalues of the Laplace and Schroedinger operators". In: *Bulletin of the American Mathematical Society* 82.5 (1976), pp. 751–753.

[LN75] E. H. Lieb and H. Narnhofer. "The thermodynamic limit for jellium". In: *Journal of Statistical Physics* 12.4 (1975), pp. 291–310.

[MV97] N. Marzari and D. Vanderbilt. "Maximally localized generalized Wannier functions for composite energy bands". In: *Physical Review B* 56.20 (1997), p. 12847.

[Mar+99] N. Marzari, D. Vanderbilt, A. De Vita, and M. Payne. "Thermal contraction and disordering of the Al (110) surface". In: *Physical Review Letters* 82.16 (1999), p. 3296.

[MP89] M. Methfessel and A. Paxton. "High-precision sampling for Brillouin-zone integration in metals". In: *Physical Review B* 40.6 (1989), p. 3616.

[MR23] D. Monaco and T. Roussigné. "Topology vs localization in synthetic dimensions". In: *Journal of Mathematical Physics* 64.1 (2023), p. 011902.

[MP76] H. Monkhorst and J. Pack. "Special points for Brillouin-zone integrations". In: *Physical Review B* 13.12 (1976), p. 5188.

[Nen83] G. Nenciu. "Existence of the exponentially localised Wannier functions". In: *Communications in Mathematical Physics* 91.1 (1983), pp. 81–85.

[Ove60] A. Overhauser. "Giant spin density waves". In: *Physical Review Letters* 4.9 (1960), p. 462.

[Ove62] A. Overhauser. "Spin density waves in an electron gas". In: *Physical Review* 128.3 (1962), p. 1437.

[Ove68] A. Overhauser. "Exchange and correlation instabilities of simple metals". In: *Physical Review* 167.3 (1968), p. 691.

[Pan07] G. Panati. "Triviality of Bloch and Bloch-Dirac bundles". In: *Annales Henri Poincaré* 8.5 (2007), pp. 995–1011.

[Pei96] R. E. Peierls. *Quantum theory of solids*. Clarendon Press, 1996.

[Phi96] J. Phillips. "Self-adjoint Fredholm operators and spectral flow". In: *Canadian Mathematical Bulletin* 39.4 (1996), pp. 460–467.

[PS16] E. Prodan and H. Schulz-Baldes. *Bulk and Boundary Invariants for Complex Topological Insulators: From K-Theory to Physics*. Mathematical Physics Studies. Springer Cham, 2016.

[RS79] M. Reed and B. Simon. *Methods of Modern Mathematical Physics, Vol. III: Scattering Theory*. Elsevier, 1979.

[Roz72] G. V. Rozenblum. "Distribution of the discrete spectrum of singular differential operators". In: *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika* 202 (1972), pp. 1012–1015.

[Ryu+10] S. Ryu, A. P. Schnyder, A. Furusaki, and A. W. W. Ludwig. "Topological insulators and superconductors: tenfold way and dimensional hierarchy". In: *New Journal of Physics* 12.6 (2010), p. 065010.

[SKR00] H. Schulz-Baldes, J. Kellendonk, and T. Richter. "Simultaneous quantization of edge and bulk Hall conductivity". In: *Journal of Physics A* 33.2 (2000), p. L27.

[SMV01] I. Souza, N. Marzari, and D. Vanderbilt. "Maximally localized Wannier functions for entangled energy bands". In: *Physical Review B* 65.3 (2001), p. 035109.

[Tho83] D. Thouless. "Quantization of particle transport". In: *Physical Review B* 27.10 (1983), p. 6083.

[Tho+82] D. Thouless, M. Kohmoto, M. Nightingale, and M. den Nijs. "Quantized Hall conductance in a two-dimensional periodic potential". In: *Physical Review Letters* 49.6 (1982), p. 405.

[Wei96] T. Weidl. "On the Lieb-Thirring constants $L_{\gamma,1}$ for $\gamma \geq 1/2$". In: *Communications in Mathematical Physics* 178.1 (1996), pp. 135–146.

[Wig38] E. Wigner. "Effects of the electron interaction on the energy levels of electrons in metals". In: *Transactions of the Faraday Society* 34 (1938), pp. 678–685.

[Yat+07] J. Yates, X. Wang, D. Vanderbilt, and I. Souza. "Spectral and Fermi surface properties from Wannier interpolation". In: *Physical Review B* 75.19 (2007), p. 195121.

[Zah05] D. Zaharioudakis. "Quadratic and cubic tetrahedron methods for Brillouin zone integration". In: *Computer Physics Communications* 167.2 (2005), pp. 85–89.

[ZC08] S. Zhang and D. Ceperley. "Hartree-Fock ground state of the three-dimensional electron gas". In: *Physical Review Letters* 100.23 (2008), p. 236404.

Résumé

Ce mémoire d'HDR comporte quatre parties :

- Nous faisons le lien entre l'existence des fonctions de Wannier en matière condensée, et l'existence d'homotopies pour une famille de projecteurs.
- Nous étudions l'apparition de mode de bords lorsqu'un matériau est coupé (systèmes semi-périodiques), et donnons un cadre général pour l'étude du spectre des opérateurs correspondants.
- Nous calculons quelques propriétés du diagramme de phase du gaz de fermions, dans l'approximation de Hartree-Fock.
- Nous étudions les inégalités de Lieb-Thirring de rang fini, et montrons un phénomène de cristallisation pour ses minimiseurs.

Abstract

This Habilitation consists in four parts:

- We link the existence of Wannier functions in condensed matter with the existence of homotopies for a family of projectors.
- We consider the emergence of edge modes when one cuts a material (half-periodic systems), and we give a general framework to study the spectrum of such operators.
- We describe some properties of the phase diagram of the fermionic gas, in the Hartree-Fock approximation.
- We study finite rank Lieb-Thirring inequalities, and exhibit a cristallisation phenomenon for its minimisers.

Mots Clés

Analyse non linéaire, théorie spectrale, inégalités fonctionnelles, physique mathématiques, matière condensée, systèmes fermioniques

Keywords

Analyse non linéaire, théorie spectrale, inégalités fonctionnelles, physique mathématiques, matière condensée, systèmes fermioniques