

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

---

# The Phase Dilemma and its Ramifications for Computational Complexity in Functional Theory

Lukas Kienesberger

---

Submitted for the degree of  
*Master of Science in Theoretical and Mathematical Physics*  
on October 25, 2023



Supervisor: Dr. Christian Schilling (Ludwig-Maximilians-Universität München)

Second Referee: Prof. Robert van Leeuwen (University of Jyväskylä)

Exam Date: October 26, 2023



# Abstract

Functional-based variational methods, such as density functional theory (DFT), have enjoyed widespread success in quantum many-body physics due to the balance they offer between accuracy and low computational cost. Reduced density matrix functional theory (RDMFT) has recently gained interest because of its ability to capture signatures of strong correlation that DFT cannot. The quest to understand and improve RDMFT has also led to explorations of other topics of intrinsic interest, such as the quantum marginal problem. It has also significantly benefited from advances in convex geometry and quantum information theory.

In this thesis, we apply the RDMFT formalism to an exactly solvable few-body system, namely the two-fermion Hubbard model in one dimension. On the one hand, this setting offers the opportunity to gain exact analytical insights about the working mechanism of RDMFT. On the other hand, it maintains a measure of extensivity, namely the chain length, to assess the functional's behavior as the Hilbert space dimension grows. Strikingly, we find that the analytical structure of the functional may render RDMFT numerically unfeasible even in this few-body setting. In particular, computing the functional is NP-complete, despite the dimension of the two-fermion Hilbert space being only quadratic in the chain length. After demonstrating this, we generalize these results to longer-range interactions between two fermions and find evidence that exponential complexity persists for any finite-range interaction. Finally, we provide numerical evidence that this is still true for infinite-range interactions.



# Contents

<b>1</b>	<b>Introduction</b>	<b>7</b>
1.1	Outline	8
<b>2</b>	<b>Background on DFT and RDMFT</b>	<b>9</b>
2.1	Density Functional Theory (DFT)	9
2.2	One-Particle Reduced Density Matrix	11
2.3	One-Body Operators	12
2.4	Symmetries of the 1RDM	12
2.5	Reduced Density Matrix Functional Theory (RDMFT)	13
2.6	Levy-Lieb Constrained Search	14
2.7	Legendre-Fenchel Transforms and Convex Hull	15
2.7.1	Convexity and $\hat{h}$ -representability	16
2.8	Mixed States and Valone Functional	17
2.8.1	Variational Computation of $E_0$	18
2.9	Degeneracy and Strict Convexity	18
2.10	Simplifications from Symmetries	19
<b>3</b>	<b>Background on Complexity Theory</b>	<b>21</b>
3.0.1	3-Satisfiability (3SAT)	21
3.0.2	2-Partition	22
3.0.3	Subset Sum	22
3.1	P and NP	22
3.1.1	The Class P	22
3.1.2	The Class NP	23
3.2	NP-hardness	23
3.2.1	Pseudopolynomial Algorithm for Subset-Sum	23
3.2.2	Weak NP-completeness	24
3.3	Optimization Problems	24
3.3.1	Non-Integer Number Problems	25
<b>4</b>	<b>Hubbard Model with two Fermions</b>	<b>27</b>
4.0.1	Hamiltonian	27
4.1	Levy Functional	28
4.2	Finding the Ground State Energy	30
4.2.1	Normalization	31
4.2.2	Exchange Force	32
4.3	Ground State Degeneracy	33
4.3.1	Degeneracy Region	33
4.3.2	Nondegenerate Regions $S_\alpha$	35
4.4	Convexity	35
4.5	Working with Real Coefficients	38

<b>5</b>	<b>Tweaking the Interaction</b>	<b>41</b>
5.1	Degeneracy Structure for $D = 3$	41
5.2	Functional for $D = 3$	42
5.3	Analytical Description of $\mathcal{R}$	47
5.3.1	Two-Fold Degeneracies	47
<b>6</b>	<b>Cells and Complexity of <math>\tilde{\mathcal{F}}</math></b>	<b>51</b>
6.0.1	Chapter Outline:	51
6.0.2	Motivation and Setup	51
6.1	Number of $\sigma$ -regions (Hubbard Interaction)	52
6.1.1	Error Bounds for Approximations	53
6.1.2	Monte Carlo Sampling	54
6.2	Complexity (Hubbard Interaction)	55
6.2.1	NP-Completeness of the Phase Dilemma	56
6.3	Permutation Symmetry (Hubbard Interaction)	57
6.4	Connectedness (Hubbard Interaction)	58
6.5	Number of $\sigma$ -regions (2 Fermions, Arbitrary Finite-Range Interaction)	60
6.5.1	Hyperplane arrangements	64
6.5.2	Lower bound on $ \Sigma_{\tilde{W}}^0 $	65
6.5.3	Lower bound on $ \Sigma_{\tilde{W}} $	65
6.5.4	Numerical Analysis	67
<b>7</b>	<b>Summary and Conclusion</b>	<b>73</b>
<b>A</b>	<b>Number of Sectors of Hyperplane Arrangements</b>	<b>75</b>
<b>B</b>	<b>Density of <math>\mathcal{W}_I^r</math></b>	<b>77</b>
<b>C</b>	<b>Density of <math>\mathcal{W}_{\tilde{U}}^r</math></b>	<b>79</b>

# Chapter 1

## Introduction

Finding the ground state of a quantum-mechanical system described by a Hamiltonian  $\hat{H}$  is a ubiquitous challenge in solid-state physics, quantum chemistry, and beyond. When attempting to understand large quantum systems, knowledge of the ground state's salient properties can elucidate the structure of chemical compounds and provide a foundation for studying excitations and dynamics. Moreover, the ground state can contain signature features of strong correlations and long-range entanglement [1]. On the other hand, determining the entire spectrum of a many-body Hamiltonian is unfeasible because the Hilbert space dimension increases exponentially with the system size. For the same reason, knowledge of the full spectrum would yield superfluous information not needed to characterize key observables of interest. Determining the ground state nevertheless requires determining it in the exponentially large Hilbert space, itself a daunting task. Realistic systems amenable to wave function methods are few and far between, and more economical approaches are needed in most scenarios. Density-matrix renormalization group (DMRG) methods [1, 2] construct trial wave functions out of local tensors and iteratively optimize these local tensors to approach the ground state. Dynamical mean-field theory [3, 4] develops a local theory of a single lattice site, self-consistently treating the remaining system as a bath. Density functional theory [5, 6] treats the local particle density as the central object and aims to find the ground state density rather than the ground state wave function.

This thesis focuses on reduced density matrix functional theory, an extension of DFT that will be explained in detail in section 2. Let us, for now, outline DFT's principal appeal and the shortcomings RDMFT aims to alleviate. DFT works well for systems where quantities of interest can be expressed in terms of the total ground state energy and the local density of particles in the ground state. In computing those quantities, one bypasses explicit wave function representations in favor of local particle densities. Hohenberg and Kohn [5] showed that the particle density is, in principle, sufficient to serve as the primary variable to represent the ground state. Their pivotal result found widespread use when coupled with the efficient approach pioneered by Kohn and Sham [6], who provided a computationally feasible approximation mapping interacting systems to non-interacting systems. Together with such methods, the DFT framework offers an excellent balance of accuracy and computational feasibility for many quantum chemistry and solid-state physics problems. There are, however, systems where this approximation is insufficiently accurate. In strongly correlated systems, the ground state is often highly entangled and cannot be represented as a product state (or Slater determinant for fermionic systems), precluding an accurate representation of the true ground state by a non-interacting state. More generally, DFT cannot easily capture strong correlations because it can only be used to probe energies and local densities, revealing little to no information about entanglement and correlation.

Reduced density matrix functional theory (RDMFT) generalizes DFT and works with the 1-particle reduced density matrix (1RDM) as the central quantity. Being a single-particle quantum state, the 1RDM can be transformed between different basis representations. The position representation recovers the ground state particle density from the ground state 1RDM's diagonal. On the other hand, being a reduced density matrix, one can analyze the 1RDM's spectrum to learn about entanglement in the ground state wave function. In particular, fractional occupation numbers in the 1RDM's eigenbasis are a signature of strong correlations.

DFT and RDMFT work in bosonic, fermionic, and distinguishable subsystems. Being initially focused on quantum chemistry and solid-state physics applications, most of the existing literature aims at tackling fermionic systems. More recently, RDMFT has been extended to bosonic systems and used to study Bose-Einstein condensation [7, 8]. The two methods equally apply to continuum and lattice systems, the main historical focus being continuous electron systems. This thesis's primary goal is to investigate the analytical

structure of the RDMFT functional (to be defined in section 2) and relate it to the method's computational complexity. To this end, we study fermions on a lattice. Working with a lattice system allows better control over the number of degrees of freedom. The choice of fermionic statistics is less motivated by practical considerations and more by the prevalence of fermionic systems in the existing DFT and RDMFT literature.

## 1.1 Outline

The outline of this thesis is as follows: In chapter 2, the mathematical framework of DFT will be explained in detail and then extended to RDMFT. The thesis focuses on characterizing exact functionals, so prominent approximation methods such as Kohn-Sham will not be discussed. Chapter 3 provides an elementary and self-contained introduction to computational complexity theory. In principle, all the information in the first two chapters can also be found in most textbooks, reviews, and theses on the topics concerned. (See [7, 9], for example) Chapter 4 then applies RDMFT to a system of two interacting fermions on a finite one-dimensional lattice. This relatively simple and exactly solvable system allows us to compute the RDMFT functional without resorting to approximations and fully characterize its properties. All of chapter 4 assumes that the fermions interact via the well-known Hubbard on-site interaction, and extensions to longer-range interactions, as well as more general considerations about the structure of the functional, will be discussed in chapter 5. Chapter 6 presents the main results gathered in this thesis. Therein, we provide a rigorous lower bound on the complexity of the functional's analytical structure. We link the complexity to the interaction range and find that, surprisingly, the complexity bound relaxes for longer-range interactions. We also characterize the symmetries of the functional under transformations of its domain.

# Chapter 2

## Background on DFT and RDMFT

This chapter reviews the elements of density functional theory (DFT) and reduced density matrix functional theory (RDMFT). We begin by introducing DFT, including its mathematical underpinnings, in section 2.1. In section 2.5, we introduce RDMFT, building on the intuition of DFT. In the intervening sections, we define one-particle reduced density matrices, one-body operators, and their symmetry aspects. This serves to lay the groundwork for a concise introduction to RDMFT. Finally, in sections 2.6-2.10, we elaborate on the properties that elevate RDMFT above DFT and explain its additional subtleties.

Throughout this text, we will be considering a one-particle Hilbert space  $\mathcal{H}_1$ , and the fermionic  $N$ -particle Hilbert space  $\mathcal{H}_N = \mathcal{H}_1^{\wedge N}$ . When referring to an unspecified, generic basis for  $\mathcal{H}_1$ , we denote the basis  $\{|\alpha\rangle\}_\alpha$ , and  $\alpha$  may or may not be finite or countable. All concrete examples, however, will have finite dimension, so the basis will be of the form  $\{|\alpha\rangle\}_{\alpha=1}^L$  and the dimension of  $\mathcal{H}_N$  will be  $D = \binom{L}{N}$ . The basis vectors of the (non-symmetrized) tensor product space  $\mathcal{H}_1^{\otimes N}$  will be denoted  $|\alpha_1\rangle \otimes \cdots \otimes |\alpha_N\rangle$ , and the corresponding antisymmetrized Slater determinants are written as  $|\alpha_1, \dots, \alpha_N\rangle \in \mathcal{H}_1^{\wedge N}$ . Density functional theory and reduced density matrix functional theory, as outlined below, can be equally well applied to bosonic systems [7, 10–12], but we do not consider those here.

### 2.1 Density Functional Theory (DFT)

DFT established itself as a workhorse for solving the ground state problem in the wake of Hohenberg and Kohn's [5] and Kohn and Sham's [6] seminal works. DFT is concerned with Hamiltonians of the form  $\hat{H} = \hat{W} + \hat{t} + \hat{v}$ , consisting of a kinetic energy term  $\hat{t}$ , an interaction term  $\hat{W}$  and a single-particle potential operator  $\hat{v}$ . The interaction  $\hat{W}$  is typically a two-body operator but may, in principle, involve any number of particles. Common examples are the Coulomb interaction  $\hat{W}|\mathbf{r}, \mathbf{r}'\rangle \propto |\mathbf{r} - \mathbf{r}'|^{-1}|\mathbf{r}, \mathbf{r}'\rangle$  and the Hubbard on-site interaction to be discussed in section 4. Let  $\{|\alpha\rangle\}_{\alpha=1}^L$  be a basis for the (first-quantized) single-particle Hilbert space, which we assume throughout to be finite-dimensional for concreteness. The single-particle potential operator is restricted to having the form

$$\hat{v} = \sum_{\alpha} v_{\alpha} |\alpha\rangle\langle\alpha|; \quad (2.1)$$

If  $|\alpha\rangle = |x\rangle$  is the position basis, then  $v_{\alpha} = v_x$  is indeed a local potential.

Given  $\hat{W}$  and  $\hat{v}$ , we denote  $|\psi\rangle$  the ground state of the Hamiltonian  $\hat{H}$ . We call  $\rho_{\alpha} = \langle\psi|\hat{n}_{\alpha}|\psi\rangle$  the *ground state density*, with  $\hat{n}_{\alpha} = \hat{a}_{\alpha}^{\dagger}\hat{a}_{\alpha}$  the number operator. Herein, we only consider systems with a fixed particle number  $N$ , so  $|\psi\rangle$  is a bosonic or fermionic  $N$ -particle wave function. The interaction  $\hat{W}$  is fixed in DFT, while the potential  $\hat{v}$  can be adjusted at will. Schematically, this logical chain may be summarized as (ignoring degeneracies)

$$\hat{v} \longmapsto \hat{H} \longmapsto |\psi\rangle \longmapsto \rho, \quad (2.2)$$

with each quantity obtained upon specifying the preceding one.

**Theorem (Hohenberg-Kohn):** Let  $v_{\alpha}$  and  $v'_{\alpha}$  be two potentials. Define the Hamiltonians  $\hat{H} = \hat{W} + \hat{v}$  and  $\hat{H}' = \hat{W} + \hat{v}'$ ; denote by  $|\psi\rangle$  and  $|\psi'\rangle$  their respective ground states. Assume in addition that  $|\psi\rangle$  is not a ground state of  $\hat{H}'$  and  $|\psi'\rangle$  is not a ground state of  $\hat{H}$ . Then  $\rho \neq \rho'$ .

*Proof:* The Hohenberg-Kohn theorem can be proved by contradiction. Suppose that two potentials  $v_{\alpha}, v'_{\alpha}$ ,

with  $v_\alpha - v'_\alpha \neq \text{const.}$ , give rise to the same density  $\rho_\alpha = \rho'_\alpha$ . By the assumption that  $|\psi\rangle$  is not a ground state of  $\hat{H}'$ , the ground states satisfy the inequality

$$\langle \psi' | \hat{H}' | \psi' \rangle < \langle \psi | \hat{H}' | \psi \rangle = \langle \psi | \hat{H} | \psi \rangle + \sum_{\alpha} \rho_{\alpha} (v'_{\alpha} - v_{\alpha}). \quad (2.3)$$

Swapping the primed and unprimed quantities, we obtain

$$\langle \psi | \hat{H} | \psi \rangle < \langle \psi' | \hat{H} | \psi' \rangle = \langle \psi' | \hat{H}' | \psi' \rangle + \sum_{\alpha} \rho'_{\alpha} (v_{\alpha} - v'_{\alpha}). \quad (2.4)$$

Adding (2.3) and (2.4) and using  $\rho = \rho'$ , we have

$$\langle \psi' | \hat{H}' | \psi' \rangle + \langle \psi | \hat{H} | \psi \rangle < \langle \psi | \hat{H} | \psi \rangle + \langle \psi' | \hat{H}' | \psi' \rangle, \quad (2.5)$$

a contradiction. We conclude that  $\rho \neq \rho'$ .

In the literature, the condition that  $\hat{H} = \hat{W} + \hat{t} + \hat{v}$  and  $\hat{H}' = \hat{W} + \hat{t} + \hat{v}'$  do not share a ground state is often ignored. However, this condition is important since its failure can lead to violations of the HK theorem [13]. The condition can be shown to hold when  $\alpha \equiv x \in \mathbb{R}$  is continuous and the ground state wave function  $\psi_0(x_1, \dots, x_N)$  is analytic. To show this, assume that  $|\psi\rangle$  is a ground state of both  $\hat{H}$  and  $\hat{H}'$ . Then clearly,

$$(\hat{v} - \hat{v}')|\psi\rangle = (E_0 - E'_0)|\psi\rangle. \quad (2.6)$$

Multiplying  $\langle \alpha_1, \dots, \alpha_N |$  into the above, and using (2.1), we obtain

$$\sum_{j=1}^N (v_{\alpha_j} - v'_{\alpha_j}) \langle \alpha_1, \dots, \alpha_N | \psi \rangle = (E_0 - E'_0) \langle \alpha_1, \dots, \alpha_N | \psi \rangle. \quad (2.7)$$

The expression  $\sum_{j=1}^N (v_{\alpha_j} - v'_{\alpha_j})$  can be viewed as a local potential on  $\mathbb{C}^N \simeq \mathcal{H}_1^{\otimes N}$ , and eq. (2.7) states that  $\langle \alpha_1, \dots, \alpha_N | \psi \rangle$  is an eigenfunction under multiplication by  $\sum_{j=1}^N (v_{\alpha_j} - v'_{\alpha_j})$ . Assume now that  $\hat{v} - \hat{v}' \neq \text{const}$  so that  $v_{\alpha} - v'_{\alpha} \neq v_{\tilde{\alpha}} - v'_{\tilde{\alpha}}$  for some  $\alpha, \tilde{\alpha}$ . Now we apply (2.7) to two basis states that differ only in the first entry:

$$\left[ v_{\alpha} - v'_{\alpha} + \sum_{j=2}^N (v_{\alpha_j} - v'_{\alpha_j}) \right] \langle \alpha, \dots, \alpha_N | \psi \rangle = (E_0 - E'_0) \langle \alpha, \dots, \alpha_N | \psi \rangle. \quad (2.8)$$

$$\left[ v_{\tilde{\alpha}} - v'_{\tilde{\alpha}} + \sum_{j=2}^N (v_{\alpha_j} - v'_{\alpha_j}) \right] \langle \tilde{\alpha}, \dots, \alpha_N | \psi \rangle = (E_0 - E'_0) \langle \tilde{\alpha}, \dots, \alpha_N | \psi \rangle. \quad (2.9)$$

If  $\langle \tilde{\alpha}, \dots, \alpha_N | \psi \rangle = 0$  and  $\langle \alpha, \dots, \alpha_N | \psi \rangle = 0$  are both finite, we can cancel them in both equations and conclude that the brackets appearing in (2.8) and (2.9) are equal, which cannot be as we assume  $v_{\alpha} - v'_{\alpha} \neq v_{\tilde{\alpha}} - v'_{\tilde{\alpha}}$ . Therefore, at least one of the matrix elements must vanish, a conclusion valid for arbitrary  $\alpha_2, \dots, \alpha_N$ , and  $\alpha, \tilde{\alpha}$  as long as  $v_{\alpha} - v'_{\alpha} \neq v_{\tilde{\alpha}} - v'_{\tilde{\alpha}}$ . However, the basis elements  $|\alpha\rangle, |\tilde{\alpha}\rangle$  are arbitrary, and applying the argument to all possible pairs  $(\alpha, \tilde{\alpha})$  leads to conclude that  $\langle \alpha, \dots, \alpha_N | \psi \rangle \neq 0$  only if  $\alpha \in U \subset \{1, \dots, L\}$ , where  $U$  is some subset over which  $v_{\alpha} - v'_{\alpha}$  is constant (such that the two brackets are identical if  $\alpha, \tilde{\alpha}$  are sampled from this subset, and otherwise one of the matrix elements in (2.8)-(2.9) is zero, so the pair of equations is never violated). The conclusion that the wave function is zero wherever  $v(x) - v'(x)$  is not constant then implies that  $\psi(x, \dots, x_N)$  vanishes for all  $x$ , which by antisymmetry implies it is identically zero, and this cannot be. In this case, we are left to conclude that  $|\psi\rangle$  cannot be an eigenstate of  $\hat{H}'$ , so the HK theorem must always hold. In the non-degenerate case, this can be concisely summarized by the relation

$$\hat{v} \xrightarrow{\text{one-to-one}} |\psi\rangle \xleftarrow{\text{one-to-one}} \rho. \quad (2.10)$$

In discrete systems, the above reasoning does not apply because there is no concept of an analytic function on a discrete set. If  $\hat{v} - \hat{v}'$  is constant over an extended domain, then  $\hat{H}$  and  $\hat{H}'$  may share a ground state. As a result, DFT on finite lattice systems is richer than DFT in the continuum, as shown in [14, 15]. For our purposes, it suffices to state that the loophole weakens the first HK theorem to the weaker version stated originally, and the relation (2.10) does not, in general, hold. Nevertheless, the weaker version of the theorem is sufficient as a foundation for density functional theory. The stronger version also doesn't generalize to nonlocal potentials,

which will be explained in section 2.5. Let us briefly describe the construction of DFT using the weaker version of the first HK theorem, as stated earlier. For this purpose, we invert the theorem statement as follows: Assume that  $v_\alpha - v'_\alpha \neq \text{const.}$  and let  $|\psi\rangle$  and  $|\psi'\rangle$  be the respective ground states. The equality  $\rho = \rho'$  holds only if  $|\psi'\rangle$  is a ground state of  $\hat{H}$  and  $|\psi\rangle$  is a ground state of  $\hat{H}'$ . Put differently,  $\Psi = \Psi'$ , where  $\Psi, \Psi' \subset \mathcal{H}_N$  are the ground state manifolds of  $\hat{H}, \hat{H}'$ . Hence, even though the map  $\hat{v} \mapsto \rho[\hat{v}]$  may not be injective, the theorem shows that  $\Psi$  can be uniquely determined from  $\rho$ . This permits defining the *Hohenberg-Kohn functional*

$$\mathcal{F}^{(\text{HK})}[\rho] = \langle \psi | \hat{W} | \psi \rangle; \quad |\psi\rangle \in \Psi[\rho]. \quad (2.11)$$

To show that this function is well-defined, we need to show that  $\langle \psi | \hat{W} | \psi \rangle = \langle \psi' | \hat{W} | \psi' \rangle$  for  $|\psi\rangle, |\psi'\rangle \in \Psi$ . This is easy to see since both wave functions have the same density  $\rho$  and both belong to the ground state manifold of  $\hat{H}$ :

$$\langle \psi | \hat{W} | \psi \rangle = \langle \psi | \hat{H} | \psi \rangle - \sum_{\alpha} \rho_{\alpha} v_{\alpha} = \langle \psi' | \hat{H} | \psi' \rangle - \sum_{\alpha} \rho_{\alpha} v_{\alpha} = \langle \psi' | \hat{W} | \psi' \rangle. \quad (2.12)$$

Note that the domain of definition HK functional (2.11) is not immediately clear. The Hohenberg-Kohn theorem assumes that  $\rho$  is a density corresponding to the ground state of a Hamiltonian  $\hat{H} = \hat{v} + \hat{W}$ . For a given interaction  $\hat{W}$ , it is not obvious whether a local potential  $v_{\alpha}$  exists such that  $\rho$  is the ground state density of  $\hat{H}$ . If there exists such a  $v_{\alpha}$ , the density  $\rho$  is termed *v-representable*. The domain of definition of  $\mathcal{F}^{(\text{HK})}$  is the set of *v-representable* densities.

The HK functional allows to write the ground state energy as

$$E_0[\hat{v}] = \min_{\rho} \left[ \mathcal{F}^{(\text{HK})}[\rho] + \sum_{\alpha} \rho_{\alpha} v_{\alpha} \right]. \quad (2.13)$$

(The minimization runs over the set of  $\hat{v}$ -representable densities.) To see why this is true, one need only look at (2.11):  $\mathcal{F}^{(\text{HK})}$  first assigns to  $\hat{\rho}$  the quantity  $\langle \psi | \hat{W} | \psi \rangle$ , where  $|\psi\rangle$  fulfills two criteria: It is the ground state of *some* Hamiltonian, and its particle density is  $\rho$ . Therefore, the quantity inside the minimization is just  $\langle \psi | \hat{H} + \hat{v} | \psi \rangle$ . It attains its minimum value  $E_0$  when  $|\psi\rangle$  is the ground state of  $\hat{H}$ . Eq. (2.13) is known as the second Hohenberg-Kohn theorem.

## 2.2 One-Particle Reduced Density Matrix

We now venture into reduced density matrix functional theory, the generalization of DFT alluded to in the introduction and explained in full in section 2.5. Before doing so, we must define the pivotal variable that will later serve as the argument of the RDMFT functional. This variable is called the one-particle reduced density matrix.

Consider an arbitrary (anti)symmetric  $N$ -particle wave function  $|\psi\rangle$ , and the corresponding pure  $N$ -particle density matrix  $\hat{\gamma} = |\psi\rangle\langle\psi|$ . Let us define an operator  $\hat{\gamma}$  on the single-particle Hilbert space via

$$\begin{aligned} \langle \alpha | \hat{\gamma} | \alpha' \rangle &= N \sum_{\alpha_1, \dots, \alpha_{N-1}} (\langle \alpha_1 | \otimes \dots \otimes \langle \alpha_{N-1} | \otimes \langle \alpha |) \hat{\Gamma} (| \alpha_1 \rangle \otimes \dots \otimes | \alpha_{N-1} \rangle \otimes | \alpha' \rangle) \\ &= N \sum_{\alpha_1, \dots, \alpha_{N-1}} \psi(\alpha_1, \dots, \alpha_{N-1}, \alpha) \psi(\alpha_1, \dots, \alpha_{N-1}, \alpha')^* \end{aligned} \quad (2.14)$$

In practice,  $\{|\alpha\rangle\}_{\alpha}$  is often the position basis  $|\alpha\rangle = |x\rangle$ , and for continuous systems, the above sum is an integral. However, we will only work with lattice systems and keep the discrete sum notation, bearing in mind that DFT and RDMFT work for both continuous and discrete systems.  $N$  is added as a normalization factor so the identity  $\text{Tr}[\hat{\gamma}] = N$  holds.  $\hat{\gamma}$  is called the *one-body reduced density matrix (1RDM)*. In (2.14), we trace over all except the  $N$ -th particle. This choice is arbitrary, and (2.14) would be the same expression if we chose to retain any of the other  $N$  degrees of freedom instead of the  $N$ -th. This fact owes to the symmetry of the wave function under particle exchange,  $\psi(\alpha_1, \dots, \alpha_j, \dots, \alpha_{N-1}, \alpha) = \pm \psi(\alpha_1, \dots, \alpha, \dots, \alpha_{N-1}, \alpha_j)$ .

From the first line of (2.14), it is clear that  $\hat{\gamma}$  is, up to the normalization factor, obtained by partially tracing over the first  $N-1$  particles, which, due to the wave function (anti)symmetry, is identical to a partial trace over any other subset of  $N-1$  particles. A commonly used shorthand is thus

$$\hat{\gamma} = N \text{Tr}_{N-1}[\hat{\Gamma}]. \quad (2.15)$$

## 2.3 One-Body Operators

In DFT, we separated the Hamiltonian into a local potential, a nonlocal kinetic term, and a many-body interaction operator. The DFT functional  $\mathcal{F}^{(\text{HK})}$  encapsulated information about  $\hat{W} + \hat{t}$ ; with the functional at hand, one can compute the ground state energy for arbitrary  $\hat{v}$ . The RDMFT functional, defined in section 2.5, will only depend on  $\hat{W}$  and allow computation of the ground state energy for arbitrary kinetic Hamiltonians  $\hat{h} \equiv \hat{t} + \hat{v}$ . This section is dedicated to delimiting precisely the conditions that  $\hat{h}$  must fulfill.

Let  $\hat{h}$  be any operator that satisfies the relation

$$(\langle \alpha'_1 | \otimes \cdots \otimes \langle \alpha'_N |) \hat{h} (| \alpha_1 \rangle \otimes \cdots \otimes | \alpha_N \rangle) = \sum_{j=1}^N \langle \alpha'_j | \hat{h} | \alpha_j \rangle \prod_{i \neq j} \langle \alpha'_i | \alpha_i \rangle = \sum_{j=1}^N \langle \alpha'_j | \hat{h} | \alpha_j \rangle \prod_{i \neq j} \delta_{\alpha_i \alpha'_i}. \quad (2.16)$$

Operators of this kind are called *one-body operators*. Let  $|\psi\rangle$  be an  $N$ -body wave function in the (anti)symmetrized subspace, so  $|\psi\rangle$  satisfies

$$\psi(\alpha_1, \dots, \alpha_N) = \langle \alpha_1, \dots, \alpha_N | \psi \rangle = (\langle \alpha_1 | \otimes \cdots \otimes \langle \alpha_N |) |\psi\rangle. \quad (2.17)$$

The expectation value of  $\hat{h}$  is then

$$\begin{aligned} \langle \psi | \hat{h} | \psi \rangle &= \sum_{\alpha_1, \dots, \alpha_N} \sum_{\alpha'_1, \dots, \alpha'_N} \psi(\alpha'_1, \dots, \alpha'_N) \psi^*(\alpha_1, \dots, \alpha_N) \sum_{j=1}^N \langle \alpha'_j | \hat{h} | \alpha_j \rangle \prod_{i \neq j} \delta_{\alpha_i \alpha'_i} \\ &= \sum_j \sum_{\alpha_1, \dots, \alpha_N} \sum_{\alpha'_j} \psi(\alpha_1, \dots, \alpha'_j, \dots, \alpha_N) \psi^*(\alpha_1, \dots, \alpha_j, \dots, \alpha_N) \langle \alpha'_j | \hat{h} | \alpha_j \rangle \\ &= \sum_j \sum_{\alpha_1, \dots, \alpha_N} \sum_{\alpha'_j} \psi(\alpha_1, \dots, \alpha_N, \dots, \alpha'_j) \psi^*(\alpha_1, \dots, \alpha_N, \dots, \alpha'_j) \langle \alpha'_j | \hat{h} | \alpha_j \rangle \\ &= \sum_j \sum_{\alpha_1, \dots, \alpha_N} \sum_{\alpha'_N} \psi(\alpha_1, \dots, \alpha_j, \dots, \alpha'_N) \psi^*(\alpha_1, \dots, \alpha_j, \dots, \alpha'_N) \langle \alpha'_N | \hat{h} | \alpha_N \rangle \\ &= N \sum_{\alpha_1, \dots, \alpha_N} \sum_{\alpha'_N} \psi(\alpha_1, \dots, \alpha'_N) \psi^*(\alpha_1, \dots, \alpha'_N) \langle \alpha'_N | \hat{h} | \alpha_N \rangle \\ &= \sum_{\alpha_N} \sum_{\alpha'_N} \langle \alpha_N | \hat{\gamma} | \alpha'_N \rangle \langle \alpha'_N | \hat{h} | \alpha_N \rangle \\ &= \text{Tr}[\hat{\gamma} \hat{h}]. \end{aligned} \quad (2.18)$$

In the first line, we have used (2.16). We used the wave function's (anti)symmetry to get from the second to the third line. In the fourth line, we relabeled integration variables, and in the last line, we used the definition (2.14) of the 1RDM. (2.18) shows that  $\hat{\gamma}$  contains sufficient information to evaluate the expectation value of any one-body operator  $\hat{h}$ .

## 2.4 Symmetries of the 1RDM

There is one last point that we shall discuss before revealing the machinations of RDMFT. Symmetries are particularly relevant for this thesis, as they are exploited extensively throughout chapter 4. The reader may skip this section for now as it is not required to understand any of the subsequent sections until 2.10.

An essential property of the 1RDM is that it inherits certain symmetries of the underlying many-particle state. In particular, let  $\hat{u}$  be a unitary transformation acting on the Hilbert space  $\mathcal{H}_1$  of a single particle. We can lift  $\hat{u}$  to the  $N$ -particle level by forming the  $N$ -fold tensor product  $\hat{u}^{\otimes N}$ . Now, let  $\hat{\Gamma}$  be any  $N$ -particle ensemble state commensurate with this  $N$ -particle symmetry transformation:  $\hat{\Gamma} = \hat{u}^{\otimes N} \hat{\Gamma} (\hat{u}^\dagger)^{\otimes N}$ . Taking the partial trace of this relation, we obtain,

$$\begin{aligned} \hat{\gamma} &= N \text{Tr}_{N-1}[\hat{u}^{\otimes N} \hat{\Gamma} (\hat{u}^\dagger)^{\otimes N}] = N \text{Tr}_{N-1}[(\hat{I}^{\otimes N-1} \otimes \hat{u}) \hat{\Gamma} (\hat{I}^{\otimes N-1} \otimes \hat{u}^\dagger)] \\ &= \hat{u} \hat{\gamma} \hat{u}^\dagger. \end{aligned} \quad (2.19)$$

In the second step, we have used that if  $\hat{o}_A, \hat{o}_B$  are operators acting on subsystems  $A$  and  $B$ , then

$$\text{Tr}_B[\hat{o}_A \otimes \hat{o}_B \hat{X}] = \text{Tr}_B[\hat{o}_A \otimes \hat{I} \hat{X} \hat{I} \otimes \hat{o}_B]. \quad (2.20)$$

In words, the partial trace is cyclic in operators acting on the traced-out system. In the last equality, we have used

$$\mathrm{Tr}_B[\hat{\rho}_A \otimes \hat{\rho}_B \hat{X}] = \hat{\rho}_A \mathrm{Tr}_B[\hat{I} \otimes \hat{\rho}_B \hat{X}]. \quad (2.21)$$

Both relations can be proved by expressing the operators in a concrete basis. In conclusion, if the  $N$ -particle state is invariant under  $\hat{u}^{\otimes N}$ , then the 1RDM is invariant under  $\hat{U}$ .

Let  $\hat{g}$  be a generator of the unitary symmetry  $\hat{u}$  so that  $\hat{u} = e^{i\hat{g}}$ . Let us extend this symmetry to a one-parameter family  $\hat{u}(t) = e^{i\hat{g}t}$ , and lift it to the  $N$ -particle level to form  $\hat{U}(t) = \hat{u}(t)^{\otimes N}$ . Taking the time derivative, we have

$$\frac{d}{dt}\hat{U}(t) = i \sum_{j=0}^{N-1} (\hat{I}^{\otimes j} \otimes \hat{g} \otimes \hat{I}^{\otimes N-1-j})\hat{U}(t) \equiv i\hat{G}\hat{U}(t). \quad (2.22)$$

The uniqueness of first-order ODE solutions implies that  $\hat{G}$  is the generator of  $\hat{U}$ .

Assume now that  $[\hat{G}, \hat{\Gamma}] = 0$ , implying that  $\hat{U}(t)\hat{\Gamma}\hat{U}^\dagger(t) = \hat{\Gamma}$ . Let us substitute  $\hat{U} = \hat{U}(t)$  in (2.19), and take the time-derivative. This yields  $\hat{\gamma} = i[\hat{g}, \hat{\gamma}]$ . We see that if  $\hat{G}$  is a hermitian symmetry operator obtained by lifting  $\hat{g}$  to the  $N$ -particle level, and  $\hat{G}$  is a symmetry of  $\hat{\Gamma}$ , then  $\hat{g}$  is a symmetry of  $\hat{\gamma}$ . This can have major implications on the form of the matrix  $\langle \alpha | \hat{\gamma} | \alpha' \rangle$ :

Let  $\{\hat{g}_i\}_{i=1}^M$ ,  $M > 0$  be a complete set of hermitian operators. That is to say, they are simultaneously diagonalizable, and if  $|\phi\rangle \in \mathcal{H}$  is an eigenstate of  $\hat{g}_i$  for every  $i$ , then  $|\phi\rangle$  can be uniquely determined from the eigenvalues  $g_i$ . Based on the above discussion, we can deduce the following:

**Lemma 2.4.1:** *Let  $\{\hat{g}_i\}_{i=1}^M$ ,  $M > 0$  be a complete set of hermitian operators, and let  $\hat{\Gamma}$  be an  $N$ -particle state that commutes with all the lifted operators  $\hat{G}_i$ . (This by no means determines  $\hat{\Gamma}$  because the  $\hat{G}_i$  do not form a complete set in the  $N$ -particle Hilbert space.) Then  $\hat{\gamma} = N \mathrm{Tr}_{N-1} \hat{\Gamma}$  is diagonal in the simultaneous eigenbasis of  $\{\hat{g}_i\}_{i=1}^M$ .*

*Proof:* As we have just shown,  $\hat{g}_i$  commutes with  $\hat{\gamma}$  for all  $i$ , so  $\hat{\gamma}$  leaves the eigenspaces of each  $\hat{g}_i$  invariant. Because  $\{\hat{g}_i\}_{i=1}^M$  is a complete set,  $\hat{\gamma}$  must be diagonal in the simultaneous eigenbasis.

## 2.5 Reduced Density Matrix Functional Theory (RDMFT)

Let us recapitulate the general idea of DFT. In DFT, one computes (usually approximately) the Hohenberg-Kohn functional  $\mathcal{F}^{\mathrm{HK}}$  for a Hamiltonian of interest. The Hamiltonian usually has the form  $\hat{W} + \hat{t}$  with  $\hat{W}$  a multi-particle interaction and  $\hat{t}$  a non-local single-particle operator. Once the functional is known, it can be used to find the ground-state energy and density for any Hamiltonian of the form  $\hat{H} = \hat{W} + \hat{t} + \hat{v}$ , where  $\hat{v}$  is an arbitrary local potential.  $\hat{W} + \hat{t}$  is baked into the functional and cannot be varied.

RDMFT [10, 16–38] relaxes the condition that only  $\hat{v}$  can be varied, and allows varying  $\hat{t}$  as well. The combined non-local one-body operator is denoted  $\hat{h} = \hat{t} + \hat{v}$ . As we shall see, construction of the below-defined RDMFT functional allows solving the Hamiltonian  $\hat{H} = \hat{W} + \hat{h}$  in polynomial time for arbitrary one-particle operators  $\hat{h}$  and fixed  $\hat{W}$ . Moreover, it yields the ground-state 1RDM  $\hat{\rho} = N \mathrm{Tr}_{N-1} [|\psi_0\rangle\langle\psi_0|]$  at the end of the computation, rather than a simple density  $\langle \psi_0 | \hat{n}_\alpha | \psi_0 \rangle$ . The eigenvectors of  $\hat{\rho}$  are termed *natural orbitals*, and its eigenvalues *natural occupation numbers*. Specifying this natural one-particle basis yields substantially more information about the ground state than the fixed-basis density.

We now outline the concrete formal construction of RDMFT. The starting point is (2.2), where we replace the density  $\rho$  by the 1RDM  $\hat{\gamma}$ , and the local potential by an arbitrary one-particle operator  $\hat{h}$ :

$$\hat{h} \mapsto \hat{H} \mapsto |\psi\rangle \mapsto \hat{\gamma}. \quad (2.23)$$

This chain still makes sense this way. Rather than  $\hat{H}[\hat{v}] = \hat{v} + \hat{t} + \hat{W}$ , we now define  $\hat{H}[\hat{h}] = \hat{h} + \hat{W}$ . On the other end of the chain, rather than computing  $\langle \psi | \hat{n}_\alpha | \psi \rangle$ , we now opt for  $\gamma_{\alpha\beta} = \langle \psi | \hat{a}_\beta^\dagger \hat{a}_\alpha | \psi \rangle$ . This means that we retain more information (the local density appears on the diagonal of the 1RDM). We also allow for greater freedom, as the class of single-particle operators  $\hat{h}$  includes, but is not limited to, local potentials. Also note that the duality of (2.2) is retained: In both (2.2) and (2.23), the first and last object are elements of the same vector space (with the caveat the  $\hat{\gamma}$  is positive-definite with unit trace, which  $\hat{h}$  need not be). The expectation value of a one-body operator can also be determined solely from the 1RDM using  $\langle \hat{h} \rangle = \mathrm{Tr}[\hat{\gamma}\hat{h}]$ , as explained in section 2.3.

Having set up the new variables, we turn to the generalization of the Hohenberg-Kohn theorem due to Gilbert [39]:

**Theorem (Gilbert):** Let  $\hat{h}$  and  $\hat{h}'$  be one-body operators. Define the Hamiltonians  $\hat{H} = \hat{W} + \hat{h}$  and  $\hat{H}' = \hat{W} + \hat{h}'$ ; denote by  $|\psi\rangle$  and  $|\psi'\rangle$  respective ground states. Assume in addition that  $|\psi\rangle$  is not a ground state of  $\hat{H}'$  and  $|\psi'\rangle$  is not a ground state of  $\hat{H}$ . Then  $\hat{\gamma} \neq \hat{\gamma}'$ .

*Proof:* The proof proceeds, as in the Hohenberg-Kohn theorem, by assuming that the 1RDMs are equal. Because  $|\psi\rangle$  is not a ground state of  $\hat{H}'$ , the inequality

$$\langle \psi' | \hat{H}' | \psi' \rangle < \langle \psi | \hat{H}' | \psi \rangle = \langle \psi | \hat{H} | \psi \rangle + \text{Tr}[\hat{\gamma}(\hat{h}' - \hat{h})]. \quad (2.24)$$

holds. Exchanging primed and unprimed variables, we have

$$\langle \psi | \hat{H} | \psi \rangle < \langle \psi' | \hat{H} | \psi' \rangle = \langle \psi' | \hat{H}' | \psi' \rangle + \text{Tr}[\hat{\gamma}(\hat{h} - \hat{h}')]. \quad (2.25)$$

Adding (2.24) and (2.25) leads to the contradiction

$$\langle \psi' | \hat{H}' | \psi' \rangle + \langle \psi | \hat{H} | \psi \rangle < \langle \psi | \hat{H} | \psi \rangle + \langle \psi' | \hat{H}' | \psi' \rangle, \quad (2.26)$$

proving that  $\hat{\gamma} \neq \hat{\gamma}'$ .  $\square$

As we discussed in concluding section 2.1, the Hohenberg-Kohn theorem establishes that if  $\rho$  is  $v$ -representable, the ground state manifold  $\Psi[\rho] \subset \mathcal{H}_N$  is uniquely determined by  $\rho$ . The Gilbert theorem establishes a similar correspondence. If  $\hat{\gamma}$  is the ground-state 1RDM of a Hamiltonian  $\hat{H} = \hat{h} + \hat{W}$ , we say that  $\hat{\gamma}$  is  $\hat{h}$ -representable; this property holds with respect to a fixed interaction  $\hat{W}$ . The Gilbert theorem implies that if  $\hat{\gamma}$  is  $\hat{h}$ -representable, the ground state manifold  $\Psi[\hat{\gamma}] \subseteq \mathcal{H}_N$  can be uniquely determined from  $\hat{\gamma}$ . This statement leads to the definition of the Gilbert functional

$$\mathcal{F}^{(G)}[\hat{\gamma}] = \langle \psi | \hat{W} | \psi \rangle; \quad |\psi\rangle \in \Psi[\hat{\gamma}]. \quad (2.27)$$

The Gilbert functional is seen to be well-defined using an argument analogous to the one used to define  $\mathcal{F}^{(\text{HK})}$  at the end of section 2.1.

The ground state energy of  $\hat{H}$  can be obtained from the Gilbert functional variationally:

$$E_0(\hat{h}) = \min_{\hat{\gamma}} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(G)}(\hat{\gamma}) \right); \quad (2.28)$$

the minimization is carried out over the set of  $\hat{h}$ -representable 1RDMs. To see why this must be true, note that if  $\hat{\gamma}$  is the 1RDM of the ground state  $|\psi\rangle$ , then  $\text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(G)}(\hat{\gamma}) = \langle \psi | \hat{H} | \psi \rangle$  is the ground state energy. On the other hand, if  $\hat{\gamma}'$  is an arbitrary  $v$ -representable 1RDM, it derives from the ground state  $|\psi'\rangle$  of some Hamiltonian  $\hat{H}'$ . In this case, the definition of the Gilbert functional (2.27) implies  $\text{Tr}[\hat{h}\hat{\gamma}'] + \mathcal{F}^{(G)}(\hat{\gamma}') = \langle \psi' | \hat{H}' | \psi' \rangle \geq \langle \psi | \hat{H} | \psi \rangle$ . Hence  $\hat{\gamma}$  achieves the minimal value of the functional  $\text{Tr}[\hat{h}\cdot] + \mathcal{F}^{(G)}(\cdot)$ , which in turn must be the ground state energy  $E_0$ .

Note that, in (2.28), we have elected to view  $E_0$  as a function of  $\hat{h}$ . This reflects the fact that  $\mathcal{F}^{(G)}$  is defined for a fixed interaction  $\hat{W}$ , but does not refer to  $\hat{h}$ , which only enters through the first term of (2.28). Viewing  $E_0$  as a function of  $\hat{h}$  also provides us with a duality between (2.28) and (2.27), since  $\hat{\gamma}$  and  $\hat{h}$  are matrices of the same dimension. We will explore this duality in more detail in section 2.7 below.

## 2.6 Levy-Lieb Constrained Search

Like the Gilbert functional, the Levy functional is defined on a subset of 1RDMs. It assigns to a 1RDM  $\hat{\gamma}$  the quantity

$$\mathcal{F}^{(L)}[\hat{\gamma}] = \min_{|\psi\rangle \mapsto \hat{\gamma}} \langle \psi | \hat{W} | \psi \rangle \quad (2.29)$$

where the minimization is performed over the set of all pure states  $|\psi\rangle$  that satisfy  $\hat{\gamma} = N\text{Tr}_{N-1}[|\psi\rangle\langle\psi|]$ . This constrained minimization is referred to as *Levy-Lieb constrained search* [40, 41]. The domain of definition of  $\mathcal{F}^{(L)}$  consists of all  $\hat{\gamma}$  where at least one such  $N$ -body pure state exists. Such a 1RDM  $\hat{\gamma}$  is called *pure-state  $N$ -representable*. The set of all pure-state  $N$ -representable 1RDMs is denoted  $\mathcal{P}_N^1$ . This set is invariant under one-body unitary transformations,

$$\hat{u}^\dagger \mathcal{P}_N^1 \hat{u} = \mathcal{P}_N^1, \quad (2.30)$$

so it is fully characterized by the spectra of all  $\hat{\gamma} \in \mathcal{P}_N^1$ :

$$\text{spec}(\mathcal{P}_N^1) = \{\text{spec}(\hat{\gamma}) \mid \hat{\gamma} \in \mathcal{P}_N^1\}. \quad (2.31)$$

Recalling the definition of  $\hat{h}$ -representability, it is clear that any  $\hat{h}$ -representable 1RDM is also pure-state  $N$ -representable, so the domain of  $\mathcal{F}^{(G)}$  is a subset of the domain of  $\mathcal{F}^{(L)}$ . Note that whereas  $\hat{h}$ -representability is defined relative to the interaction  $\hat{W}$ ,  $N$ -representability is independent of any reference Hamiltonian. Similarly to the Gilbert functional, the Levy functional satisfies the property

$$E_0(\hat{h}) = \min_{\hat{\gamma} \in \mathcal{P}_N^1} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(L)}(\hat{\gamma}) \right), \quad (2.32)$$

the verification of which is straightforward:

$$\begin{aligned} \min_{\hat{\gamma} \in \mathcal{P}_N^1} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(L)}(\hat{\gamma}) \right) &= \min_{\hat{\gamma} \in \mathcal{P}_N^1} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \min_{|\psi\rangle \mapsto \hat{\gamma}} \langle \psi | \hat{W} | \psi \rangle \right) = \min_{\hat{\gamma} \in \mathcal{P}_N^1} \min_{|\psi\rangle \mapsto \hat{\gamma}} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \langle \psi | \hat{W} | \psi \rangle \right) \\ &= \min_{|\psi\rangle} \langle \psi | \hat{H} | \psi \rangle = E_0(\hat{h}). \end{aligned} \quad (2.33)$$

The domain of the Gilbert functional consists of all 1RDMs  $\hat{\gamma}$  that correspond to the ground state of some Hamiltonian with interaction  $\hat{W}$ . The Gilbert theorem does not cover other 1RDMs. This implies, in particular, that specifying the domain requires advance knowledge of the Hamiltonian's solution, or features thereof. On the other hand, the Levy functional's domain  $\mathcal{P}_N^1$  encompasses all 1RDMs that are representable by an  $N$ -particle wave function (pure-state  $N$ -representable).  $\mathcal{P}_N^1$  is a superset of the Gilbert functional's domain. More importantly, its definition is independent of any particular Hamiltonian, making the Levy functional less cumbersome and of greater practical use.

## 2.7 Legendre-Fenchel Transforms and Convex Hull

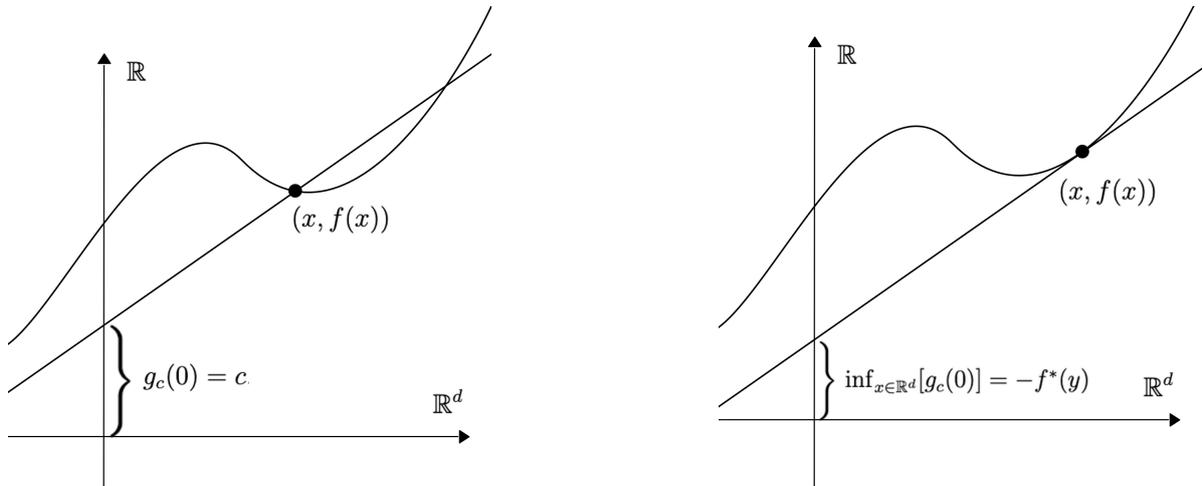


Figure 2.1: Graphical evaluation of Legendre-Fenchel transform

The variational problems (2.28) and (2.32) can be rephrased in terms of the well-known Legendre-Fenchel transform. We recall here its definition and basic properties, following [42]

Given a function  $f$  defined on a subset of  $\mathbb{R}^d$  and taking values in  $\mathbb{R}$ , its *Legendre-Fenchel transform* or *convex conjugate*  $f^* : \mathbb{R}^d \rightarrow \mathbb{R}$  is defined as

$$f^*(y) \equiv \sup_{x \in \mathcal{D}_f} [\langle y, x \rangle - f(x)], \quad (2.34)$$

or equivalently,

$$-f^*(y) = \inf_{x \in \mathcal{D}_f} [f(x) - \langle y, x \rangle]. \quad (2.35)$$

Eq. (2.35) may be understood pictorially. We initially fix both  $x$  and  $y$  and denote by  $G_y$  the graph of  $g_y : x \mapsto \langle y, x \rangle$ .  $G$  is a hyperplane in  $\mathbb{R}^d \oplus \mathbb{R}$ . If we add a constant  $c$ ,

$$g_{y,c} : x \mapsto \langle y, x \rangle + c, \quad (2.36)$$

we obtain a new hyperplane  $G_{y,c}$ . If we select  $c \equiv c_x \equiv f(x) - \langle y, x \rangle$ , then

$$g_{y,c}(0) = c = f(x) - \langle y, x \rangle,$$

as illustrated on the left-hand side of Fig. 2.1. Taking the infimum over  $x$ , we obtain  $\inf_{x \in \mathbb{R}^d} [g_c(0)] = -f^*(y)$ . This is illustrated on the right-hand side of Fig. 2.1.

Eq. (2.34) is not the only way to define the Legendre-Fenchel transform.

**Proposition 2.7.1:**

$$-f^*(y) = \sup_{c \in \mathbb{R}} \{c \mid \langle y, x \rangle + c \leq f(x) \text{ for all } x \in \mathcal{D}f\} \quad (2.37)$$

*Proof:* We need to show two criteria; firstly that for any  $a > -f^*(y)$  there exists  $x$  in the domain of  $f$  (denoted  $\mathcal{D}f$ ) such that  $f(x) < \langle y, x \rangle + c$ ; secondly, that for any  $\epsilon > 0$ , there exists  $c \in (-f^*(y) - \epsilon, -f^*(y)]$  such that  $\langle y, x \rangle + c \leq f(x)$  holds for all  $x \in \mathcal{D}f$ . To show the first criterion, note that (2.35) implies for  $c > -f^*(y)$  the existence of  $x \in \mathcal{D}f$  such that  $f(x) - \langle x, y \rangle < c$ . For the second criterion, we need only pick some arbitrary  $a$  in the interval  $a \in (-f^*(y) - \epsilon, -f^*(y)]$ ; because  $c < -f^*(y)$ , the infimum in (2.35) then implies that  $f(x) - \langle x, y \rangle \geq a$  for all  $x \in \mathcal{D}f$ .  $\square$

Eq. (2.37) may again be understood pictorially from Fig. 2.1. We consider again the family of affine functions  $g_{y,c}$  in (2.36), this time with the criterion that  $g_{y,c}(x) \leq f(x)$  everywhere, and take the supremum over  $c$  to obtain the affine function  $g(x)$  that intersects  $f$  tangentially from below. Finally, we multiply its value at  $x = 0$  by  $-1$  to obtain  $f^*(y)$ .

The definition (2.37) identifies the Legendre-Fenchel transform at  $k$  as the negative of the  $y$ -intercept  $g_{k,c}(0)$  of the affine function  $g_{k,c}(x)$ , with  $c$  set to the largest possible value that prevents  $g_{k,c}(x)$  from ever exceeding  $f(x)$ . This viewpoint is beneficial when considering the double conjugate,

$$\begin{aligned} f^{**}(x) &= \sup_{k \in \mathbb{R}^d} [\langle z, k \rangle - f^*(k)] \\ &= \sup_{k \in \mathbb{R}^d} \left[ \langle z, k \rangle + \sup_{c \in \mathbb{R}} \{c \mid c + \langle x', k \rangle \leq f(x') \text{ for all } x' \in \mathcal{D}f\} \right] \\ &= \sup_{k \in \mathbb{R}^d} \sup_{c \in \mathbb{R}} \{ \langle z, k \rangle + c \mid c + \langle x', k \rangle \leq f(x') \text{ for all } x' \in \mathcal{D}f \}. \end{aligned} \quad (2.38)$$

It is worth taking a moment to interpret the last line. The supremum is that of the expression  $g_{k,c}(x) = \langle x, k \rangle + c$ , and runs over the family of all such affine functions that are nowhere greater than  $f(x)$ . This is nothing other than the lower convex hull of  $f(x)$ , which we denote  $\bar{f}(x)$ : [42]

$$f^{**}(x) = \bar{f}(x). \quad (2.39)$$

In other words, the double Legendre transform of  $f$  is its lower convex envelope. In particular, for functions defined on a non-convex domain  $\mathcal{D}$ , the double Legendre transform extends to the convex hull  $\mathcal{D}$  (and indeed all of  $\mathbb{R}$ , as seen from (2.38)).

Let us return to the ground state energies obtained from the Gilbert and Levy functional via (2.28) and (2.32). Comparing to (2.35), it is clear that

$$E_0(\hat{h}) = -\mathcal{F}^{(G/L)*}(-\hat{h}), \quad (2.40)$$

where the superscript indicates that the identity holds for both functionals. In other words, if we view the ground state energy as a function of the one-body part of the Hamiltonian, then this function is the convex conjugate of the Gilbert and Levy functionals up to sign flips. Moreover, this identity implies via (2.39) that  $\bar{\mathcal{F}}^{(G)}(\hat{\gamma}) = \bar{\mathcal{F}}^{(L)}(\hat{\gamma})$ ; the two functionals have the same lower convex envelope (with suitable extension of the domain to its convex hull).

## 2.7.1 Convexity and $\hat{h}$ -representability

We say that  $f$  is convex at  $x \in \mathcal{D}f$  if there exist  $c \in \mathbb{R}, y \in \mathbb{R}^d$  such that  $g_{y,c}(x) = f(x)$  and  $g_{y,c}(z) \leq f(z)$  for all  $z \in \mathcal{D}f$ . By (2.39), this is equivalent to the criterion  $f(x) = f^{**}(x)$ . Using the identity (2.37), one can further show that if  $\mathcal{D}f$  is compact, then  $f$  is convex at  $x$  if and only if there exists  $y \in \mathbb{R}^d$  such that  $-f^*(y) = f(x) - \langle y, x \rangle$ . Applying this statement to the identity (2.40), we learn that  $\mathcal{F}^{(G/L)}$  is convex at  $\hat{\gamma}$  if and only if there exists a one-body operator  $\hat{h}$  such that  $E_0(\hat{h}) = \mathcal{F}^{(G/L)}(\hat{\gamma}) + \text{Tr}[\hat{\gamma}\hat{h}]$ . By either definition, (2.27) or (2.29), this amounts to  $\hat{\gamma}$  being the 1RDM for a ground state wave function of  $\hat{H} = \hat{W} + \hat{h}$ . On other words,  $\mathcal{F}^{(G/L)}$  is convex at  $\hat{\gamma}$  if and only if  $\hat{\gamma}$  is  $\hat{h}$ -representable.

## 2.8 Mixed States and Valone Functional

On the  $N$ -particle level, we have, so far, exclusively considered pure states  $|\psi\rangle\langle\psi|$ , only working with mixed states on the 1-particle level, where they arise as a consequence of tracing out the other  $N - 1$  particles. Now, we will consider mixed states on the  $N$ -particle level, obtained by forming a convex combination of pure states,

$$\hat{\Gamma} = \sum_{i=1}^m \Gamma_i |\psi_i\rangle\langle\psi_i|, \quad (2.41)$$

with  $m$  an arbitrary positive integer and  $\sum_{i=1}^m \Gamma_i = 1$ . Equivalently, the set of mixed states can be defined as consisting of all positive-semidefinite, hermitian, trace-1 operators on  $\mathcal{H}_N$ .

Up to this point, we have discussed two functionals with argument  $\hat{\gamma}$ : the Gilbert functional in section 2.5 and the Levy functional in 2.6. The third 1RDM function we discuss is the *Valone functional* [43], obtained by relaxing Levy's constrained search to mixed states:

$$\mathcal{F}^{(V)}[\hat{\gamma}] = \min_{\hat{\Gamma} \rightarrow \hat{\gamma}} \text{Tr}[\hat{W}\hat{\Gamma}] \quad (2.42)$$

The minimization runs over all (pure or mixed)  $N$ -particle states  $\hat{\Gamma}$  that satisfy to  $N\text{Tr}_{N-1}\hat{\Gamma} = \hat{\gamma}$ . The domain of  $\mathcal{F}^{(V)}$ , denoted  $\mathcal{E}_N^1$ , consists of all 1RDMs  $\hat{\gamma}$  for which at least one such  $\hat{\Gamma}$  exists. We call  $\mathcal{E}_N^1$  the set of *ensemble  $N$ -representable* 1RDMs.

**Lemma 2.9.1:**  $\mathcal{E}_N^1 = \overline{\mathcal{P}_N^1}$ , where  $\overline{\mathcal{P}_N^1}$  is the convex hull of  $\mathcal{P}_N^1$

*Proof:* Suppose that  $\hat{\gamma}_1 = \text{Tr}_{N-1}[|\psi_1\rangle\langle\psi_1|]$  and  $\hat{\gamma}_2 = \text{Tr}_{N-1}[|\psi_2\rangle\langle\psi_2|]$ . Then for  $0 < \lambda < 1$ ,

$$\lambda\hat{\gamma}_1 + (1 - \lambda)\hat{\gamma}_2 = \text{Tr}_{N-1}[\lambda|\psi_1\rangle\langle\psi_1| + (1 - \lambda)|\psi_2\rangle\langle\psi_2|], \quad (2.43)$$

so  $\lambda\hat{\gamma}_1 + (1 - \lambda)\hat{\gamma}_2 \in \mathcal{E}_N^1$ . Conversely, let  $\hat{\gamma} \in \mathcal{E}_N^1$ , so  $\hat{\gamma} = \text{Tr}_{N-1}\hat{\Gamma}$  for some  $N$ -particle state  $\hat{\Gamma}$ . The latter can be expressed as a convex combination of  $N$ -particle pure states:  $\hat{\Gamma} = \sum_{\alpha} \lambda_{\alpha} |\psi_{\alpha}\rangle\langle\psi_{\alpha}|$  where  $\sum_{\alpha} \lambda_{\alpha} = 1$  and  $0 < \lambda_{\alpha} < 1$ . Taking the partial trace of this relation, we have

$$\hat{\gamma} = \sum_{\alpha} \lambda_{\alpha} \text{Tr}_{N-1}[|\psi_{\alpha}\rangle\langle\psi_{\alpha}|], \quad (2.44)$$

a convex combinations of elements of  $\mathcal{P}_N^1$ , so  $\hat{\gamma} \in \overline{\mathcal{P}_N^1}$ .  $\square$

Once the Valone functional is known, the ground state energy is found analogously to the Levy case:

$$\begin{aligned} \min_{\hat{\gamma} \in \mathcal{E}_N^1} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(V)}(\hat{\gamma}) \right) &= \min_{\hat{\gamma} \in \mathcal{E}_N^1} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \min_{\hat{\Gamma} \rightarrow \hat{\gamma}} \text{Tr}[\hat{W}\hat{\Gamma}] \right) = \min_{\hat{\gamma} \in \mathcal{E}_N^1} \min_{\hat{\Gamma} \rightarrow \hat{\gamma}} \left( \text{Tr}[\hat{h}\hat{\gamma}] + \text{Tr}[\hat{W}\hat{\Gamma}] \right) \\ &= \min_{\hat{\Gamma}} \text{Tr}[\hat{H}\hat{\Gamma}] = E_0(\hat{h}). \end{aligned} \quad (2.45)$$

With a glance at (2.34), we see that the Valone functional's Legendre transform equals the Gilbert and Levy functionals'. The Valone functional, however, carries the additional merit of being convex. To see why, pick  $\hat{\gamma}_1, \hat{\gamma}_2 \in \mathcal{E}_N^1$ , and let  $\hat{\gamma} = \lambda\hat{\gamma}_1 + (1 - \lambda)\hat{\gamma}_2$ . Let  $\hat{\Gamma}_i, i = 1, 2$  be the  $N$ -particle ensemble state that saturates the bound in (2.42), i.e.,  $\hat{\gamma}_i = \text{Tr}_{N-1}\hat{\Gamma}_i$  and  $\min_{\hat{\Gamma} \rightarrow \hat{\gamma}_i} \text{Tr}[\hat{W}\hat{\Gamma}] = \text{Tr}[\hat{W}\hat{\Gamma}_i]$ . Forming the convex combination  $\hat{\Gamma} = \lambda\hat{\Gamma}_1 + (1 - \lambda)\hat{\Gamma}_2$ , it is clear by linearity that  $\text{Tr}[\hat{W}\hat{\Gamma}] = \lambda\mathcal{F}^{(V)}[\hat{\gamma}_1] + (1 - \lambda)\mathcal{F}^{(V)}[\hat{\gamma}_2]$ . Referring again to the definition of the Valone functional,  $\mathcal{F}^{(V)}[\hat{\gamma}] = \min_{\hat{\Gamma} \rightarrow \hat{\gamma}} \text{Tr}[\hat{W}\hat{\Gamma}]$ , and noting that  $\text{Tr}_{N-1}\hat{\Gamma} = \hat{\gamma}$ , so  $\hat{\Gamma}$  is included in the minimization, it is clear that

$$\mathcal{F}^{(V)}[\hat{\gamma}] \leq \text{Tr}[\hat{W}\hat{\Gamma}] = \lambda\mathcal{F}^{(V)}[\hat{\gamma}_1] + (1 - \lambda)\mathcal{F}^{(V)}[\hat{\gamma}_2], \quad (2.46)$$

and thus satisfies the criterion for convexity.

Let us recapitulate. In section 2.7, we showed that two functionals with the same Legendre transform have the same lower convex hull and concluded that the Gilbert and Levy functional have the same convex hull. In this section, we showed that the Valone functional itself is convex, and by (2.45), the Valone functional has the same Legendre transform as the Gilbert and Levy functionals. We can, therefore, conclude (after extending the

respective domains to their convex hull) that the Valone functional is itself the sought-after convex hull of the two previously defined functionals:

$$\overline{\mathcal{F}}^{(G)}(\hat{\gamma}) = \overline{\mathcal{F}}^{(L)}(\hat{\gamma}) = \mathcal{F}^{(V)}(\hat{\gamma}). \quad (2.47)$$

This makes the Valone functional desirable for practical reasons: Since all the functionals in question serve the end of being plugged into the minimizations (2.28), (2.32), (2.45), it is advantageous to have a convex functional at hand, where the occurrence of local minima can be ruled out. On the other hand, however, it is not a priori clear how much the functional domain grows due to the relaxation. An enlarged domain increases the search space, an unwanted side effect. [44]

### 2.8.1 Variational Computation of $E_0$

In this section and henceforth, we adopt the Levy functional as the canonical functional and drop the L superscript:

$$\mathcal{F}(\hat{\gamma}) \equiv \mathcal{F}^{(L)}(\hat{\gamma}). \quad (2.48)$$

If  $\mathcal{F}$  is differentiable, the solution to the variational problem (2.32) may be found using the criterion

$$\frac{\partial \mathcal{F}}{\partial \hat{\gamma}}[\hat{\gamma}_{\min}] = -\hat{h}. \quad (2.49)$$

If we view the functional as a potential defined on  $\mathcal{P}_N^1$ , we can interpret  $\text{grad } \mathcal{F} = \partial \mathcal{F} / \partial \hat{\gamma}$  as a force. In a loose sense, this force originates from the particle-particle interaction, which the functional encodes. It is occasionally termed *exchange force*, as it pushes 1RDM away from the domain boundary [25]. We will see concrete examples of this phenomenon in section 4.2.2. Eq. (2.49) stipulates that at the ground-state 1RDM  $\hat{\gamma}_{\min}$ , the exchange force must be exactly balanced by a force originating from the single-particle Hamiltonian  $\hat{h}$ .

## 2.9 Degeneracy and Strict Convexity

An important link exists between the analytical structure of  $\mathcal{F}^{(V)}$  and the appearance of degenerate ground states. Without making any assumptions on the fixed interaction  $\hat{W}$ , one generically expects that the ground state will be degenerate for some configurations of the one-body Hamiltonian  $\hat{h}$ . So let us assume that for some choice of  $\hat{h}$ , the ground state manifold  $\mathcal{H}_{\text{GS}}$  is spanned by the orthonormal basis  $\{|\psi\rangle_i\}_{i=1}^n$ . The partial trace maps this subspace onto some set of 1RDMs,

$$\text{Tr}_{N-1}[\mathcal{H}_{\text{GS}}] = \mathcal{R} \subset \mathcal{P}_N^1, \quad (2.50)$$

which is, by construction, a subset of the pure-state  $N$ -representable 1RDMs.

**Proposition 2.10.1:** *Consider now two arbitrary ground-state 1RDMs  $\hat{\gamma}_1, \hat{\gamma}_2 \in \mathcal{R}$ . Then*

$$\mathcal{F}^{(V)}(x\hat{\gamma}_1 + (1-x)\hat{\gamma}_2) \geq x\mathcal{F}^{(V)}(\hat{\gamma}_1) + (1-x)\mathcal{F}^{(V)}(\hat{\gamma}_2) \quad (2.51)$$

*Proof:* Assume that

$$\mathcal{F}^{(V)}(\hat{\gamma}) < x\mathcal{F}^{(V)}(\hat{\gamma}_1) + (1-x)\mathcal{F}^{(V)}(\hat{\gamma}_2), \quad (2.52)$$

where  $\hat{\gamma} = x\hat{\gamma}_1 + (1-x)\hat{\gamma}_2$ . Adding  $x\text{Tr}[\hat{\gamma}_1\hat{h}] + (1-x)\text{Tr}[\hat{\gamma}_2\hat{h}] = \text{Tr}[\hat{\gamma}\hat{h}]$  on both sides, we have

$$\mathcal{F}^{(V)}(\hat{\gamma}) + \text{Tr}[\hat{\gamma}\hat{h}] < x \left( \mathcal{F}^{(V)}(\hat{\gamma}_1) + \text{Tr}[\hat{\gamma}_1\hat{h}] \right) + (1-x) \left( \mathcal{F}^{(V)}(\hat{\gamma}_2) + \text{Tr}[\hat{\gamma}_2\hat{h}] \right) \quad (2.53)$$

$$= xE_0(\hat{h}) + (1-x)E_0(\hat{h}) = E_0(\hat{h}). \quad (2.54)$$

To get to the second line, we have used that  $\hat{\gamma}_1, \hat{\gamma}_2 \in \mathcal{R}$  are ground-state 1RDMs. Turning to the left-hand side, there must be an  $N$ -particle state  $\hat{\Gamma}$  such that  $\text{Tr}[\hat{\Gamma}(\hat{W} + \hat{h})] = \mathcal{F}^{(V)}(\hat{\gamma}) + \text{Tr}[\hat{\gamma}\hat{h}]$  by the definition (2.42). Therefore, (2.53) implies  $\text{Tr}[\hat{\Gamma}(\hat{W} + \hat{h})] < E_0(\hat{h})$ , a contradiction. This means the assumption (2.52) must be wrong.  $\square$

The Valone functional  $\mathcal{F}^{(V)}$  is (not strictly) convex everywhere, so the above claim immediately implies that  $\mathcal{F}^{(V)}$  is an affine function on the convex hull  $\overline{\mathcal{R}}$ . Proposition 2.10.1 has a converse:

**Proposition 2.10.2:** If  $\mathcal{F}^{(V)}$  is affine on some set  $\mathcal{R} \in \mathcal{P}_N^1$ , and  $\hat{\gamma} \in \mathcal{R}$  is a ground-state 1RDM of  $\hat{W} + \hat{h}$ , then all  $\hat{\gamma}' \in \mathcal{R}$  are ground-state 1RDMs of  $\hat{W} + \hat{h}$ .

*Proof:* Since  $\mathcal{F}^{(V)}(\hat{\gamma})$  is locally affine, so is  $\text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(V)}(\hat{\gamma})$ . Because  $\hat{\gamma}$  is a ground-state 1RDM, it must be a minimum of  $\text{Tr}[\hat{h}\hat{\gamma}] + \mathcal{F}^{(V)}(\hat{\gamma})$ . Therefore, the latter can only be affine on  $\mathcal{R}$  if it is constant on  $\mathcal{R}$ . But this means that every  $\hat{\gamma}' \in \mathcal{R}$  also minimizes this function and is a ground-state 1RDM.  $\square$

Previously, we had used inequality (2.46) to show that  $\mathcal{F}^{(V)}$  is convex, but so far, we had not established strict convexity. The above two propositions reveal that the regions where  $\mathcal{F}^{(V)}$  is not strictly convex (i.e., affine) can always be traced back to degenerate ground-state manifolds. Note, however, that not every appearance of degeneracy leads to this affine behavior because the region  $\mathcal{R}$  in Proposition 2.10.1 may consist of just a single point.

The sets where  $\mathcal{F}^{(V)}$  is affine can exhibit rich geometry and were recently studied in [15]. We will discuss an elementary example thereof in section 5.3.

## 2.10 Simplifications from Symmetries

Let  $\hat{G}$  be a hermitian operator that commutes with the Hamiltonian:  $[\hat{H}, \hat{G}] = 0$ . We may thus assume that the ground state is an eigenvector of  $\hat{G}$ . Furthermore, assume that  $\hat{G}$  is a one-particle symmetry,

$$\hat{G} = \sum_{j=0}^{N-1} \hat{I}^{\otimes j} \otimes \hat{g} \otimes \hat{I}^{\otimes N-1-j}, \quad (2.55)$$

as is the case for most symmetries encountered in physics. Denote by  $\mathcal{H}_G$  the  $\hat{G}$ -eigenspace corresponding to the eigenvalue  $G \in \text{spec } \hat{G}$ . The ground-state energy can now be found

$$E_0 = \min_{G \in \text{spec}[\hat{G}]} \min_{|\psi\rangle \in \mathcal{H}_G} \langle \psi | \hat{H} | \psi \rangle; \quad (2.56)$$

because we know a priori that the ground state resides in some  $\hat{G}$ -eigenspace, we search for it in each subspace and then minimize over all the subspaces. We now split the second minimization into two more:

$$\begin{aligned} \min_{|\psi\rangle \in \mathcal{H}_G} \langle \psi | \hat{H} | \psi \rangle &= \min_{[\hat{\gamma}, \hat{g}] = 0} \min_{|\psi\rangle \mapsto \hat{\gamma}} \langle \psi | \hat{H} | \psi \rangle = \min_{[\hat{\gamma}, \hat{g}] = 0} \left( \text{Tr}_{N-1}[\hat{h}\hat{\gamma}] + \min_{\mathcal{H}_G \ni |\psi\rangle \mapsto \hat{\gamma}} \langle \psi | \hat{W} | \psi \rangle \right) \\ &\equiv \min_{[\hat{\gamma}, \hat{g}] = 0} \left( \text{Tr}_{N-1}[\hat{h}\hat{\gamma}] + \mathcal{F}_G[\hat{\gamma}] \right) \end{aligned} \quad (2.57)$$

The first minimization runs over all  $\hat{\gamma}$  that commute with  $\hat{g}$ . This restriction makes sense, because any  $|\psi\rangle$  on the left-hand side resides in the symmetric subspace of  $\hat{G}$ , so  $\hat{\gamma} = \text{Tr}_{N-1}[|\psi\rangle\langle\psi|]$  commutes with  $\hat{g}$  by lemma 2.4.1. Therefore, every  $|\psi\rangle$  on the left-hand side also occurs on the right, and the minima are equal.

The functionals  $\mathcal{F}_G$  implicitly depend on  $\hat{W}$ , and we would like to adopt the viewpoint that interaction  $\hat{W}$  is fixed. At the same time, the one-body Hamiltonian  $\hat{h}$  can be adjusted at will, as we did in (2.32). Therefore, a more appropriate starting point is  $[\hat{W}, \hat{G}] = 0$  rather than  $[\hat{H}, \hat{G}] = 0$ . This prescription allows one to define the ground-state energy function

$$E_{0,G}[\hat{h}] = \min_{[\hat{\gamma}, \hat{g}] = 0} \left( \text{Tr}_{N-1}[\hat{h}\hat{\gamma}] + \mathcal{F}_G[\hat{\gamma}] \right) \quad (2.58)$$

for each subspace labeled by  $G$ . The domain of  $E_{0,G}$  consists of all  $\hat{h}$  that satisfy  $[\hat{h}, \hat{G}] = 0$ , ensuring that the total Hamiltonian remains compatible with the symmetry. In particular,  $E_{0,G}$  and  $\mathcal{F}_G$  have the same domain, so we may once again reframe (2.58) as a Legendre transform

$$E_{0,G}(\hat{h}) = -\mathcal{F}_G^*(-\hat{h}). \quad (2.59)$$

As a closing remark, we note that while the most common symmetries are one-particle symmetries, the Hamiltonian may also commute with other operators  $\hat{G}$  that cannot be represented in the form (2.55). In such cases, the symmetry may, of course, still be leveraged by finding the ground state energy and 1RDM for each

symmetry subspace separately. In this case, we defined the set  $\mathcal{P}_{N,G}^1 = \{\hat{\gamma} | \hat{\gamma} = \text{Tr}_{N-1}[|\psi\rangle\langle\psi|], \hat{G}|\psi\rangle = G|\psi\rangle\}$  of 1RDMs that are pure-state  $N$ -representable by wave functions in the  $G$ -subspace. The functional

$$\mathcal{F}_G[\hat{\gamma}] = \min_{\mathcal{H}_G \ni |\psi\rangle \mapsto \hat{\gamma}} \langle \psi | \hat{W} | \psi \rangle \quad (2.60)$$

is defined on  $\mathcal{P}_{N,G}^1$ , and yields the ground state energy as in (2.58), where  $\hat{h}$  is allowed to be any one-body operator that commutes with  $\hat{G}$  in the  $N$ -particle level. We will encounter an example of such a symmetry in the next chapter.

# Chapter 3

## Background on Complexity Theory

Some of the main results of this thesis, to be presented in chapter 6, will deal with the computational complexity of calculating the Levy functional. To properly convey the results therein, we need to introduce a few elementary concepts from complexity theory. In this chapter, we explain the computer science definition of an algorithm and mention some well-known examples. We then introduce the complexity classes P and NP in 3.1. The term NP-hardness will be explained in section 3.2. Up to that point, the main focus will be on decision problems, i.e., problems with a yes/no answer. In section 3.3, the concepts introduced theretofore are expanded to numerical optimization problems, which play a central role in functional theory.

We keep definitions somewhat informal in service of brevity. Our main goal is to avoid later confusion about what can and cannot be described by complexity theory. We mostly follow [45], but all of the material presented below can be found in most introductory textbooks on algorithm theory. Moreover, all the terminology we introduce is standard in the field. (See [45] or [9], for example.)

An *algorithm* is a well-defined computational procedure that deterministically maps an integer  $x$  to another integer  $y$ . This is sufficient to describe the solution of any real-world computational problem since any input (or output) drawn from a finite set can be deterministically encoded as an integer. A problem for which an algorithmic solution exists is called *computable*. A strict requirement for computability is thus that all inputs and outputs be finite. This definition may seem overly restrictive. In particular, the problem of implementing an arbitrary map,

$$f : U \rightarrow \mathbb{R}; \quad U \subset \mathbb{R} \tag{3.1}$$

is not computable, unless  $U$  and  $f(U)$  are known to be is most countable.

A problem with a binary output is called a *decision problem*. Decision problems are frequently encountered in computer science and are extensively used in complexity theory.

The *size* of the input refers not to the magnitude of  $x \in \mathbb{Z}$  but to the number of bits used to encode it. In binary encoding, the input size thus behaves as  $\log_2(|x|)$ . Henceforth, we will always assume that inputs are encoded in binary.

### 3.0.1 3-Satisfiability (3SAT)

As a first example of a computable problem, let us consider the 3-satisfiability problem. An instance of 3SAT consists of a set of logical maps  $\{f_l\}_{l=1}^n$  where  $f_l : \mathbb{Z}_2^N \rightarrow \mathbb{Z}_2$ ,  $N > 0$ . The domain  $\mathbb{Z}_2^N = \{0, 1\}^N$  is the set of  $N$ -component vectors with binary entries. The maps  $f_l$  all have the form

$$f_l(\mathbf{x}) = c_l(x_{i(l)}, x_{j(l)}, x_{k(l)}), \tag{3.2}$$

where  $0 < i(l), j(l), k(l) \leq N$  and  $c_l : \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2 \rightarrow \mathbb{Z}_2$  can be written as a binary clause, e.g.,  $c_l(x, y, z) = \bar{x} \vee (\bar{y} \wedge z)$ . The 3SAT problem asks:

Does there exist an input vector  $\mathbf{x} \in \{\mathbb{Z}_2\}^N$  such that  $f_l(\mathbf{x}) = 1$  for all  $l$ ?

3SAT is a decision problem because the output is binary. Any algorithm that, given any such set of maps  $\{f_l\}_{l=1}^n$ , outputs 1 if such an input string exists (otherwise outputting 0) is a solution of 3SAT.

Note that the 3SAT input, consisting of the maps  $f_l$ , is not naturally given as an integer. It can, however, be represented as one because the set of maps of the form (3.2) is finite. Specifically, the input consists of the maps  $i, j, k : \mathbb{Z}_N \rightarrow \mathbb{Z}_N$  and  $c : \mathbb{Z}_n \rightarrow \mathcal{C}(3)$ , where  $\mathcal{C}(3)$  is the set of tree-variable binary clauses. Each of the

$i, j, k$  requires  $n \log_2(N)$  bits to encode, and  $c$  takes  $8n$  bits. ( $c_l$  must be specified for each  $l \in \mathbb{Z}_n$ , and encoding  $c_l : \mathbb{Z}_2^3 \rightarrow \mathbb{Z}_2$  requires 8 bits since the value taken by the clause must be specified separately for each of the eight possible inputs.) The input size with this encoding is, therefore,  $n(3 \log_2(N) + 8)$ , and when representing the input as an integer, the integer will range from 1 to  $2^{n(3 \log_2(N) + 8)}$ .

### 3.0.2 2-Partition

This is another staple of complexity theory. While 3SAT works with binary inputs and logical clauses, 2-partition is phrased explicitly in terms of integers. The input is a multiset  $S = \{a_j\}_{j \in I}$  with  $a_j \in \mathbb{Z}_N, n > 0$  and  $I \equiv \{1, \dots, n\}$ . The task is as follows:

$$\text{Does there exist a subset } U \subset I \text{ such that } \sum_{j \in U} a_j = \sum_{j \in I \setminus U} a_j?$$

In other words, an algorithm must determine if  $S$  can be partitioned into two subsets that sum to the same integer and output 1 if this is the case. If this is not the case, it must output 0. The input consists of  $n$  elements of  $\mathbb{Z}_N$ ; thus, the input size is  $n \log_2(N)$ .

### 3.0.3 Subset Sum

The subset sum problem (SSP) is somewhat reminiscent of 2-partition. The SSP's input consists of a multiset  $S = \{a_j\}_{j \in I}$  with  $a_j \in \mathbb{Z}_N$  and  $I \equiv \{1, \dots, n\}$ , and an integer  $B > 0, B < nN$ . It poses the question

$$\text{Does there exist a subset } U \subset I \text{ such that } \sum_{j \in U} a_j = B?$$

The input size is  $n \log_2(N) + \log_2(nN)$ . Karp [46] showed that subset-sum can be reduced to 2-partition, implying that if a polynomial-time algorithm exists for 2-partition, then it can also be applied to subset-sum. We will not reproduce this proof here, but we do note that the special case of subset-sum where  $B = \frac{1}{2} \sum_{j=1}^n a_j$  is equivalent to 2-partition, since

$$\sum_{j \in U} a_j = \frac{1}{2} \sum_{j=1}^n a_j$$

is equivalent to

$$\sum_{j \in U} a_j = \sum_{j \in I \setminus U} a_j.$$

Applying Karp's result, it follows that the general subset-sum problem can be reduced to this special case, an insight that will be directly relevant in chapter 6.

## 3.1 P and NP

### 3.1.1 The Class P

The number of arithmetic operations an algorithm must perform to accomplish the mapping  $x \mapsto y$  is an essential characteristic. The number of operations usually depends on the input  $x$  and, in most cases, grows with input size.

The class of all problems that can be solved in polynomial time is called P. More precisely, if there exists an algorithm that finds the correct output  $y$  for each input  $x \in \mathbb{Z}_L$  using  $\mathbf{N}(L)$  operations, where  $\mathbf{N}(L)$  is polynomial in the input size  $\log_2(L)$ , then the problem is in P. Neither 3SAT nor subset-sum – and, by extension, 2-partition – is known (or believed) to be in P: No polynomial-time algorithm has been found that solves any of these problems. Note that for all three problems, the input size is logarithmic in  $N$ , so an algorithm that runs in  $\mathcal{O}(n^2 N)$  time is still exponential in the input size and does not imply membership in P. As we shall explain shortly, such an algorithm indeed exists for subset-sum.

As an example of a program in P, let us consider linear programming. The input  $y = (A, \mathbf{c}, \mathbf{b})$  consists of a real  $m \times n$  matrix  $A$ , and vectors  $\mathbf{c} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m$ . The correct output is the vector  $\mathbf{x} \in \mathbb{R}^n$  that maximizes  $\langle \mathbf{c}, \mathbf{x} \rangle$  subject to the element-wise constraint  $A\mathbf{x} \leq \mathbf{b}$ . Note that this violates the stipulation that input and output must be elements of a finite set. In practice,  $A, \mathbf{c}, \mathbf{b}$  are rounded to a fixed number of decimal points, so the problem can be considered discrete.

Khachiyan found the first polynomial-time solution algorithm [47] in 1979. Up to that point, it was unknown whether linear programming was in P; the previously known simplex algorithm [48] worked efficiently in most

settings but required an exponential number of arithmetic operations in the worst case. Finally, Karmarkar's projective algorithm [49] is noteworthy for being both efficient in most scenarios and having a worst-case polynomial-time complexity.

### 3.1.2 The Class NP

There are multiple definitions of the computational class NP, with varying degrees of preciseness. A pedestrian one that nonetheless does the job for us is via verification algorithms. We say that a problem is in NP if, given a proposed solution, a polynomial-time algorithm verifies that the solution is correct. Both 3SAT and 2-partition are in NP. Given a 3SAT instance  $\{c_l\}$  and a proposed solution  $\mathbf{x}$ , verification requires plugging  $x_{i(l)}, x_{j(l)}, x_{k(l)}$  into  $c_l$  for each  $l$ , and checking if all the clauses return 1. This is linear in the number of clauses  $n$ . Since the input size is linear in  $n$ , the verification takes  $\mathcal{O}(n)$  operations. There is an additional subtlety here, as plugging in  $\mathbf{x}$  may scale polynomially in  $N$ , which would indeed be exponential in the input size. Ascertaining that this is not the case requires specifying a precise model for how the information in  $\mathbf{x}$  is accessed, which is beyond our discussion's scope. As for 2-partition, a proposed solution consists of a subset  $U \subset I$ . To verify it, it suffices to compute the sums  $\sum_{j \in U} a_j$  and  $\sum_{j \in I \setminus U} a_j$ , which can be done in polynomial time in  $n$ . It is also subject to the aforementioned subtlety, which we will not go into.

## 3.2 NP-hardness

As mentioned in 3.1.1, neither 3SAT nor 2-partition has a known polynomial-time solution algorithm. While this is the case for all NP-hard problems, it does not explicitly enter the definition of NP-hardness. Rather, a problem is called *NP-hard* if any problem in NP can be reduced to it in polynomial time. To be precise, reducing from decision problem  $A$  to decision problem  $B$  means transforming the input  $x_A$  of  $A$  into an input  $x_B(x_A)$  of  $B$  in such a way that  $y_A(x_A) = y_B(x_B(x_A))$  for all  $x_A$ . In particular, if a polynomial-time solution exists for  $B$ , then one for  $A$  exists. Hence, if every problem in NP can be reduced to  $B$ , then so can, in particular, 3SAT. No known polynomial solution exists for 3SAT, so none exists for  $B$ . Hence, no known polynomial solution exists for any problem that is NP-hard.

### 3.2.1 Pseudopolynomial Algorithm for Subset-Sum

In section 3.1.1, we noted that no algorithm polynomial in the input size exists for subset-sum, 2-partition, or 3SAT. However, we have also remarked that a solution polynomial in both  $N$  and  $n$  may still exist, because  $N$  itself is exponential in the input size. This is indeed the case for subset-sum and 2-partition. We will now explain a well-known example of such a solution for the subset-sum problem [9].

Let us define

$$X(S, B) = \begin{cases} 1 & \exists U \subset I : \sum_{j \in U} a_j = B \\ 0 & \text{otherwise,} \end{cases} \quad (3.3)$$

where  $I = \{1, \dots, n\}$  and  $n = |S|$ . This is just a formal way to define the subset-sum problem. We define the prefix sets  $S_j = \{a_1, \dots, a_j\} \subset S$  for  $0 \leq j \leq |S|$  and take note of the recursive property

$$X(S_j, B) = \begin{cases} X(S_{j-1}, B - a_j) \vee X(S_{j-1}, B) & a_j < B, \\ X(S_{j-1}, B) & a_j > B, \\ 1 & a_j = B, \end{cases} \quad (3.4)$$

where  $\vee$  denotes the logical OR. Let us briefly explain why the recursion holds. If there exists a subset of  $S_j$ , the elements of which sum to  $B$ , then either this subset contains  $a_j$ , in which case  $S_{j-1}$  must have a subset that sums to  $a_j - B$ , or it excludes  $a_j$ , in which case  $S_{j-1}$  must have a subset that sums to  $B$ . In the former case,  $X(S_{j-1}, B - a_j) = 1$ , while in the latter case  $X(S_{j-1}, B) = 1$ . Conversely, if either is true, then  $X(S_j, B) = 1$ . The second line in (3.4) is needed because  $X(S, B)$  is only defined for  $B > 0$ , so if  $B - a_j \leq 0$ , only the second case should be considered. Finally, if  $a_j = B$ , then the subset  $\{a_j\} \subset S_j$  sums to  $B$  so  $X(S_j, B) = 1$ . A procedure for solving the subset-sum problem using the recursion (3.4) is shown in algorithm 1. The array  $r$  can have 0, 1, and  $-\infty$  as its entries. At any point in the algorithm,  $r[j, b] = -\infty$  means that  $X(S_j, b)$  is unknown. Otherwise,  $r[j, b]$  is the known value of  $X(S_j, b)$ . The function  $x(j, b)$  is supposed to return the value of  $X(S_j, b)$ . It first checks whether  $S_j$  is the empty set (only if  $j = 0$ ). If so,  $X(S_j, b) = 1$  if and only if  $b = 0$ .

If, on the other hand,  $S_j$  is nontrivial, one must first check if  $X(S_j, b)$  is already known, i.e., if  $r[j, b] \geq 0$ . If this is true,  $X(S_j, b) = r[j, b]$  is returned. Otherwise,  $X(S_j, b)$  is calculated with the recursion (3.4). Finally, to solve the original problem of determining  $X(S, B)$ , one need only solve  $X(S_j, b)$  for  $j = n$  and  $b = B$ .

For the algorithm to solve the subset-sum problem, it must satisfy two criteria. First, it must eventually terminate when given any finite input. Second, it must return the correct value  $X(S, B)$  when it terminates. The algorithm terminates because that  $j$  decreases with each recursion, and the algorithm finishes when  $j = 0$ . As for correctness, it is clear that  $x(0, b)$  returns the correct value for any  $b$  (1 if  $b = 0$  and 0 if  $b > 0$ ). From (3.4), it follows that if  $x(j-1, b)$  returns the correct value for all  $b$ , then  $x(j, b)$  does so, too. Therefore,  $x(n, B)$  returns the correct value by induction.

The technique of dividing the problem  $X(n, B)$  into subproblems  $X(j, b)$  and recursively evaluating the subproblems without repetition is called *dynamic programming*. In many cases, it permits a significant reduction in computational effort:

**Proposition 3.2.1:** *The computational complexity of algorithm 1 is  $\mathcal{O}(nB)$ .*

*Proof:* We begin by pointing out that each function call to  $x(j, b)$  takes constant time (assuming that the cost of the array lookup  $r[j, b]$  does not depend on the array size). More precisely, the number of arithmetic operations in  $x(j, b)$ , excluding the overhead from the recursive calls, is  $\mathcal{O}(1)$ . Next, note that each call to  $x(j, b)$  sets  $r[j, b]$  to a non-negative value before returning. On the other hand,  $x(j, b)$  only places a recursive call if  $r[j, b] < 0$ . Therefore, the number of calls to  $x(j, b)$  is bounded by the number of unique argument tuples  $(j, b)$ . Since  $b = B$  and  $j = n$  initially, and both are non-increasing and non-negative, this number is  $(n+1)(B+1) \in \mathcal{O}(nB)$ . In summary, the non-recursive overhead of each call is  $\mathcal{O}(1)$ , and the number of calls is  $\mathcal{O}(nB)$ , so the computational complexity of the whole algorithm is  $\mathcal{O}(nB)$ .  $\square$

Despite the  $\mathcal{O}(nB)$  computation time, the subset-sum problem is not in P. To understand why, remember that the input size of a problem is defined as the number of bits used to encode the input and that a problem is in P if there is a solution algorithm polynomial in the input size. As previously mentioned, the input size in binary encoding is  $b = n \log_2(N) + \log_2(nN)$  for subset-sum, so  $N \sim 2^{b/(n+\log_2(n))}$ . The bound  $\mathcal{O}(nB)$  does guarantee polynomial complexity in the input size since there is no fixed relation between  $B$  and  $N$ . As a result, this bound does not, in principle, guarantee membership in P. In practice, one expects a scaling  $B \sim nN$  (recall that  $N$  is the maximal magnitude of the integers  $\{a_j\}_{j=1}^n$ ) so that  $B \sim n2^{b/(n+\log_2(n))}$ . Algorithm 1 then takes  $\sim nB = n^2 2^{b/(n+\log_2(n))}$  operations. No algorithm polynomial in  $b$  is known. An algorithm that has polynomial complexity in the magnitude of the input integers but is still exponentially difficult in overall input size is called *pseudopolynomial*. As we have just seen, algorithm 1 is such a case.

### 3.2.2 Weak NP-completeness

The subset-sum problem is NP-hard [46]. It is also clear that it is in NP: Given a proposed solution  $U \subset S$ , the correctness of the solution can be verified by summing the  $m < n$  integers in  $U$  and checking that the sum equals  $B$ . The time this takes is linear in  $m < n$  and the number of bits used to encode each integer. Therefore, the subset-sum problem is NP-complete. One sub-classification of NP-complete problems is the distinction between *strongly* and *weakly* NP-complete problems. A problem is called weakly NP-complete if it is NP-complete but has a pseudopolynomial solution algorithm. Otherwise, it is strongly NP-complete. Recall that the definition of pseudopolynomial constrains the input's magnitude, which only makes sense for number problems. Therefore, any NP-complete problem that is not a number problem is strongly NP-complete. This applies, in particular, to 3SAT.

## 3.3 Optimization Problems

One commonality of NP-complete problems is that all are decision problems. A correct yes/no answer must be returned, given some input. In practice, however, most computational problems are optimization problems that seek to determine the optimal value some quantity ought to take. Extending the NP-completeness classification to optimization algorithms requires linking them to solutions to NP-complete decision problems. For example, consider the following task:

Given a multiset  $S = \{a_j\}_{j \in I}$  with  $a_j \in \mathbb{Z}_+$  and  $I \equiv \{1, \dots, n\}$ , as well as an integer  $A > 0$ , determine  $\max_{U \subset I} \sum_{j \in U} a_j$  subject to the constraint  $\sum_{j \in U} a_j < A$ .

This optimization problem asks for the subset  $U \subset I$  that maximizes the sum of its elements without exceeding  $A$ . It is very similar to the subset-sum problem. The latter takes an integer  $B$  and determines where a subset  $U$  exactly sums to  $B$ . One can show that the optimization problem is at least as hard as the decision problem, meaning that if the optimization problem is solved, the decision problem is, too. To prove this, consider an arbitrary input  $(S, B)$  of the subset-sum decision problem. Now, invoke the (presumed solved) optimization problem with the constraint set to  $A = B + 1$ . If there exists a subset  $U$  such that  $\sum_{j \in U} a_j = B$ , then the largest-sum subset  $U'$  satisfying  $\sum_{j \in U'} a_j < A$  is  $U' = U$ , which sums to  $A - 1$ . The converse is easily seen to be true as well. On the other hand, there is no such subset if and only if  $\sum_{j \in U'} a_j < A - 1$ . Therefore, the answer to the decision problem is

$$\text{SUBSET-SUM}(S, B) = \begin{cases} 1 & \text{if } \sum_{j \in U'} a_j = A - 1 \\ 0 & \text{if } \sum_{j \in U'} a_j < A - 1 \end{cases} \quad (3.5)$$

Formally, we have transformed the input  $(S, B)$  of the decision problem into an input  $(S, A)$  of the optimization problem and transformed the output back to a decision problem output via (3.5). Both transformations take polynomial time. Therefore, if the optimization problem can be solved in polynomial time, so can the decision problem, meaning the former is at least as hard as the latter.

### 3.3.1 Non-Integer Number Problems

So far, we have considered only integer problems, such as Subset-Sum, and non-numerical problems, such as 3SAT. In practice, many problems involve non-integer variables. Both the 2-partition and subset-sum problems defined in section 3 can be easily generalized to the non-integer case by relaxing the  $a_j \in \mathbb{Z}$  requirement to  $a_j \in \mathbb{R}$ . Finding a general solution to the non-integer subset sum problem must be at least as difficult as the integer problem because the latter is a special case.

In principle, algorithm 1 still returns the correct solution in the non-integer case. In section 3.2.1, we showed that the computational complexity of algorithm 1 is bounded by  $\mathcal{O}(nB)$  (c.f. Proposition 3.2.1). This claim no longer holds. The non-recursive overhead of  $\mathbf{x}(j, b)$  is still  $\mathcal{O}(1)$ , and  $\mathbf{x}(j, b)$  gets called at most once for every tuple  $(j, b)$ . However, the number of tuples is no longer  $(n+1)(B+1)$  because  $b$  can be any real number between 0 and  $B$ . The number  $b$  in  $(j, b)$  may take any value obtainable by subtracting a subset of  $\{a_j\}_{j=1}^n$  from  $B$ . The number of possible values is  $\sim 2^n$ , so there is no longer any polynomial bound on the running time.

---

**Algorithm 1** Dynamic programming solution for Subset-Sum.

---

```

1: function  $x(j, b)$ :
2:   if  $j = 0$  then
3:     if  $b = 0$  then
4:       return 1
5:     else
6:       return 0
7:     end if
8:   end if
9:   if  $r[j, b] \geq 0$  then
10:    return  $r[j, b]$ 
11:  else
12:     $a_j \leftarrow S[j]$ 
13:    if  $b - a_j < 0$  then
14:       $r[j, b] \leftarrow x(j - 1, b)$ 
15:    else if  $b - a_j > 0$  then
16:       $r[j, b] \leftarrow x(j - 1, b) \vee x(j - 1, b - a_j)$ 
17:    else
18:       $r[j, b] \leftarrow 1$ 
19:    end if
20:    return  $r[j, b]$ 
21:  end if
22: end function
23: function  $X(S, B)$ :
24:    $n = |S|$ 
25:   Let  $r$  be a 2D array of size  $n \times B$ .
26:   for  $j = 1 \dots n$  do
27:     for  $b = 1 \dots B$  do
28:        $r[i, b] \leftarrow -\infty$ 
29:     end for
30:   end for
31:   return  $x(n, B)$ 
32: end function

```

---

# Chapter 4

## Hubbard Model with two Fermions

In this chapter, we apply the theory outlined in chapter 2 to a concrete model. As mentioned in the introduction, DFT and RDMFT are successful in practice thanks to the balance of accuracy and efficiency they offer for systems with many particles. However, exact mathematical characteristics of  $\mathcal{F}$  that relate to system properties are all but impossible to obtain for systems with many degrees of freedom because this would necessitate an analytical solution to the ground state problem. Nonetheless, valuable insights can be gained by studying toy models with low-dimensional Hilbert spaces [14, 15, 25, 50]. This thesis is in the same vein, and we will restrict ourselves to models small enough to be exactly solvable. We will limit the particle number to  $N = 2$  and work on a one-dimensional chain of arbitrary length, retaining a notion of system size. Complexity results will be expressed in terms of the chain length.

This chapter introduces the one-dimensional Hubbard model in its full generality, subsequently making the restrictions needed to have an exact solution. We apply the analytical machinery explained in chapter 2 and apply it step-by-step to the problem. We then take a detour and examine what happens if, rather than fixing a type of interaction and varying system size, we select a suitably small system size and allow all possible interaction Hamiltonians. Finally, we explain the implications of time-reversal invariance and whether this simplifies the evaluation of  $\mathcal{F}$ .

### 4.0.1 Hamiltonian

Consider a one-dimensional chain of  $L$  sites, populated by  $N$  fermions, governed by the well-known Hubbard Hamiltonian

$$\hat{H} = -t \sum_{i=1}^L \sum_{s=\uparrow\downarrow} \hat{f}_{i+1,s}^\dagger \hat{f}_{i,s} + \hat{f}_{i,s}^\dagger \hat{f}_{i+1,s} + V \sum_{i=1}^L \hat{n}_{i,\uparrow} \hat{n}_{i,\downarrow} \equiv \hat{h} + \hat{W}. \quad (4.1)$$

To make the problem approachable analytically, we make several drastic simplifications. Consider the subspace  $\mathcal{S}_{P,2}$  of states satisfying

$$\hat{N}|\psi\rangle = 2|\psi\rangle, \quad \hat{P}|\psi\rangle = P|\psi\rangle, \quad \hat{\mathbf{S}}^2|\psi\rangle = \hat{S}_z|\psi\rangle = 0.$$

In words, we look at two fermions forming a spin singlet with total momentum  $P$ . Such a state can be conveniently expressed in first quantization as

$$|\psi\rangle = \sum_{p=0}^{L-1} \psi_p |p \uparrow, P-p \downarrow\rangle, \quad (4.2)$$

where  $P-p$  is understood to be defined modulo  $L$ . Note that the number of degrees of freedom is less than  $L$ , since the constraint of spin-antisymmetry implies

$$\psi_p = \psi_{P-p} \quad (4.3)$$

for all  $p$ . We define the index subset

$$\Omega_P = \{0 \leq \alpha < L \mid \alpha \leq [P - \alpha \bmod L]\}, \quad (4.4)$$

such that  $\{\psi_\alpha\}_{\alpha \in \Omega_P}$  are the linearly independent coefficients. Explicitly, the resulting basis for  $\mathcal{S}_{P,2}$  is

$$\{|\alpha \uparrow, \alpha, \downarrow\rangle\}_{\alpha \in [P-\alpha \bmod L]} \cup \left\{ \frac{|P-\alpha \uparrow, \alpha, \downarrow\rangle + |\alpha \uparrow, P-\alpha, \downarrow\rangle}{\sqrt{2}} \right\}_{\alpha < [P-\alpha \bmod L]}. \quad (4.5)$$

The left-hand set has a single element if  $P$  is even and is otherwise empty. The dimension of  $\mathcal{S}_{P,2}$  depends on  $P$  and  $L$  as shown in table 4.1.

Using the definition (2.14), the matrix elements of the 1RDM are

$$\langle p\sigma|\hat{\gamma}|q\sigma'\rangle = \delta_{pq}\delta_{\sigma\sigma'}|\psi_p|^2. \quad (4.6)$$

$\hat{\gamma}$  is diagonal in the simultaneous basis of  $\hat{\sigma}_z$ ,  $\hat{\boldsymbol{\sigma}}^2$ , and  $\hat{P}$ . This was a foregone conclusion by the final claim in section 2.4, as these operators form a complete set and  $\hat{\Gamma} = |\psi\rangle\langle\psi|$  respects all three symmetries.

## 4.1 Levy Functional

Let us recapitulate our goal. We seek to determine the symmetry-restricted functional  $\mathcal{F}_G$  outlined in section 2.10. The symmetry operators are  $\hat{P}$ ,  $\hat{\sigma}_z$  and  $\hat{\boldsymbol{\sigma}}^2$ . Note that  $\hat{\boldsymbol{\sigma}}^2$  is not a one-particle symmetry, so the constraint  $\hat{\boldsymbol{\sigma}}^2 = 0$  on the wave function cannot be represented on the 1RDM level. We must, therefore, revert to the case outlined in the final paragraph of section 2.10.

To compute the functional, our first step is to determine its domain. The domain is  $\mathcal{P}_{N,G}^1$ , the set of 1RDMs that are  $N$ -representable with wave functions in the  $G$ -subspace, where  $G = (P, \sigma_z, \boldsymbol{\sigma}^2)$ . We have already established the most general wave function (4.2) and its corresponding 1RDM (4.6). From (4.6) and (4.3), we see that the 1RDM is diagonal and satisfies the spin constraint

$$\langle p\uparrow|\hat{\gamma}|p\uparrow\rangle = \langle p\downarrow|\hat{\gamma}|p\downarrow\rangle \quad (4.7)$$

and the momentum constraint

$$\langle p\sigma|\hat{\gamma}|p\sigma\rangle = \langle P-p\sigma|\hat{\gamma}|P-p\sigma\rangle, \quad (4.8)$$

hence there are as many linearly independent (real) entries of  $\hat{\gamma}$  as there are (complex) wave function coefficients. The domain of  $\mathcal{F}_G$  thus has dimension  $D$  as defined in table 4.1. Additional nonholonomic constraints are imposed by the normalization condition  $\sum_{p=0}^{L-1} |\psi_p|^2 = 1$ . If  $p \neq P-p \pmod L$ , then the two coefficients in the basis expansion must be equal by (4.3). Together with the normalization condition, this implies  $|\psi_p|^2 \leq 1/2$ . On the other hand, if  $p = P-p \pmod L$ , then (4.3) is trivial, and normalization only implies  $|\psi_p|^2 \leq 1$ . The resulting normalization constraint on the diagonal elements of  $\hat{\gamma}$  is

$$\begin{aligned} 0 \leq \langle p\sigma|\hat{\gamma}|q\sigma\rangle &\leq 1 && \text{if } p = P-p \pmod L, \\ 0 \leq \langle p\sigma|\hat{\gamma}|q\sigma\rangle &\leq 1/2 && \text{if } p \neq P-p \pmod L. \end{aligned} \quad (4.9)$$

We define the rescaled occupation numbers

$$\lambda_\alpha = \begin{cases} \langle \alpha\sigma|\hat{\gamma}|\alpha\sigma\rangle & \text{if } \alpha = P-\alpha \pmod L, \\ 2\langle \alpha\sigma|\hat{\gamma}|\alpha\sigma\rangle & \text{if } \alpha \neq P-\alpha \pmod L. \end{cases} \quad (4.10)$$

Recalling the definition (4.4), it follows from (4.6) and (4.3) that a set of linearly independent occupation numbers is given by  $\{\lambda_\alpha\}_{\alpha \in \Omega_P}$ . The domain of  $\mathcal{F}_G$  is thus

$$\mathcal{D}(\mathcal{F}_G) = \mathcal{P}_{N,G}^1 \simeq \left\{ (\lambda_\alpha)_{\alpha \in \Omega_P} \mid 0 \leq \lambda_\alpha \leq 1, \sum_{\alpha \in \Omega_P} \lambda_\alpha = 1 \right\} = \Delta^{D-1}, \quad (4.11)$$

a  $(D-1)$ -dimensional hypersimplex. We will henceforth write  $\mathcal{F}_G[\boldsymbol{\lambda}]$  in place of  $\mathcal{F}_G[\hat{\gamma}]$ , with the understanding that  $\boldsymbol{\lambda}$  is a vector that parameterizes the constrained  $2L \times 2L$  matrix  $\hat{\gamma} \in \mathcal{P}_{N,G}^1$ . Recall that  $G = (P, \sigma_z, \boldsymbol{\sigma}^2)$  labels the symmetry subspace and that we have restricted ourselves to the spin singlet sector while leaving the momentum  $P$  arbitrary. From (4.11) and table 4.1, we see that the domain differs depending on which symmetry sector one is considering.

Having established its domain, we seek to determine  $\mathcal{F}_G$ , based on the general formula (2.60). Given  $\hat{\gamma} \in \mathcal{D}(\mathcal{F}_G)$ , we must first characterize the set  $\mathcal{P}_{N,G}^1(\hat{\gamma}) = \{|\psi\rangle \in \mathcal{H}_G \mid \hat{\gamma} = \text{Tr}_{N-1}[|\psi\rangle\langle\psi|]\}$ , which the minimization in (2.60) is performed over. We will do so by considering the most general wave function in  $\mathcal{H}_G$ , given in (4.2), and determining the constraints on  $\psi_p$  needed to guarantee  $\hat{\gamma} = \text{Tr}_{N-1}[|\psi\rangle\langle\psi|]$ . This is easy, as we have already found the general form of  $\hat{\gamma}$  in (4.6). It shows that the magnitudes  $|\psi_\alpha|^2 = \langle \alpha\sigma|\hat{\gamma}|\alpha\sigma\rangle$  are fixed once  $\hat{\gamma}$  is

	$L$ even	$L$ odd
$P$ even	$D = \frac{L}{2} + 1$	$D = \frac{L+1}{2}$
$P$ odd	$D = \frac{L}{2}$	$D = \frac{L+1}{2}$

Table 4.1: (complex) degrees of freedom  $D$  for two-Fermion singlet with fixed total momentum  $P$  (without imposing normalization constraint)

specified, and only the phases  $\eta_\alpha \equiv \exp(i \arg \psi_\alpha)$  can be varied. Hence, the minimization in (2.60) runs over the possible values of  $\eta_\alpha$  for all  $\alpha \in \Omega_P$ . To determine (2.29), we compute the matrix elements

$$\begin{aligned} \langle p \uparrow, P-p \downarrow | \hat{W} | q \uparrow, P-q \downarrow \rangle &= \frac{V}{L^2} \sum_{k,l=1}^L \sum_{m,n=1}^L e^{2\pi i(pk+(P-p)l-qm-(P-q)n)/L} \sum_{j=1}^L \langle k \uparrow, l \downarrow | \hat{n}_{j,\uparrow} \hat{n}_{j,\downarrow} | m \uparrow, n \downarrow \rangle \\ &= \frac{V}{L^2} \sum_{j=1}^L \sum_{k,l=1}^L \sum_{m,n=1}^L e^{2\pi i(pk+(P-p)l-qm-(P-q)n)/L} \delta_{jk} \delta_{jl} \delta_{jm} \delta_{jn} = \frac{V}{L}. \end{aligned} \quad (4.12)$$

The variables  $j, k, l, m, n$  are in position space, i.e., they label sites in the chain. The mean interaction energy for the state (4.2) then follows as

$$\langle \psi | \hat{W} | \psi \rangle = \frac{V}{L} \sum_{p,q=0}^{L-1} \psi_p^* \psi_q = \frac{V}{L} \left| \sum_{p=0}^{L-1} \psi_p \right|^2 = \frac{V}{L} \left| \sum_{\alpha \in \Omega_P} \sqrt{\zeta_\alpha \lambda_\alpha} \eta_\alpha \right|^2, \quad (4.13)$$

where we used (4.6) and (4.10), and introduced the numerical factors

$$\zeta_\alpha = \begin{cases} 2 & \text{if } \alpha \neq (P-\alpha) \pmod{L} \\ 1 & \text{if } \alpha = (P-\alpha) \pmod{L}. \end{cases} \quad (4.14)$$

The minimization (2.29) then reduces to

$$\mathcal{F}_G[\boldsymbol{\lambda}] = \frac{V}{L} \min_{\boldsymbol{\eta} \in U(1)^D} \left| \sum_{\alpha \in \Omega_P} \sqrt{\zeta_\alpha \lambda_\alpha} \eta_\alpha \right|^2. \quad (4.15)$$

If  $\sqrt{\zeta_\alpha \lambda_\alpha} \leq \sum_{\beta \neq \alpha} \sqrt{\zeta_\beta \lambda_\beta}$  for all  $\alpha$ , the sum can be made to vanish by choosing the  $\eta_\alpha$  such that  $\sqrt{\zeta_\alpha \lambda_\alpha} \eta_\alpha$  are the side lengths of a polygon in the complex plane. More specifically,  $\eta_\alpha \sqrt{\zeta_\alpha \lambda_\alpha}$  is an arrow in the complex plane, and the arrows can be arranged into a closed loop by adjusting their orientation  $\eta_\alpha$ . Hence, the functional  $\mathcal{F}_G[\boldsymbol{\lambda}]$  vanishes in this case. If, on the other hand,  $\sqrt{\zeta_\alpha \lambda_\alpha} > \sum_{\beta \neq \alpha} \sqrt{\zeta_\beta \lambda_\beta}$  for some  $\alpha$ , the minimum is achieved by choosing

$$\eta_\alpha = \begin{cases} 1 & \text{if } \beta = \alpha \\ -1 & \text{if } \beta \neq \alpha, \end{cases} \quad (4.16)$$

i.e., we orient the arrows such that the longest arrow is pointed opposingly to all others. Putting the two cases together, we arrive at

$$\mathcal{F}_G[\boldsymbol{\lambda}] = \begin{cases} \frac{V}{L} \left( \sqrt{\zeta_\alpha \lambda_\alpha} - \sum_{\beta \neq \alpha} \sqrt{\zeta_\beta \lambda_\beta} \right)^2 & \text{if } \sqrt{\zeta_\alpha \lambda_\alpha} > \sum_{\beta \neq \alpha} \sqrt{\zeta_\beta \lambda_\beta} \text{ for some } \alpha \\ 0 & \text{otherwise} \end{cases} \quad (4.17)$$

Note that  $\mathcal{F}_G$  implicitly depends on  $P$ : On the one hand, its domain (4.11) depends on  $D$ , which in turn depends on  $P$  via table 4.1; on the other hand,  $\zeta_\alpha$  also implicitly depends on  $P$ .

We have found that  $\mathcal{F}_G$  vanishes on a subset of its domain. An alternative way to determine this subset is to note that  $\hat{W}$  is positive-semidefinite, so  $\langle \psi | \hat{W} | \psi \rangle = 0$  if and only if  $\hat{W} | \psi \rangle = 0$ . Therefore,

$$\mathcal{F}_G^{-1}(0) = \{ \hat{\gamma} \mid \text{there exists } |\psi\rangle \text{ such that } \text{Tr}_{N-1}(|\psi\rangle\langle\psi|) = \hat{\gamma} \text{ and } \hat{W}|\psi\rangle = 0 \} \quad (4.18)$$

From (4.1), it is easy to see that  $\hat{W}|\psi\rangle = 0$  if and only if for all  $j$  we have

$$0 = \langle j \uparrow, j \downarrow | \psi \rangle = \frac{1}{L} \sum_{p=0}^{L-1} \psi_p e^{2\pi i P j / L} = \frac{e^{2\pi i P j / L}}{L} \sum_{\alpha \in \Omega_P} \sqrt{\zeta_\alpha \lambda_\alpha} \eta_\alpha, \quad (4.19)$$

This sum can be made to vanish if and only if the upper condition in (4.17) is not met, recovering the zero region of (4.17).

Let us denote by  $\mathcal{R}_D = \mathcal{F}_G^{-1}(0)$  the zero region, explicitly referencing the number of degrees of freedom  $D$ . To understand  $\mathcal{R}_D$  better, recall that  $\beta \in \mathbb{S}^{2D-1} \subset \mathbb{C}^D$  is mapped by (4.6) and (4.10) to  $\lambda \in \Delta^{D-1}$ . Its  $D$  boundary faces,

$$\Delta_\beta^{D-2} \equiv \Delta^{D-1} \cap \{\lambda_\beta = 0\}, \quad (4.20)$$

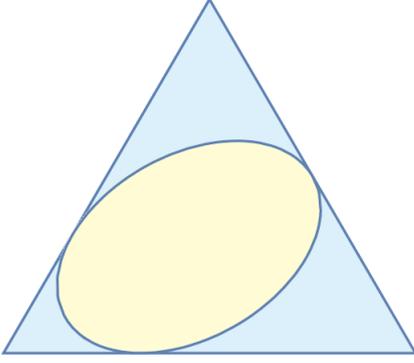
are themselves hypersimplices of dimension one lower. The zero region can be expressed as

$$\mathcal{R}_D = \left\{ \lambda \in \Delta^{D-1} \mid \exists \psi \in \mathbb{S}^{2D-1} : \lambda_\alpha = \zeta_\alpha |\psi_\alpha|^2, \sum_{\alpha \in \Omega_P} \zeta_\alpha \psi_\alpha = 0 \right\} \subset \Delta^{D-1}. \quad (4.21)$$

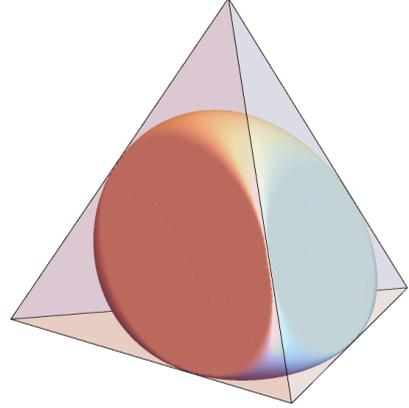
In particular, if we intersect  $\mathcal{R}_D$  with a boundary face, we obtain

$$\mathcal{R}_D \cap \Delta_\beta^{D-2} = \left\{ \lambda \in \Delta^{D-1} \mid \exists \psi \in \mathbb{S}^{2D-3} : \lambda_\alpha = \zeta_\alpha |\psi_\alpha|^2, \sum_{\alpha \in \Omega_P, \alpha \neq \beta} \zeta_\alpha \psi_\alpha = 0 \right\} \sim \mathcal{R}_{D-1}. \quad (4.22)$$

Due to the presence of  $\zeta_\alpha$ , the regions  $\mathcal{R}_D$  take different shapes depending on the values of  $P$  and  $L$ . However, the difference amounts to a mere linear transformation: The form of the equation defining  $\mathcal{R}_D$  only depends on  $D$ . This is illustrated in Fig. 4.1 for  $D = 4$  (three degrees of freedom when accounting for normalization).



(a) Case  $L = 5, P = 2$ . The yellow region is obtained by intersecting the region in 4.1b with one of the bounding faces.



(b) Case  $L = 6, P = 2$ .

Figure 4.1: Shape of zero region for two and three degrees of freedom. The domain of  $\mathcal{F}_G$  is the simplex  $\Delta^{D-1}$ , in this case with  $D = 2, 3$ .

## 4.2 Finding the Ground State Energy

Having determined the functional, one can find the ground state energy, starting from (2.58). Since the functional has been computed separately for each of the  $L$  symmetry sectors labeled by  $P = 0, \dots, L-1$ , one must execute the minimization (2.58) for every sector and finally choose the lowest of the  $L$  minima, as in (2.56). We begin by computing the kinetic energy, which only depends on the 1RDM:

$$\begin{aligned} \text{Tr}(\hat{\gamma}\hat{h}) &= -2t \sum_{p=0}^{L-1} \sum_{\sigma=\uparrow,\downarrow} \langle p\sigma | \hat{\gamma} | p\sigma \rangle \cos(2\pi p/L) \\ &= -2t \sum_{\alpha \in \Omega_P} \lambda_\alpha (\cos(2\pi\alpha/L) + \cos(2\pi(P-\alpha)/L)). \end{aligned} \quad (4.23)$$

Let us consider the special case  $P = 0$ , so the momentum minimizing the kinetic energy is  $\alpha_{\min} = 0$ . In this case (and also for general  $P$ ), the kinetic energy is minimized when  $\lambda_\alpha = 0$  for  $\alpha \neq \alpha_{\min}$ . The points where all except one  $\lambda_\alpha$  vanish are the vertices of the simplex  $\mathcal{D}(\mathcal{F}_G)$ , where the functional reaches a local maximum as seen from (4.17); minimizing the kinetic energy means maximizing the interaction energy.

More generally, suppose  $\hat{h}$  is an arbitrary symmetry-respecting single-particle operator:

$$\hat{h}_{pq} = \delta_{p-q} h_p, \quad (4.24)$$

where we have defined  $\mathbf{h}$  to be the vector with entries the diagonal elements of  $\hat{h}$ . The kinetic energy (4.23) reduces to

$$\text{Tr}(\hat{\gamma}\hat{h}) = \sum_{p=0}^{L-1} \sum_{\sigma=\uparrow,\downarrow} \langle p\sigma|\hat{\gamma}|p\sigma\rangle h_p = \sum_{\alpha \in \Omega_P} \lambda_\alpha (h_{P-\alpha} + h_\alpha) = \boldsymbol{\lambda} \cdot \tilde{\mathbf{h}}, \quad (4.25)$$

where we have defined

$$\tilde{h}_\alpha \equiv h_\alpha + h_{P-\alpha}; \quad \alpha \in \Omega_P \quad (4.26)$$

Substituting this expression in (2.32), we may write

$$E_0(\tilde{\mathbf{h}}) = \min_{\boldsymbol{\lambda} \in \mathcal{D}(\mathcal{F}_G)} (\mathcal{F}_G[\boldsymbol{\lambda}] + \boldsymbol{\lambda} \cdot \tilde{\mathbf{h}}). \quad (4.27)$$

### 4.2.1 Normalization

A technical comment is in order. Note that  $\tilde{\mathbf{h}}$  is an element of  $\mathbb{R}^D$  while  $\boldsymbol{\lambda}$ , being confined to  $\mathcal{D}(\mathcal{F}_G) = \Delta^{D-1} \subset \mathbb{R}^D$ , has only  $D-1$  degrees of freedom. However, the duality between  $\hat{\gamma}$  and  $\hat{h}$  arising from the Legendre-Fenchel transform (2.32) prescribes that  $\hat{\gamma}$  and  $\hat{h}$  live in the same vector space. This discrepancy is resolved by picking  $\alpha \in \Omega_P$  at will, and making the replacement  $\lambda_\alpha = 1 - \sum_{\beta \neq \alpha} \lambda_\beta$  such that

$$\boldsymbol{\lambda} \cdot \tilde{\mathbf{h}} = \sum_{\beta \in \Omega_P} \lambda_\beta \tilde{h}_\beta = \sum_{\beta \neq \alpha} \lambda_\beta \tilde{h}_\beta + \left(1 - \sum_{\beta \neq \alpha} \lambda_\beta\right) \tilde{h}_\alpha \equiv \boldsymbol{\lambda}' \cdot \tilde{\mathbf{h}}' + \tilde{h}_\alpha, \quad (4.28)$$

where we have defined new vectors  $\boldsymbol{\lambda}' \in \mathbb{R}^{D-1}$  and  $\tilde{\mathbf{h}}' \in \mathbb{C}^{D-1}$ : The new occupation number vector,

$$\lambda'_\beta = \lambda_\beta, \quad \beta \in \Omega_P \setminus \{\alpha\}, \quad (4.29)$$

is just  $\boldsymbol{\lambda}$  without the entry corresponding to  $\alpha$ ; The vector  $\tilde{\mathbf{h}}'$  is defined as

$$\tilde{h}'_\beta = \tilde{h}_\beta - \tilde{h}_\alpha, \quad \beta \in \Omega_P \setminus \{\alpha\}. \quad (4.30)$$

The expression for the Levy ground state energy (4.27) then becomes

$$E_0(\tilde{\mathbf{h}}) = \min_{\boldsymbol{\lambda}' \in \mathcal{M}_\alpha(\mathcal{D})} (\mathcal{F}_G[\mathcal{M}_\alpha^{-1}(\boldsymbol{\lambda}')] + \boldsymbol{\lambda}' \cdot \tilde{\mathbf{h}}') + \tilde{h}_\alpha, \quad (4.31)$$

where we have defined the map

$$\mathcal{M}_\alpha : \boldsymbol{\lambda} \mapsto \boldsymbol{\lambda}' \quad (4.32)$$

implementing the parameter change from barycentric to cartesian coordinates for the simplex. Hence  $\mathcal{M}_\alpha(\mathcal{D})$  is the  $D-1$  dimensional simplex embedded directly in  $\mathbb{R}^{D-1}$  rather than in  $\mathcal{X} \subset \mathbb{R}^D$ . (4.31) identifies the eliminated coordinate  $\tilde{h}_\alpha$  as an additive constant in the ground state energy that does not affect the ground state occupation number vector.

To illustrate, let us consider the case  $P=2, L=5$ . The relation between the coefficients  $\lambda_\alpha$  and the occupation numbers  $\langle p\sigma|\hat{\gamma}|p\sigma\rangle$  follows from (4.10):

$$\langle 2\sigma|\hat{\gamma}|2\sigma\rangle = \langle 0\sigma|\hat{\gamma}|0\sigma\rangle = \lambda_0/2; \quad \langle 1\sigma|\hat{\gamma}|1\sigma\rangle = \lambda_1; \quad \langle 3\sigma|\hat{\gamma}|3\sigma\rangle = \langle 4\sigma|\hat{\gamma}|4\sigma\rangle = \lambda_3/2. \quad (4.33)$$

The kinetic energy evaluates to

$$\begin{aligned} \text{Tr}(\hat{\gamma}\hat{h}) &= \sum_{p=0}^{L-1} \sum_{\sigma=\uparrow,\downarrow} \langle p\sigma|\hat{\gamma}|p\sigma\rangle h_p \\ &= \lambda_0(h_0 + h_2) + 2\lambda_1 h_1 + \lambda_3(h_3 + h_4) \\ &= \lambda_0 \tilde{h}_0 + \lambda_1 \tilde{h}_1 + \lambda_3 \tilde{h}_3 \\ &= \lambda_0(\tilde{h}_0 - \tilde{h}_3) + \lambda_1(\tilde{h}_1 - h_3) + \tilde{h}_3 \\ &= \boldsymbol{\lambda}' \cdot \tilde{\mathbf{h}}' + \tilde{h}_3, \end{aligned} \quad (4.34)$$

where we have (arbitrarily) chosen  $\alpha = 3$ .

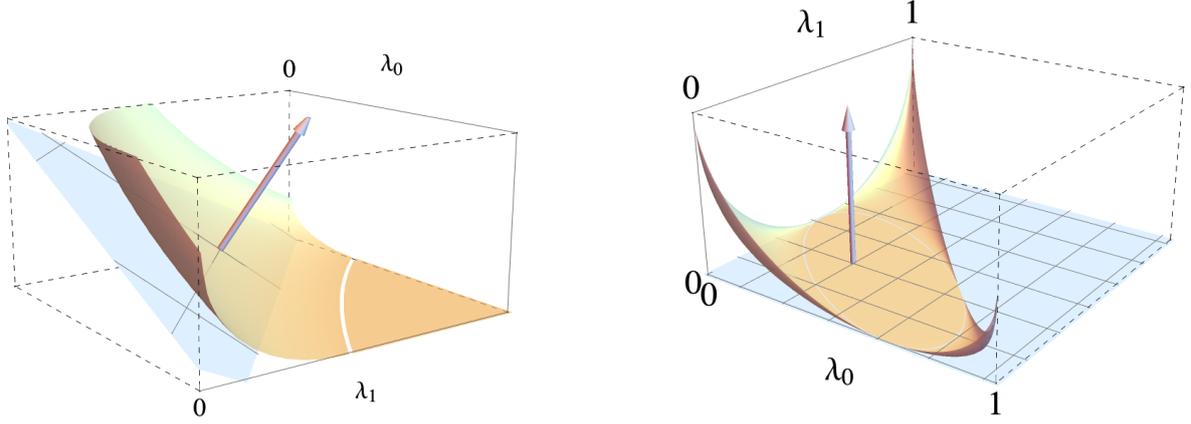
Recall that  $E_0$  can be expressed in terms of the Legendre-Fenchel transform of the functional:

$$E_0(\tilde{\mathbf{h}}') = -\mathcal{F}_G^*(-\tilde{\mathbf{h}}')$$

We now employ the graphical interpretation of the Legendre-Fenchel transform discussed previously. We first consider the plane defined by

$$g_c : \boldsymbol{\lambda}' \mapsto \boldsymbol{\lambda}' \cdot \tilde{\mathbf{h}}' + c. \quad (4.35)$$

Then, we find the lowest value  $c = c_{\min}$  guaranteeing  $g_c(\boldsymbol{\lambda}') = \mathcal{F}_G(\boldsymbol{\lambda}')$  for some  $\boldsymbol{\lambda}'$ . In other words, the plane is made to be tangent to the graph of  $\mathcal{F}_G$ , touching it from below as in Fig. 4.2. Hence, the ground state energy is  $E_0 = \tilde{h}_3 + g_{c_{\min}}(0)$ .



(a)  $\hat{h} \neq 0$  case. The depicted normal vector points in the direction  $(-\tilde{h}_0, -\tilde{h}_1, 1)$ .

(b)  $\hat{h} = 0$  case. The resulting ground-state energy is  $E_0 = \tilde{h}_3 = 0$ .

Figure 4.2: Graphical evaluation of the ground state energy

### 4.2.2 Exchange Force

In Fig. 4.2, the gradient diverges at the domain boundary. Since  $\lambda_i$  are the natural occupation numbers, the domain boundary corresponds to states where one or more orbitals are unoccupied. Recalling the discussion in section 2.8.1, Fig. 4.2 signifies a diverging force that repels the ground-state 1RDM from the boundary. This so-called exchange force has been linked to fermionic exchange symmetry [25] and was previously investigated in systems close to half filling, like the case in Fig. 4.2 which shows  $\mathcal{F}_G$  for  $N = 2$  fermions on  $L = 5$  sites. It is interesting to study whether this exchange force still manifests in the more dilute case with arbitrary  $L$ .

Let us compute the gradient, starting from the analytical expression (4.17). The nontrivial case is the one where  $\sqrt{\lambda_\alpha} > \sum_{\beta \neq \alpha} \sqrt{\lambda_\beta}$  for some  $\alpha$ . In this case, we obtain

$$\frac{\partial \mathcal{F}_G}{\partial \lambda_\beta} = \frac{V}{L} \sqrt{\frac{\zeta_\beta}{\lambda_\beta}} \left( \sqrt{\zeta_\alpha \lambda_\alpha} - \sum_{\varepsilon \neq \alpha} \sqrt{\zeta_\varepsilon \lambda_\varepsilon} \right) \times \begin{cases} 1 & \text{if } \alpha = \beta \\ -1 & \text{if } \alpha \neq \beta \end{cases} \quad (4.36)$$

Note that we have so far regarded  $\mathcal{F}_G$  as a function of  $D$  independent variables, when, in reality, only  $D - 1$  of the  $\lambda$ 's are independent due to the condition  $\sum_\beta \lambda_\beta = 1$ . Hence, we must project the gradient into the linear subspace parallel to  $\text{aff}(\mathcal{D}(\mathcal{F}_G))$ , the affine hull of  $\mathcal{D}(\mathcal{F}_G)$ . Since  $\text{aff}(\mathcal{D}(\mathcal{F}_G))$  is defined by the constraint  $\sum_\alpha \lambda_\alpha = 2$ , its normal vector is  $(1, \dots, 1)^T$ . If we regard  $\partial \mathcal{F}_G / \partial \boldsymbol{\lambda}$  as a column vector, the projection operation into the parallel linear subspace is

$$P \frac{\partial \mathcal{F}_G(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} = \frac{\partial \mathcal{F}_G(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} - \frac{1}{D} \left( (1, \dots, 1) \cdot \frac{\partial \mathcal{F}_G(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \right) \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad (4.37)$$

which leads to

$$\left[ P \frac{\partial \mathcal{F}_G(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \right]_\beta = \sqrt{\frac{V \mathcal{F}_G(\boldsymbol{\lambda})}{L}} \begin{cases} \sqrt{\zeta_\alpha / \lambda_\alpha} - \frac{1}{K} \left[ \sqrt{\zeta_\alpha / \lambda_\alpha} - \sum_{\varepsilon \neq \alpha} \sqrt{\zeta_\varepsilon / \lambda_\varepsilon} \right] & \text{if } \alpha = \beta \\ -\sqrt{\zeta_\beta / \lambda_\beta} - \frac{1}{K} \left[ \sqrt{\zeta_\alpha / \lambda_\alpha} - \sum_{\varepsilon \neq \alpha} \sqrt{\zeta_\varepsilon / \lambda_\varepsilon} \right] & \text{if } \alpha \neq \beta \end{cases} \quad (4.38)$$

Let us see how the functional behaves when approaching a bounding face  $\lambda_\beta = 0$ ,  $\beta \neq \alpha$ . We set  $\lambda_\beta = \epsilon$  and let  $\epsilon \rightarrow 0$ . In this limit,

$$P \frac{\partial \mathcal{F}_G(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \rightarrow -\sqrt{\frac{V \mathcal{F}_G(\boldsymbol{\lambda}) \boldsymbol{\zeta}}{L \epsilon}} \left( \mathbf{e}_\beta - \frac{1}{K} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \right) + O(1), \quad (4.39)$$

where  $\mathbf{e}_\beta$  is a vector with components  $(\mathbf{e}_\beta)_\alpha = \delta_{\alpha\beta}$  and the square root of  $\boldsymbol{\zeta}$  is taken entry-wise. We see that the leading  $O(\epsilon^{-1/2})$  part is just the projection of  $\mathbf{e}_\beta$  into the physical (affine) subspace  $\mathcal{X}$ . In other words, the gradient becomes an inward-pointing normal vector to the  $\lambda_\beta = 0$  face of the Pauli hypersimplex. This concludes our discussion of the first case in (4.17).

If, on the other hand, the boundary point in question satisfies  $\sqrt{\zeta_\alpha \lambda_\alpha} < \sum_{\beta \neq \alpha} \sqrt{\zeta_\beta \lambda_\beta}$  for all  $\alpha$ , the point lies in the interior the intersection calculated in (4.22), and the function remains identically zero to the boundary. In this case, the gradient also vanishes,

$$P \frac{\partial \mathcal{F}_G}{\partial \boldsymbol{\lambda}} = 0, \quad (4.40)$$

and no exchange force manifests. This is again due to the simplification in dealing only with the case  $N = 2$ . For  $N \approx L$ , the vanishing of one occupation number necessarily implies that some sites must become doubly occupied. This is not the case for  $N = 2$ : the exchange force only appears close to the vertices, where a few of the  $\lambda_\beta$  dominate.

The variational criterion (2.49) lends a more precise interpretation to the exchange force. (We return for convenience to the notation of the previous subsection, using  $\boldsymbol{\lambda}'$  and  $\tilde{\mathbf{h}}'$ , and working in the coordinates of  $\mathcal{D}' \subset \mathbb{R}^D$  rather than  $\mathcal{D} \subset \mathbb{R}^{D-1}$ .) For the ground state occupation vector  $\boldsymbol{\lambda}'_0$  to reach the vicinity of a vertex (i.e. the region  $\sqrt{\zeta_\alpha \lambda_\alpha} > \sum_{\beta \neq \alpha} \sqrt{\zeta_\beta \lambda_\beta}$ ), the kinetic term  $\tilde{\mathbf{h}}'$  must be proportional to the exchange force at  $\boldsymbol{\lambda}'_0$ :

$$\tilde{\mathbf{h}}' = -\frac{\partial \mathcal{F}_G}{\partial \boldsymbol{\lambda}'}[\boldsymbol{\lambda}'_0]. \quad (4.41)$$

If one tries to approach a bounding face, the magnitude of the kinetic term required to achieve a distance  $\epsilon$  from the face diverges as  $\epsilon^{-1/2}$ . An example is shown in Fig. 4.3.

## 4.3 Ground State Degeneracy

Recall how the closed-form expression (4.17) was obtained from (4.15). We thought of the complex numbers  $\sqrt{\zeta_\alpha \lambda_\alpha} \eta_\alpha$  as vectors in the complex plane;  $\boldsymbol{\lambda} \in \mathcal{F}_G^{-1}(0)$  holds if the phases  $\eta_\alpha$  (i.e. the orientations of the vectors) can be adjusted to arrange the vectors into the sides of a polyhedron, and the sum vanishes. Otherwise, the minimizing phases are given by (4.16). We analyze the ground state degeneracy for the two cases in turn.

### 4.3.1 Degeneracy Region

The region where the choice of  $\boldsymbol{\eta}$  is not unique is, strictly speaking, not given by  $\mathcal{F}_G^{-1}(0)$ , but by its interior: If  $\boldsymbol{\lambda}$  is on the boundary of  $\mathcal{F}_G^{-1}(0)$ , we have  $\sqrt{\lambda_\alpha} = \sum_{\beta \neq \alpha} \sqrt{\lambda_\beta}$ . The choice (4.16) yields  $\mathcal{F}_G = 0$ , but no other choices of  $\boldsymbol{\eta}$  do. We denote

$$\mathcal{D}_0 = \text{int}(\mathcal{F}_G^{-1}(0)) = \left\{ \boldsymbol{\lambda} \in \Delta^{D-1} \mid \sqrt{\lambda_\alpha} < \sum_{\beta \neq \alpha} \sqrt{\lambda_\beta} \text{ for all } \alpha \right\}. \quad (4.42)$$

If  $\boldsymbol{\lambda} \in \mathcal{D}_0$ , the exact constraints for  $\boldsymbol{\eta}$  to yield  $\mathcal{F}_G = 0$  are nontrivial. However, one can heuristically argue what the number of freedom degrees for  $\boldsymbol{\eta}$  ought to be. Suppose  $\boldsymbol{\lambda}$  has three nonvanishing components. Once  $\eta_1$  has been chosen, there is a unique choice of  $\eta_2$  and  $\eta_3$  such that the three vectors form a triangle. Since  $\eta_1$  corresponds to a choice of overall phase for  $|\psi\rangle$ , we conclude the overall number of degrees of freedom to be zero. More generally, if  $\boldsymbol{\lambda}$  has  $n$  nonvanishing components, the first  $n - 2$  components in  $\boldsymbol{\eta}$  may be picked from some open subset of  $U(1)^{n-2}$ , while the last two will be uniquely determined; hence there are in general  $D - 2$  real degrees of freedom in  $|\psi\rangle$ , for fixed  $\boldsymbol{\lambda} \in \mathcal{F}_G^{-1}(0)$ .

The only choice of  $\mathbf{h}$  yielding a ground state density  $\boldsymbol{\lambda} \in \mathcal{D}_0$  is  $h_\beta = h$  independent of  $\beta$ , which can be seen from (4.30). The number of real degrees of freedom in choosing the ground state  $|\psi\rangle$  is  $\dim(\mathcal{F}_G^{-1}(0)) = D - 1$ , in addition to the  $D - 2$  DOF in  $\boldsymbol{\eta}$ , giving  $2D - 3$  DOF in total.

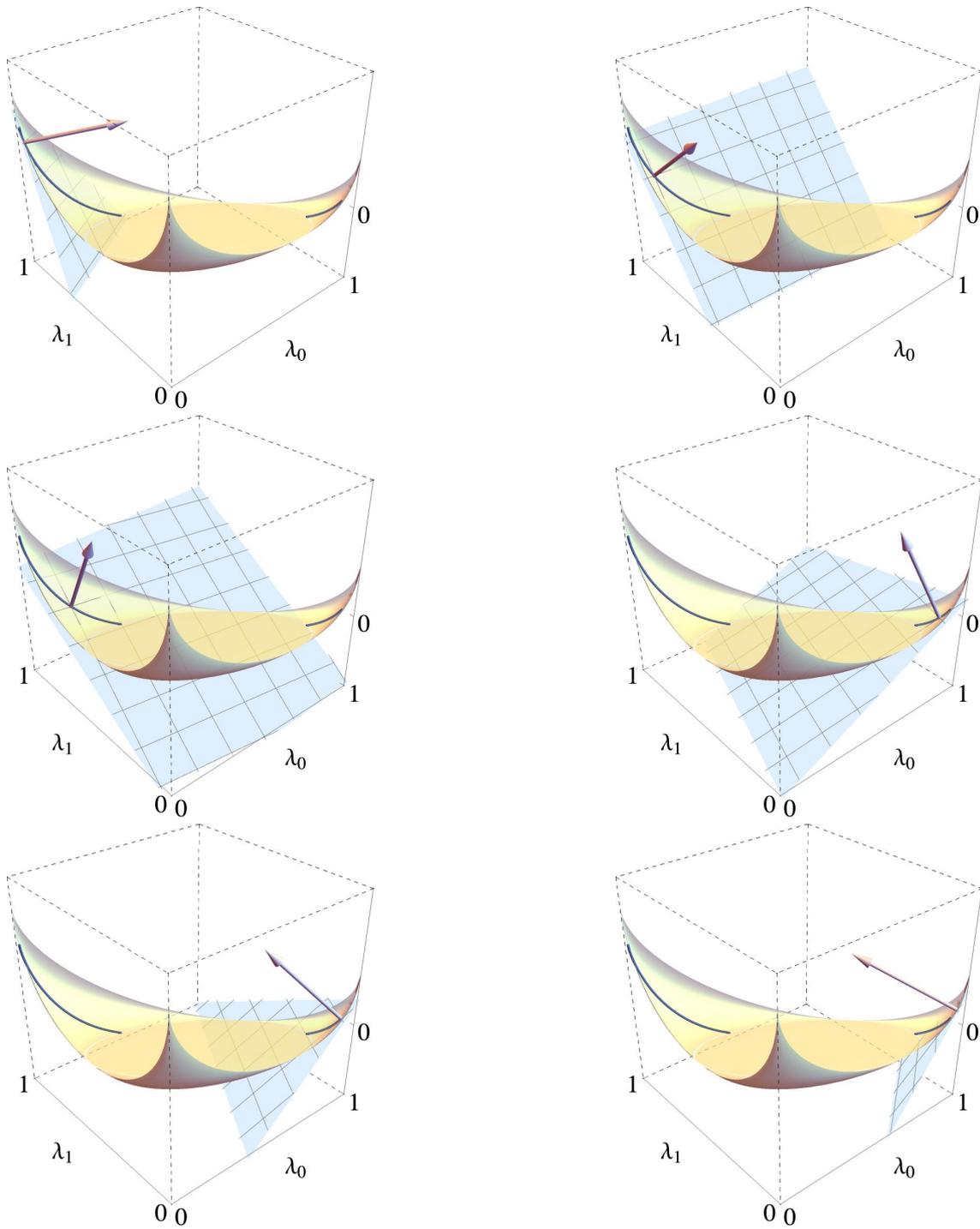


Figure 4.3: When trying to reach ground state occupations close the bounding face  $\lambda_0 + \lambda_1 = 1$ , exceedingly steep tangent planes are required, manifesting in large values of  $\tilde{h}'$ .

To crosscheck this reasoning, note that the choice  $h_\beta = h$  is equivalent to  $\hat{h} = 0$ . In this case, any wave function satisfying (4.19) is a ground state wave function. The constraint (4.19) removes one of the original  $D$  complex degrees of freedom. Accounting for the overall phase of  $|\psi\rangle$ , we have again  $2D - 3$  real degrees of freedom.

As a final remark, we note that for  $h_\beta = h$ , the Hilbert space decomposes into a one-dimensional subspace with eigenvalue  $E = V$  and a  $(D - 1)$ -dimensional subspace with eigenvalue  $E = 0$ .

### 4.3.2 Nondegenerate Regions $S_\alpha$

If  $\lambda \notin \mathcal{D}_0$ , (4.16) is the unique solution to the minimization problem. This, however, does not imply that every  $\lambda \notin \mathcal{D}_0$  has a unique ground state  $|\psi_0\rangle$ . On the contrary, if  $\lambda \in \partial\mathcal{D}_0$ ,  $|\psi_0\rangle$  may be the ground state of the Hamiltonian  $\hat{H} = \hat{W}$  with  $\hat{h} = 0$  and thus be part of the larger ground state manifold defined by  $\hat{W}|\psi\rangle = 0$ .

We want to argue that the  $(D - 1)$ -fold degeneracy only occurs in the case  $h_\beta = h$ . To this end, note that by (4.12),  $\hat{W}$  is given by  $|\phi\rangle\langle\phi|$  with

$$|\phi\rangle = \mathcal{K} \sum_{p=0}^{L-1} |p \uparrow, P - p \downarrow\rangle \quad (4.43)$$

and  $\mathcal{K}$  some normalization constant. In general, a hermitian operator  $\hat{A} \in \text{End}(\mathbb{R}^D)$  will have a  $(D - 1)$ -dimensional eigenspace if and only if  $\hat{A} = a\hat{P} + b\hat{I}$  for some rank-1 projector  $\hat{P}$  and  $a, b \in \mathbb{R}$ . Since  $\hat{H} = \hat{W} + \hat{h} = |\phi\rangle\langle\phi| + \hat{h}$ , this is true for  $\hat{H}$  only if  $\hat{h} = h\hat{I}$  with  $h \in \mathbb{R}$ . (Remember that  $\hat{h}$  is assumed diagonal, whereas  $\hat{W} = |\phi\rangle\langle\phi|$  is assumed off-diagonal.)

In summary, the above implies that all eigenstates are nondegenerate unless  $\hat{h} = h\hat{I}$ . Let us, for concreteness, consider the case  $D = 3$ . In fig. 4.4, we plot the spectrum of  $\hat{H}$  over  $h_1$  and  $h_0$ , with  $h_3 = 1 - h_1 - h_0$ .

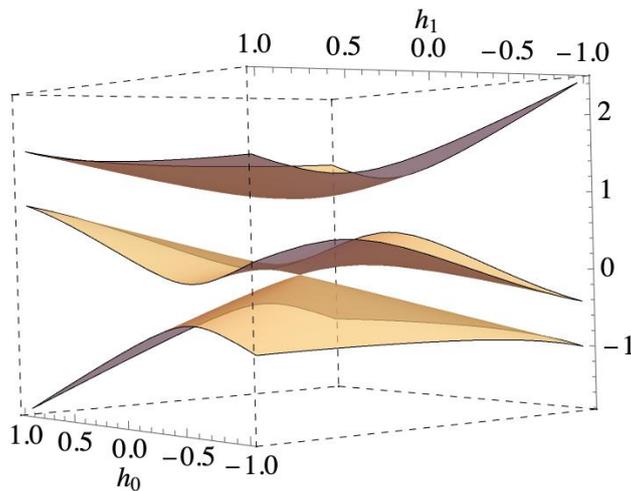


Figure 4.4: Eigenvalues of  $\hat{H} = \hat{W} + \hat{h} = |\phi\rangle\langle\phi| + \hat{h}$  plotted over  $h_0$  and  $h_1$  for the case  $D = 3, P = 2, L = 5, V = 1$ . The condition  $h_0 + h_1 + h_3 = 0$  is enforced by adding  $-\text{Tr}(\hat{h})\hat{I}$  to the Hamiltonian. It is apparent that the spectrum is non-degenerate for all  $\hat{h} \neq 0$ .

## 4.4 Convexity

Here, we aim to prove that  $\mathcal{F}_G$  is convex. To this end it is convenient to make the replacement  $\zeta_\alpha \lambda_\alpha \rightarrow \lambda_\alpha$ . Since this amounts to rescaling the domain and functional, the functional's convexity properties are unaffected.

The functional (4.17) decomposes  $\mathcal{D}(\mathcal{F}_G)$  into  $D$  subsets:

$$\mathcal{S}_\alpha \equiv \left\{ \lambda \in \Delta^{D-1} \mid \sqrt{\lambda_\alpha} > \sum_{\beta \neq \alpha} \sqrt{\lambda_\beta} \right\}; \quad \alpha \in \Omega_P \quad (4.44)$$

are the  $D - 1$  regions where the functional is nontrivial, while

$$\mathcal{R} \equiv \mathcal{F}_G^{-1}(0)$$

is the zero region discussed earlier (where we denoted it  $\mathcal{R}_D$ ). Note that  $\mathcal{F}_G$  is analytic in  $\mathcal{R}$  and in  $\mathcal{S}_\alpha$ , while we cannot expect it to be analytic at the boundary.

Remember that the domain  $\mathcal{D} = \Delta^{D-1}$  is in the affine subspace  $\{\sum_{\alpha \in \Omega_P} \lambda_\alpha = 1\} \subset \mathbb{R}^D$ . In the following, working in cartesian rather than barycentric coordinates will be more convenient. We use the coordinate transformation  $\mathcal{M}_\beta$  defined in (4.32). Recall that  $\mathcal{M}_\beta$  removes the  $\beta$  entry of  $\boldsymbol{\lambda}$  to yield  $\boldsymbol{\lambda}'$ , and that it is injective because the deleted entry can be reconstructed using  $\lambda_\alpha = 1 - \sum_{\beta \neq \alpha} \lambda_\beta$ . The explicit form of the reparameterized functional is thus

$$\mathcal{F}_G(\mathcal{M}_\beta^{-1}(\boldsymbol{\lambda}')) = \frac{V}{L} \left( \left( 1 - \sum_{\beta \neq \alpha} \lambda'_\beta \right)^{1/2} - \sum_{\beta \neq \alpha} (\lambda'_\beta)^{1/2} \right)^2 = \frac{V}{L} \mathcal{G}(\boldsymbol{\lambda}'), \quad (4.45)$$

where we define

$$\mathcal{G}(\boldsymbol{\lambda}') \equiv \left( 1 - \sum_{\beta \neq \alpha} \lambda'_\beta \right)^{1/2} - \sum_{\beta \neq \alpha} (\lambda'_\beta)^{1/2}.$$

Moreover, let us note that the parameterization of (4.44) as

$$\mathcal{M}_\beta(\mathcal{S}_\alpha) = \left\{ \boldsymbol{\lambda}' \in \Delta^{D-1} \mid \left( 1 - \sum_{\beta \neq \alpha} \lambda'_\beta \right)^{1/2} > \sum_{\beta \neq \alpha} (\lambda'_\beta)^{1/2} \right\}. \quad (4.46)$$

For the remainder of this section, we will drop the prime from  $\boldsymbol{\lambda}'$ , always referring to  $\boldsymbol{\lambda} \in \Delta^{D-1} \subset \mathbb{R}^{D-1}$ .

The heavy lifting needed to establish convexity is done by the following Lemma:

**Lemma 4.3.1:** *Let  $\boldsymbol{\lambda}$  be in the interior of  $\mathcal{S}_\alpha$ , such that  $\mathcal{F}_G$  is analytic at  $\boldsymbol{\lambda}$ , and let*

$$H_{\beta\varepsilon}(\boldsymbol{\lambda}) = \frac{\partial^2 \mathcal{F}_G}{\partial \lambda_\beta \partial \lambda_\varepsilon}(\boldsymbol{\lambda})$$

*be the entries of the Hessian matrix.  $H(\boldsymbol{\lambda})$  is positive-semidefinite.*

*Proof:* Explicit computation shows

$$\frac{\partial^2 \mathcal{F}_G}{\partial \lambda_k \partial \lambda_l} = \frac{V}{L} \left( \frac{\partial^2 \mathcal{G}}{\partial \lambda_\beta \partial \lambda_\tau} \mathcal{G} + \frac{\partial \mathcal{G}}{\partial \lambda_\beta} \frac{\partial \mathcal{G}}{\partial \lambda_\tau} \right) = \frac{V}{2L} \left( \frac{1}{2} A_{\beta\varepsilon} \mathcal{G} + 2B_{\beta\varepsilon} \right), \quad (4.47)$$

with

$$A_{\beta\varepsilon}(\boldsymbol{\lambda}) = - \left( 1 - \sum_{\delta} \lambda_\delta \right)^{-3/2} + \delta_{\beta\varepsilon} \lambda_\beta^{-3/2}, \quad (4.48)$$

and

$$B(\boldsymbol{\lambda}) = (\text{grad } \mathcal{F}_G(\boldsymbol{\lambda}))^T \text{grad } \mathcal{F}_G(\boldsymbol{\lambda}). \quad (4.49)$$

Clearly,  $B$  is positive-semidefinite; since  $\mathcal{G}$  is non-negative in  $\mathcal{S}_D$ , it only remains to be shown that  $A$  is positive-semidefinite. To this end, pick an arbitrary  $\mathbf{v} \in \mathbb{R}^{D-1}$ , and evaluate

$$\mathbf{v}^T \mathbf{A} \mathbf{v} = -(1 - \|\boldsymbol{\lambda}\|_1)^{-3/2} \|\mathbf{v}\|_1^2 + \sum_{\beta} \lambda_\beta^{-3/2} v_\beta^2; \quad (4.50)$$

Note that the  $l$ -norm on  $\mathbb{R}^{D-1}$  is defined as  $\|\mathbf{x}\|_l = (\sum_{\alpha} x_\alpha^l)^{1/l}$  for  $l \geq 1$ . We must show that (4.50) is non-negative for all  $\mathbf{v}$ .

As a first step, we establish

$$\sum_{\beta} \lambda_\beta^{-3/2} v_\beta^2 \geq \|\boldsymbol{\lambda}\|_{3/2}^{-3/2} \|\mathbf{v}\|_1^2, \quad (4.51)$$

which immediately follows from the Cauchy-Schwarz inequality applied to  $x_\beta = \lambda_\beta^{3/4}$  and  $y_\beta = \lambda_\beta^{-3/4} v_\beta$ . We substitute (4.51) in (4.50), yielding

$$\mathbf{v}^T \mathbf{A} \mathbf{v} \geq \|\mathbf{v}\|_1^2 \left( -(1 - \|\boldsymbol{\lambda}\|_1)^{-3/2} + \|\boldsymbol{\lambda}\|_{3/2}^{-3/2} \right). \quad (4.52)$$

Now, since  $\boldsymbol{\lambda} \in \mathcal{S}_\alpha$  by assumption, (4.46) implies

$$(1 - \|\boldsymbol{\lambda}\|_1)^{-3/2} \leq \left( \sum_\alpha \lambda_\alpha^{1/2} \right)^{-3} = \|\boldsymbol{\beta}\|_1^{-3}, \quad (4.53)$$

where we have defined  $\psi_\alpha = \lambda_\alpha^{1/2}$ . Substituting this in (4.52),

$$\mathbf{v}^T \mathbf{A} \mathbf{v} \geq \|\mathbf{v}\|_1^2 (-\|\boldsymbol{\beta}\|_1^{-3} + \|\boldsymbol{\beta}\|_3^{-3}). \quad (4.54)$$

Now, use the norm inequality

$$\|\boldsymbol{\beta}\|_l \leq \|\boldsymbol{\beta}\|_1 \quad (4.55)$$

with  $l = 3$  to see that  $\|\boldsymbol{\beta}\|_3^{-3} > \|\boldsymbol{\beta}\|_1^{-3}$ . This finally shows that  $\mathbf{v}^T \mathbf{A} \mathbf{v} \geq 0$ , concluding the proof of Lemma 4.3.1.  $\square$

Let  $\mathcal{D}_0$  be defined as in (4.42), and define

$$\mathcal{F}_\alpha : \Delta^{D-1} \rightarrow \mathbb{R} : \quad \mathcal{F}_\alpha(\boldsymbol{\lambda}) = \frac{V}{L} \left( \sqrt{\lambda_\alpha} - \sum_{\beta \neq \alpha} \sqrt{\lambda_\beta} \right)^2, \quad (4.56)$$

such that Eq. (4.17), with  $\zeta_\alpha = 1$ , can be written as

$$\mathcal{F}_G[\boldsymbol{\lambda}] = \begin{cases} \mathcal{F}_\alpha(\boldsymbol{\lambda}) & \text{if } \sqrt{\lambda_\alpha} > \sum_{\beta \neq \alpha} \sqrt{\lambda_\beta} \text{ for some } \alpha \\ 0 & \text{otherwise.} \end{cases} \quad (4.57)$$

The following lemma is a trivial consequence of Eq. (4.17):

**Lemma 4.3.2:**  $\mathcal{F}_\alpha(\boldsymbol{\lambda}) = 0$  if and only if  $\boldsymbol{\lambda} \in \partial \mathcal{D}_0$ , where the latter denotes the boundary of  $\mathcal{D}_0$ .  $\square$

Another useful result is the following:

**Lemma 4.3.3:** Assume that  $D > 2$ . If  $\mathcal{F}_\alpha(\boldsymbol{\lambda}) = \mathcal{F}_\beta(\boldsymbol{\lambda}) = 0$  for  $\alpha \neq \beta$  then  $\boldsymbol{\lambda} \in \partial \Delta^{D-1}$ .

*Proof:* The equality implies that both  $\sqrt{\lambda_\alpha} = \sum_{\gamma \neq \alpha} \sqrt{\lambda_\gamma}$  and  $\sqrt{\lambda_\beta} = \sum_{\gamma \neq \beta} \sqrt{\lambda_\gamma}$  hold, implying  $\lambda_\alpha = \lambda_\beta$  and  $\lambda_\gamma = 0$  for all  $\gamma \neq \alpha, \beta$ . Therefore,  $\boldsymbol{\lambda}$  has at least one vanishing component and lies on the boundary of the simplex.

Next, we prove another auxiliary result:

**Lemma 4.3.4:** Define  $f : [-1, 1] \rightarrow \mathbb{R}$  by

$$f(t) = \begin{cases} 0 & \text{if } t < 0, \\ g(t) & \text{if } t > 0, \end{cases} \quad (4.58)$$

where  $g(x) : [-1, 1] \rightarrow \mathbb{R}$  is twice differentiable and satisfies  $g'(t) > 0$  and  $g(0) = g'(0) = 0$ . Then  $f(t)$  is convex.

*Proof:* We prove that  $f'(t)$  exists everywhere and is nondecreasing. Existence follows because  $\lim_{t \downarrow 0} f'(t) = \lim_{t \downarrow 0} g'(t) = 0$  and  $\lim_{t \uparrow 0} f'(t) = 0$ . It is nondecreasing because, for  $t < 0$ ,  $f'(t) = 0$  by definition, for  $t > 0$ ,  $f'(t) = g'(t) > 0$  and for  $t = 0$ ,  $f'(t) = 0$  as we just showed.  $\square$

**Proposition 4.3.5:**  $\mathcal{F}_G$  is convex.

*Proof:* Consider two arbitrary points  $\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 \in \text{Int } \Delta^{D-1}$  and define  $f(t) = \mathcal{F}_G(\boldsymbol{\lambda}(t))$  where  $\boldsymbol{\lambda}(t) = t\boldsymbol{\lambda}_1 + (1-t)\boldsymbol{\lambda}_2$ ,  $0 < t < 1$ . If we can show that  $f(t)$  is convex, we will have shown that  $\mathcal{F}_G$  is convex on the interior of its domain and, by continuity, on the entire domain.

For  $h(t)$  to be convex, it suffices for it to be differentiable everywhere and for  $h'$  to be nondecreasing. This follows from Lemma 4.3.1 for all  $t$  where  $\boldsymbol{\lambda}(t)$  is in the interior of either  $\mathcal{D}_0$  or  $\mathcal{S}_\alpha$ . Lemmas 4.3.2 and 4.3.3 and the convexity of  $\mathcal{D}_0$  make clear that  $\boldsymbol{\lambda}(t)$  crosses the boundaries between  $\mathcal{D}_0$  and the  $\mathcal{S}_\alpha$ 's at most twice (in particular, finitely many times). Let  $t_0$  be a crossing point, i.e.  $\boldsymbol{\lambda}(t_0) \in \partial \mathcal{D}_0$ , and choose  $\epsilon > 0$  so that no other crossing point lies inside  $[t_0 - \epsilon, t_0 + \epsilon]$ . If the crossover is from  $\mathcal{D}_0$  into  $\mathcal{S}_\alpha$ , then  $f(t) = h(t_0 + \epsilon t)$  fulfills the conditions of Lemma 4.3.4. If the crossover is from  $\mathcal{S}_\alpha$  into  $\mathcal{D}_0$ , then  $f(t) = h(t_0 - \epsilon t)$  fulfills the conditions. In either case, Lemma 4.3.4 establishes that  $h'$  exists and is nondecreasing at  $t_0$  and thus on  $[0, 1]$ .  $\square$

Proposition 4.3.5 establishes that, for the two-fermion Hubbard system we are considering, the Levy functional coincides with the Valone functional (c.f. section 2.8). Thus, not only does it have a straightforward closed-form expression, but it is also convex. As we will see shortly, this property cannot be taken for granted, even in such a simple system.

## 4.5 Working with Real Coefficients

To reduce notational clutter, we dispose of the subscript  $G$  that reminds us of the symmetry subspace and write  $\mathcal{F}$  instead of  $\mathcal{F}_G$ . The functional  $\tilde{\mathcal{F}}$ , to be defined below, has the same physical origin as  $\mathcal{F}$  and should, strictly speaking, be denoted  $\tilde{\mathcal{F}}_G$ , which we refrain from in service of readability.

In momentum space, the Hamiltonian is a real matrix, so the ground state may a priori be assumed to have a strictly real basis expansion, and one may replace minimization in (4.15) by minimization over real phases  $\sigma_k = \pm 1$  [22, 23]:

$$\tilde{\mathcal{F}}[\boldsymbol{\lambda}] = \tilde{\mathcal{F}}[(\lambda_\alpha)] = \frac{V}{2L} \min_{\boldsymbol{\sigma} \in \{\pm 1\}^D} \left| \sum_{\alpha \in \Omega_P} \sqrt{\zeta_\alpha} \lambda_\alpha \sigma_\alpha \right|^2. \quad (4.59)$$

For general  $L$ , there no longer is a straightforward analytical expression as in the case of complex phases  $\eta_k \in U(1)$ . To still gain some insights, we consider the concrete case  $D = 3$ . Physically, this might, for example, correspond to  $L = 6$  and  $P = 1$ , as per table 4.1. In this case, the (three) phase values needed to achieve the minimum in (4.59) are

$$\sigma_\alpha = \begin{cases} -1 & \text{if } \alpha = \alpha' \\ 1 & \text{otherwise,} \end{cases} \quad (4.60)$$

where  $\alpha'$  is the momentum with the highest occupation  $\lambda_\alpha$ . This divides the simplex  $\mathcal{D}(\tilde{\mathcal{F}})$  into three regions, as shown in Fig. 4.5. At the boundary of those regions,  $\tilde{\mathcal{F}}$  is nonanalytic, as seen in Fig. 4.6a.

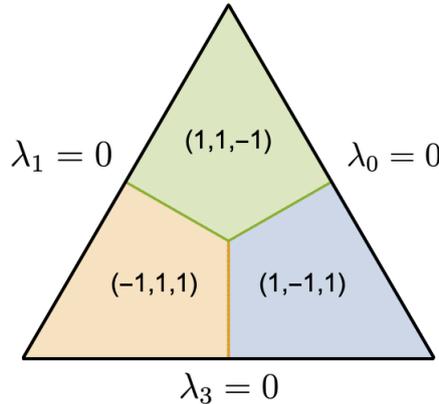


Figure 4.5: Subdomains (6.1) of the  $\boldsymbol{\lambda}$ -simplex corresponding to different choices of  $\boldsymbol{\sigma} = (\sigma_0, \sigma_1, \sigma_3)$ . The parameters  $\lambda_i, i = 0, 1, 3$  serve as barycentric coordinates. The case depicted is  $L = 6, P = 1$ .

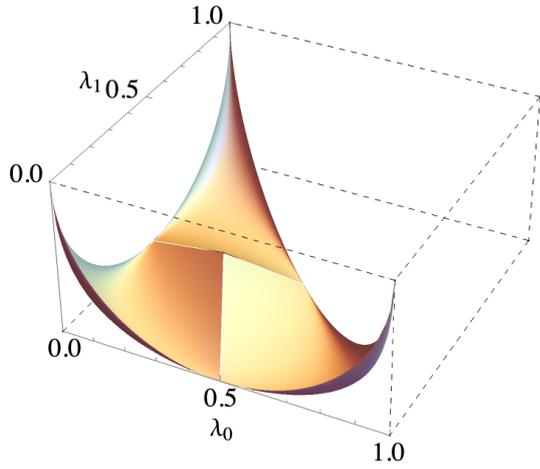
The two functionals  $\mathcal{F}$  and  $\tilde{\mathcal{F}}$  do not agree on the entirety of their domain. Still, they are physically equivalent since the minimization (2.32) only depends on the convex hull of  $\tilde{\mathcal{F}}$ . We see that  $\mathcal{F}$  and  $\tilde{\mathcal{F}}$  have the same convex hull and thus yield the same ground state energy for every choice of  $h$ .

For general values of  $D$ , as is the case for  $D = 3$ , the domain gets subdivided into cells, each corresponding to a choice of signs  $\boldsymbol{\sigma}_\alpha \in \{\pm 1\}^D$ . At the cell boundaries, the functional is continuous, but generally not differentiable, producing kinks in the functional's graph. We can assert that these kinks will always point upward: When moving across the kink on a parametric curve  $\gamma(t)$ , the derivative of  $\mathcal{F}(t) \equiv \tilde{\mathcal{F}}(\gamma(t))$  jumps from  $\tilde{\mathcal{F}}'(t_0^-)$  to  $\tilde{\mathcal{F}}'(t_0^+) < \tilde{\mathcal{F}}'(t_0^-)$ . More precisely, we claim

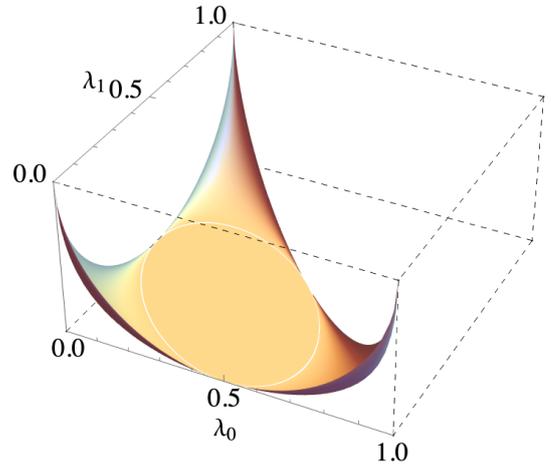
$$\lim_{t \rightarrow t_0^-} \tilde{\mathcal{F}}'(t) \geq \lim_{t \rightarrow t_0^+} \tilde{\mathcal{F}}'(t). \quad (4.61)$$

This implies that if  $\tilde{\mathcal{F}}'(t_0)$  is discontinuous at  $t_0$ , the limiting value  $t \rightarrow t_0^-$  must exceed the limiting value for  $t \rightarrow t_0^+$ . To prove this, assume  $\tilde{\mathcal{F}}'(t_0)$  is discontinuous at  $t_0$ , i.e.

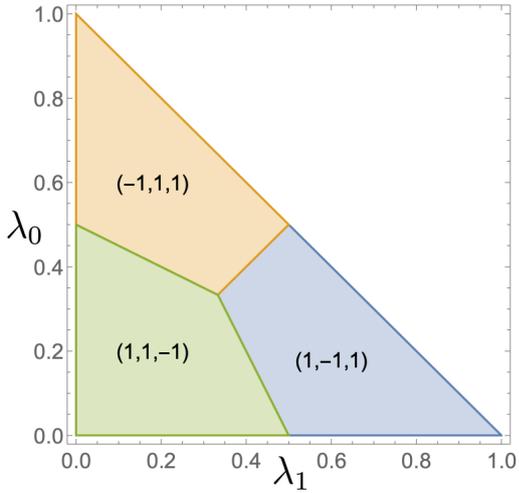
$$\lim_{t \rightarrow t_0^-} \tilde{\mathcal{F}}'(t) \neq \lim_{t \rightarrow t_0^+} \tilde{\mathcal{F}}'(t) \quad (4.62)$$



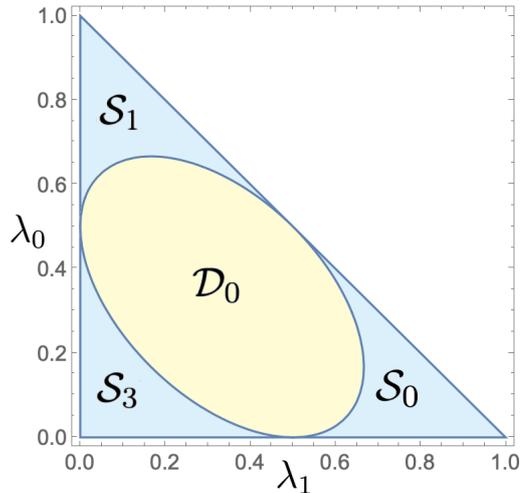
(a) Functional when minimizing over  $\sigma = \{-1, 1\}^D$



(b) Functional when minimizing over general  $\eta \in U(1)^D$



(c) Subdomains of the  $\lambda$ -simplex where  $\mathcal{F}$  is analytical. This corresponds to Fig. 4.5 in cartesian coordinates.



(d) Subdomains of the  $\lambda$ -simplex, as defined in (4.44) and (4.42). This corresponds to Fig. 4.1a in cartesian coordinates.

Figure 4.6: Functional for the Hubbard model for two fermions of total momentum  $P = 1$  populating six sites. The  $\eta \in U(1)^D$  case is analogous to a functional found for a different system investigated in [14].

We recall that each of the  $2^L$  choices of  $\sigma_\alpha$  in (4.59) represents a choice of smooth function

$$\tilde{\mathcal{F}}_\sigma[\hat{\gamma}] = \frac{V}{2L} \left| \sum_{\alpha \in \Omega_P} \sqrt{\lambda_\alpha} \sigma_\alpha \right|^2. \quad (4.63)$$

Therefore,  $\tilde{\mathcal{F}}(t)$  can be non-analytic at  $t_0$  only if

$$\begin{aligned} \tilde{\mathcal{F}}(t) &= \tilde{\mathcal{F}}_\sigma(t) \text{ for } t < t_0, \\ \tilde{\mathcal{F}}(t) &= \tilde{\mathcal{F}}_{\sigma'}(t) \text{ for } t > t_0 \end{aligned} \quad (4.64)$$

with  $\sigma'_\alpha \neq \sigma_\alpha$ . Remembering that the choices stem from the minimization in (4.59), it is clear that we must have

$$\begin{aligned} \tilde{\mathcal{F}}_{\sigma'_\alpha}(t) &\leq \tilde{\mathcal{F}}_{\sigma_\alpha}(t) \text{ for } t > t_0 \\ \tilde{\mathcal{F}}_{\sigma'_\alpha}(t) &\geq \tilde{\mathcal{F}}_{\sigma_\alpha}(t) \text{ for } t < t_0 \end{aligned} \quad (4.65)$$

Which by smoothness implies

$$\tilde{\mathcal{F}}'_{\sigma'_\alpha}(t_0) \leq \tilde{\mathcal{F}}'_{\sigma_\alpha}(t_0). \quad (4.66)$$

It follows from (4.62) and (4.64) that this inequality is strict, which, combined again with (4.64) implies the claim (4.61).

## Chapter 5

# Tweaking the Interaction

In this chapter, we study how the functional plotted in (4.6b) changes when the interaction Hamiltonian  $\hat{W}$  differs from the Hubbard interaction (4.1). In particular, we classify all possible ways the functional can fail to be strictly convex. This chapter only deals with the Levy functional (2.29), and the minimization is carried out over the entire Hilbert space of wave functions, not only those with real coefficients as in section 4.5. The last chapter deals with the real-coefficient functional in much greater detail.

In the previous chapter, we have seen that the Levy functional for the Hubbard on-site interaction with  $N = 2$  can be found analytically for an arbitrary finite one-dimensional lattice. We want to investigate how the analytical structure of  $\mathcal{F}$  changes if specific tweaks are made to the interaction itself. To this end, we continue in the setting of two fermions on an  $L$ -site lattice, but we drop the Hubbard on-site repulsion as the interaction of choice. However, we continue to dictate that  $\hat{W}$  is a translation-invariant two-body interaction that respects spin symmetry so that we can continue to restrict our attention to spin singlets with total momentum  $P$ , and so that the functional domain is again the simplex (4.11). These simplifications also require that  $\hat{h}$  respects these symmetries as before, so it may be assumed diagonal as outlined in section 2.10.

### 5.1 Degeneracy Structure for $D = 3$

We consider the case where the subspace dimension is  $D = 3$ . For specificity, we may take the space in question to be the  $P = 2$  subspace of the  $L = 5$  chain discussed previously and adopt the notation

$$|1\rangle \equiv |0 \uparrow, 2 \downarrow\rangle, \quad |2\rangle \equiv |1 \uparrow, 1 \downarrow\rangle, \quad |3\rangle \equiv |3 \uparrow, 4 \downarrow\rangle. \quad (5.1)$$

We have seen that the only one-particle Hamiltonian leading to a degenerate ground state for the simple Hubbard on-site interaction is  $\hat{h} = h\hat{I}$ . We want to know whether richer behavior is possible for more general  $\hat{W}$ .

Recall from section 2.9 that the Levy functional  $\mathcal{F}_G$  is strictly convex, except on subsets that are images of degenerate subspaces under the partial trace map. We want to see what these subsets can look like. In particular, we want to characterize the set of single-particle Hamiltonians  $\hat{h}$  that lead to a degenerate ground state. We want to classify the forms this set can take based on the properties of  $\hat{W}$ .

In general, the linear operator  $\hat{H} : \mathbb{C}^3 \mapsto \mathbb{C}^3$  will have a degenerate eigenvalue if and only if it has the form

$$\hat{H} = \hat{W} + \hat{h} = a|\phi\rangle\langle\phi| + b\hat{I} \quad (5.2)$$

for some rank-one projector  $|\phi\rangle\langle\phi|$  and  $a, b \in \mathbb{R}$ . We parameterize the interaction as

$$\hat{W} = \begin{pmatrix} 0 & w_3 & w_2 \\ w_3^* & 0 & w_1 \\ w_2^* & w_1^* & 0 \end{pmatrix}, \quad (5.3)$$

where we have assumed the diagonal elements to vanish since they can be absorbed into  $\hat{h}$ . Only the first term in (5.2) can contribute to the off-diagonal elements of  $\hat{W}$ . Writing  $|\phi\rangle = \phi_1|1\rangle + \phi_2|2\rangle + \phi_3|3\rangle$ , it follows that

$$w_3 = a\phi_1\phi_2^*, \quad w_2 = a\phi_1\phi_3^*, \quad w_1 = a\phi_2\phi_3^*. \quad (5.4)$$

To solve for  $\phi_i$ , we distinguish three cases:

*Case 1: All off-diagonal entries are non-zero,  $w_1, w_2, w_3 \neq 0$ .*

In this case, (5.4) is easily solved, yielding

$$\phi_1 = \sqrt{\frac{w_2 w_3}{a w_1}}, \quad \phi_2 = \sqrt{\frac{w_3 w_1}{a w_2}}, \quad \phi_3 = \sqrt{\frac{w_1 w_2}{a w_3}}. \quad (5.5)$$

We do not assume that  $w_{1,2,3} \in \mathbb{R}$ , and either branch of the square root may be taken in the above. It then follows from (5.2) that  $h_i = a|\phi_i|^2 + b$  defines the unique  $\hat{h}$  leading to a degenerate ground state.

*Case 2:  $w_1, w_2 \neq 0, w_3 = 0$  (+ permutations)*

This contradicts (5.4), so there is no  $\hat{h}$  satisfying (5.2).

*Case 3:  $w_1 \neq 0, w_2 = w_3 = 0$  (+ permutations)*

It is clear from (5.4) that  $\phi_1 = 0$  and  $\phi_3 = w_1/a\phi_2$ , so that

$$\hat{H} = \begin{pmatrix} b & 0 & 0 \\ 0 & b + a|\phi_2|^2 & w_1 \\ 0 & w_1 & b + w_1^2/a|\phi_2|^2 \end{pmatrix} \equiv \begin{pmatrix} b & 0 \\ 0 & b\hat{I}_2 + \hat{A} \end{pmatrix} \quad (5.6)$$

It is easy to see that  $\hat{A}$  is singular, with  $a|\phi_2|^2 + w_1^2/a|\phi_2|^2$  the only non-zero eigenvalue. Thus  $\hat{H}$  has one doubly degenerate eigenvalue  $b$  and one non-degenerate eigenvalue  $b + a|\phi_2|^2 + w_1^2/a|\phi_2|^2$ . It is also clear that

$$\hat{h} = \begin{pmatrix} b & 0 & 0 \\ 0 & b + a|\phi_2|^2 & 0 \\ 0 & 0 & b + w_1^2/a|\phi_2|^2 \end{pmatrix}, \quad (5.7)$$

leaving  $b$  and  $a|\phi_2|^2$  as degrees of freedom in  $\hat{h}$ . For the ground state to be degenerate, we need the degenerate eigenvalue to be the lower eigenvalue, which is true if and only if  $a > 0$ . Because adding a constant to the Hamiltonian is trivial, we also require  $\hat{h}$  to be traceless, yielding  $h_1 = b = -\frac{1}{3}(a|\phi_2|^2 + w_1/a|\phi_2|^2)$ . Hence, the ground state is degenerate if and only if  $h_1 < 0$ . A more practical parameterization is, therefore,

$$\hat{h} = \begin{pmatrix} h_1 & 0 & 0 \\ 0 & 1 - h_1 - h_3 & 0 \\ 0 & 0 & h_3 \end{pmatrix}, \quad (5.8)$$

rendering the full Hamiltonian in the form

$$\hat{H} = \begin{pmatrix} h_1 & 0 & 0 \\ 0 & 1 - h_1 - h_3 & w_1 \\ 0 & w_1 & h_3 \end{pmatrix}. \quad (5.9)$$

From (5.7) it follows that

$$h_1 = \frac{1}{3} \left( 1 - (h_3 - h_1) - \frac{w_1^2}{h_3 - h_1} \right), \quad (5.10)$$

which yields

$$h_1 = \frac{1}{4} \left( 1 + h_3 \pm \sqrt{9h_3^2 - 6h_3 + 1 + 8w_1^2} \right). \quad (5.11)$$

The ground state is degenerate if and only if  $h_1 < 0$ , so we are interested in the lower branch of the square root. Its graph is shown in Fig. 5.1b.

## 5.2 Functional for $D = 3$

We now compute the functional (2.60) for the Hamiltonians considered in section 5.1, treating each of the three cases in turn.

*Case 1: Using (5.5), we can rewrite the interaction (5.3) as*

$$\hat{W} = a|\phi\rangle\langle\phi| + a \begin{pmatrix} |\phi_1|^2 & 0 & 0 \\ 0 & |\phi_2|^2 & 0 \\ 0 & 0 & |\phi_3|^2 \end{pmatrix}. \quad (5.12)$$

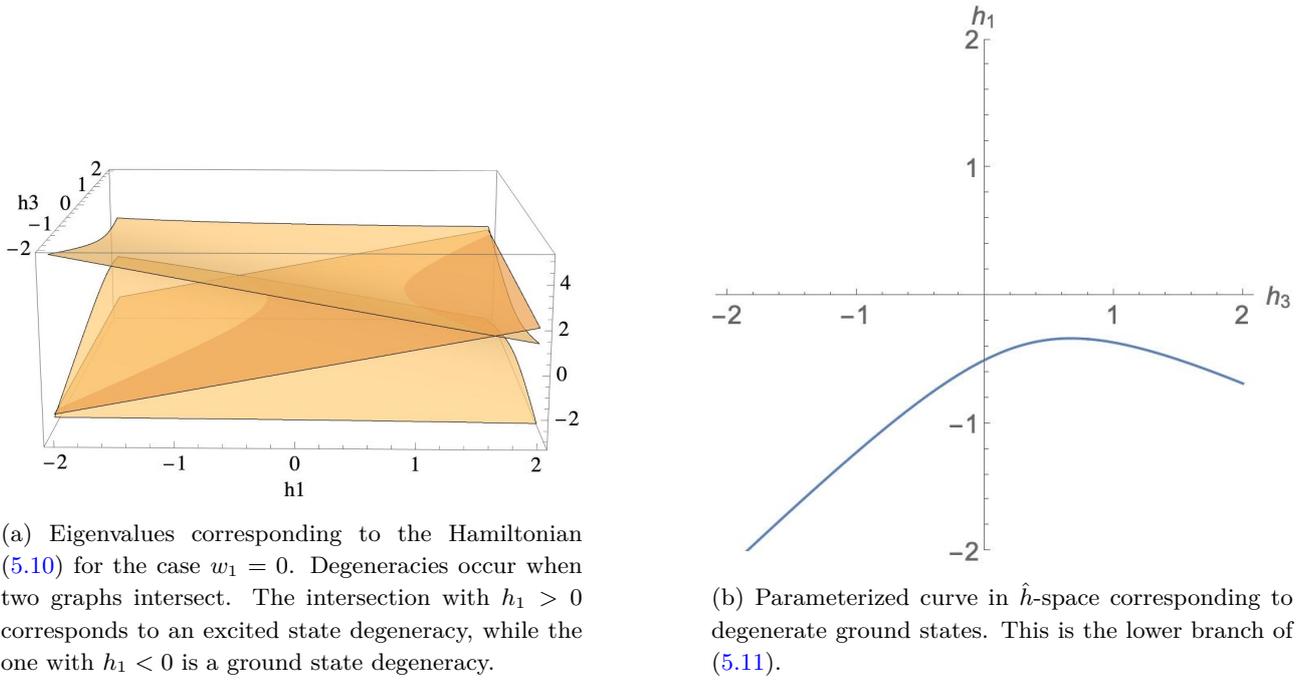


Figure 5.1

Writing the general wave function ansatz

$$|\psi\rangle = \eta_1 \sqrt{\lambda_1} |1\rangle + \eta_2 \sqrt{\lambda_2} |2\rangle + \eta_3 \sqrt{\lambda_3} |3\rangle \quad (5.13)$$

in terms of the basis (5.1), we get the expectation value

$$\langle \psi | \hat{W} | \psi \rangle = a \left( |\langle \phi | \psi \rangle|^2 + \sum_{\alpha=1}^3 |\phi_\alpha|^2 \lambda_\alpha \right). \quad (5.14)$$

The second term is independent of the  $\eta_\alpha$  and therefore irrelevant for the minimization in (2.60). We abbreviate the second term as  $\mathcal{C}(\boldsymbol{\lambda})$ . Expanding  $|\psi\rangle$  in the first term, the functional follows as

$$\mathcal{F}[\boldsymbol{\lambda}] = \min_{\boldsymbol{\eta} \in U(1)^3} \left\{ \left| \sum_{\alpha=1}^3 \sqrt{a} \phi_\alpha \sqrt{\lambda_\alpha} \eta_\alpha \right|^2 \right\} + \sum_{\alpha=1}^3 |\phi_\alpha|^2 \lambda_\alpha \equiv \mathcal{A}(\boldsymbol{\lambda}) + \mathcal{C}(\boldsymbol{\lambda}). \quad (5.15)$$

The sum inside the braces is just identical to (4.15) with  $\sqrt{\lambda_\alpha}$  replaced by  $\sqrt{a} \phi_\alpha \sqrt{\lambda_\alpha}$ . If the entries of  $\hat{W}$  are complex, so are the  $\phi_i$ , but this can be undone by redefining  $\eta_\alpha \rightarrow \eta_\alpha \phi_\alpha^*$ . The minimum is thus reached by

$$\mathcal{A}(\boldsymbol{\lambda}) = \begin{cases} a \left( |\phi_\beta| \sqrt{\lambda_\beta} - \sum_{\alpha \neq \beta} |\phi_\alpha| \sqrt{\lambda_\alpha} \right)^2 & \text{if } |\phi_\beta| \sqrt{\lambda_\beta} > \sum_{\alpha \neq \beta} |\phi_\alpha| \sqrt{\lambda_\alpha} \text{ for some } \beta \\ 0 & \text{otherwise} \end{cases} \quad (5.16)$$

The functional (5.16) is similar in form to the Hubbard functional (4.17). It only differs by the coordinate rescaling  $\sqrt{\lambda_\alpha} \rightarrow \sqrt{a} |\phi_\alpha| \sqrt{\lambda_\alpha}$ , and the additional linear term  $\mathcal{C}(\boldsymbol{\lambda})$ . This preserves the analytic structure in (4.17). The domain still decomposes into three corner regions where the functional is strictly convex and a central ellipse where the functional is affine. The  $\mathcal{A}$ -part of the functional is depicted in Fig. 5.2.

*Case 2:* The interaction Hamiltonian reduces to

$$\hat{W} = \begin{pmatrix} 0 & 0 & w_2 \\ 0 & 0 & w_1 \\ w_2^* & w_1^* & 0 \end{pmatrix}; \quad (5.17)$$

with the ansatz (5.13), we obtain

$$\langle \psi | \hat{W} | \psi \rangle = 2\text{Re} \left( w_1 \eta_2 \eta_3^* \sqrt{\lambda_2 \lambda_3} + w_2 \eta_1 \eta_3^* \sqrt{\lambda_1 \lambda_3} \right). \quad (5.18)$$

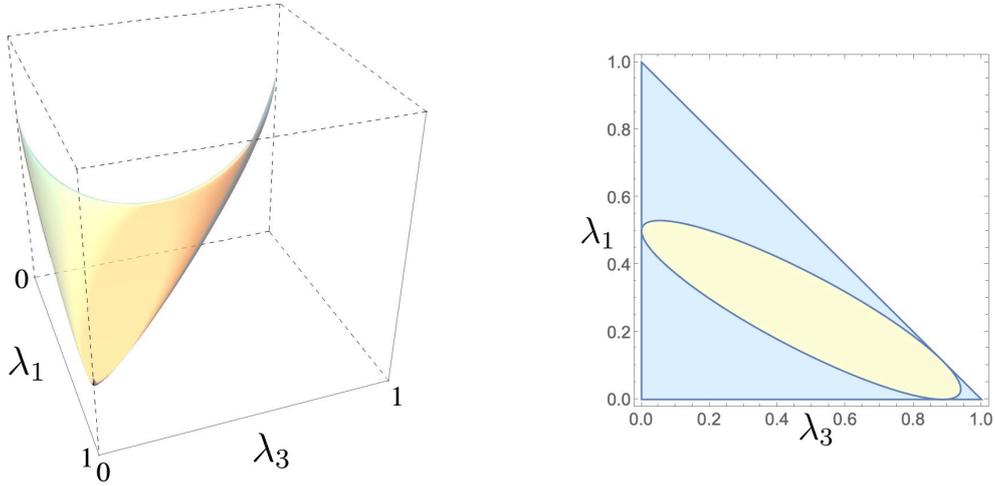
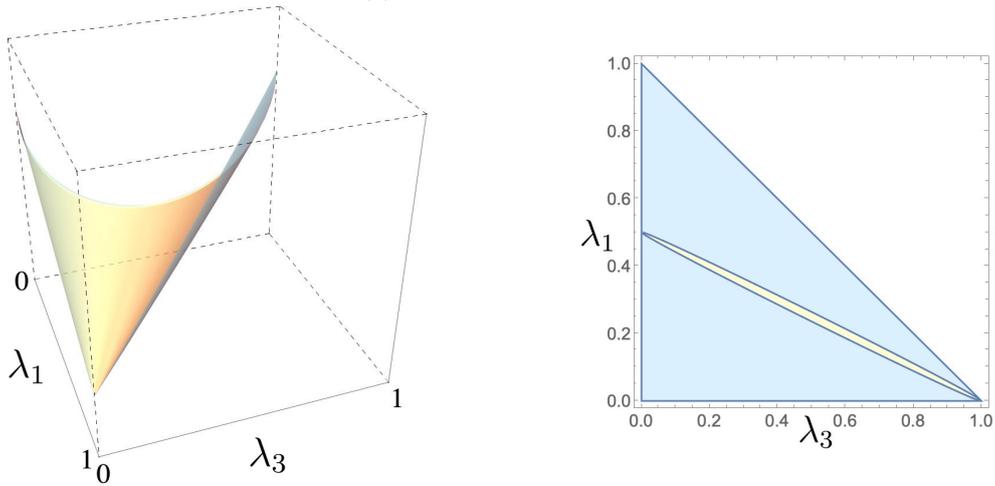
(a)  $w_1 = 1, w_2 = w_3 = 0.4$ (b)  $w_1 = 1, w_2 = w_3 = 0.2$ 

Figure 5.2:  $\mathcal{A}(\lambda)$  as given by (5.16), for two realizations of  $\hat{W}$ . On the right-hand side, the yellow regions consist of all occupation number vectors corresponding to degenerate ground states, manifesting as flat pieces of the graph of  $\mathcal{A}$ . For  $w_1 = w_2 = w_3 = 1$ , the functional is, up to a constant, identical to the one for the Hubbard interaction shown in Fig. 4.6b. The yellow region vanishes for  $w_1 = 1, w_2 = w_3 = 0$ , reducing the function to that depicted in Fig. 5.4a.

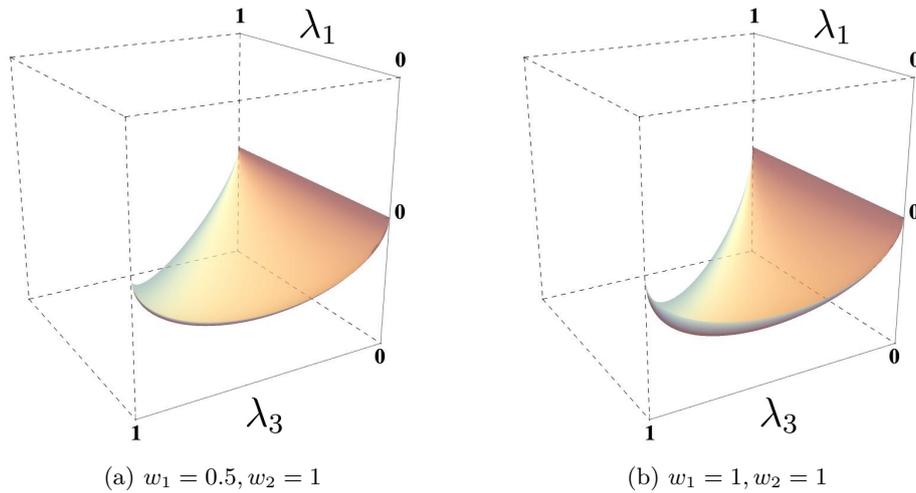


Figure 5.3: Functional (5.19) for the interaction Hamiltonian of the form (5.17). The functional is strictly convex everywhere, reflecting the absence of degeneracies as predicted for *Case 2* in section 5.1.

For  $\lambda$  fixed, the minimum is reached by adjusting  $\eta_1$  and  $\eta_2$  such that both terms are real and negative, leaving  $\eta_3$  arbitrary. The functional is thus

$$\mathcal{F}_G[\lambda] = -2\sqrt{\lambda_3} \left( |w_1|\sqrt{\lambda_2} + |w_2|\sqrt{\lambda_1} \right) = -2\sqrt{\lambda_3} \left( |w_1|\sqrt{1 - \lambda_1 - \lambda_3} + |w_2|\sqrt{\lambda_1} \right). \quad (5.19)$$

It is depicted in Fig. 5.3. The functional is strictly convex everywhere, implying there is no choice of  $\hat{h}$  that leads to a degenerate ground state. This is consistent with what we found in section 5.1.

*Case 3:* With the interaction of the simple form

$$\hat{W} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & w_1 \\ 0 & w_1^* & 0 \end{pmatrix}, \quad (5.20)$$

we once again use (5.13) to calculate

$$\langle \psi | W | \psi \rangle = (w_1 \eta_2^* \eta_3 + w_1^* \eta_3^* \eta_2) \sqrt{\lambda_2 \lambda_3}. \quad (5.21)$$

For fixed  $\lambda$ , this is minimized by setting  $\eta_2 = -\eta_3 e^{i \arg w_1}$ , yielding the functional

$$\mathcal{F}[\lambda] = -2w_1 \sqrt{\lambda_2 \lambda_3}. \quad (5.22)$$

This functional has a new form we haven't encountered before, and we will illustrate the minimization graphically as we did for the functional (4.17). To this end, we convert from barycentric to cartesian coordinates as described in section 4.2.1. Since this is a continuation of the treatment in 5.1, we pick  $\lambda_2$  as the coordinate we eliminate by normalization. This gives

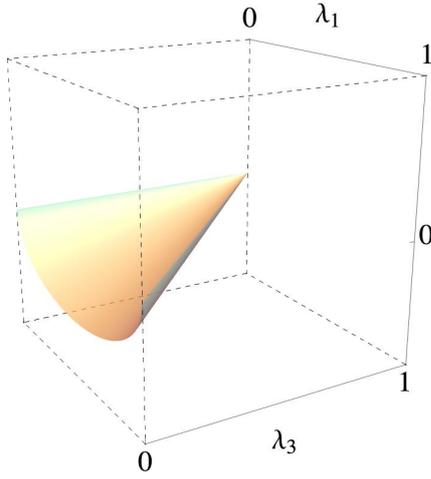
$$\tilde{\mathbf{h}}' = \begin{pmatrix} h_1 - h_2 \\ h_3 - h_2 \end{pmatrix} = \begin{pmatrix} 2h_1 + h_3 - 1 \\ 2h_3 + h_1 - 1 \end{pmatrix} \quad (5.23)$$

for the vector parameterizing the kinetic energy operator; in the second step we have used the aforementioned  $\text{Tr}[\hat{h}] = 0$  constraint we introduced in section 5.1. The functional in cartesian coordinates reads

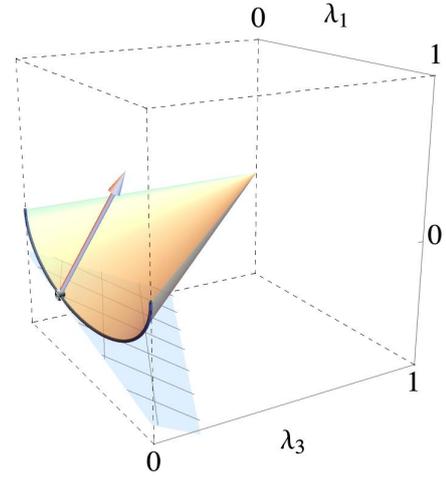
$$\mathcal{F}[\mathcal{M}_2^{-1}(\lambda')] = -2w_1 \sqrt{(1 - \lambda'_1 - \lambda'_3)\lambda'_3}, \quad (5.24)$$

where  $\mathcal{M}_2$  is the transformation from barycentric to cartesian coordinates as in (4.32). Note that if we pick  $\lambda'_0 \in \Delta^{D-1}$  at will and define the parametric curve

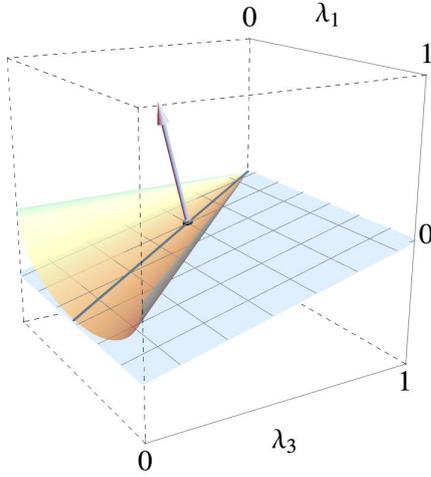
$$\lambda'(c) = \begin{pmatrix} 1 - c(2\lambda'_{0,3} + \lambda'_{0,1} - 1) \\ c\lambda'_{0,3} \end{pmatrix}, \quad (5.25)$$



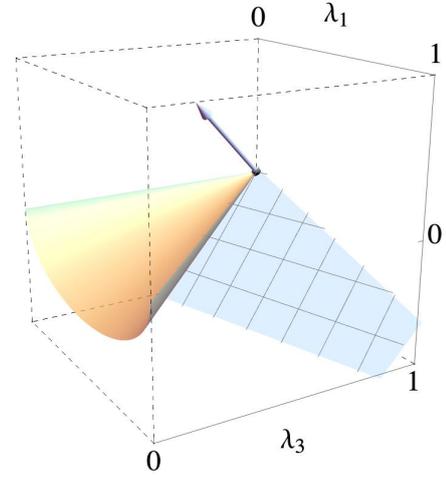
(a) Functional for the Hamiltonian (5.9). Unlike the functional in Fig. 4.3, its graph does not have a single 2-dimensional flat piece. Rather, the graph consists of a parameterized family of 1-dimensional flat pieces.



(b) Nondegenerate ground state problem. The ground state energy is the lesser eigenvalue of the lower-right block in (5.9). The ground state occupation numbers on the blue curve correspond to all such choices of  $\mathbf{h}'$ .



(c) Degenerate ground state problem. The eigenvalue  $h_1$  in (5.9) is doubly degenerate. The blue line corresponds to all ground state densities corresponding to the same  $\mathbf{h}'$ .



(d) Nondegenerate ground state problem. Here, the choice of  $\mathbf{h}$  makes  $h_1$  the unique lowest eigenvalue of (5.9). All such Hamiltonians have the same ground state  $|1\rangle$ .

Figure 5.4: Ground state problem for the Hamiltonian (5.9). The arrow has components  $(h'_1, h'_3, 1)$ . The ground state is signified by the  $(\lambda_1, \lambda_3)$ -coordinates of the gray dot.

the functional satisfies  $\mathcal{F}[\boldsymbol{\lambda}'(c)] = c\mathcal{F}[\boldsymbol{\lambda}']$ , showing that  $\mathcal{F}$  is linear along parameterized lines of the form (5.25). If  $\mathbf{h}'$  is antiparallel to  $\partial\mathcal{F}/\partial\boldsymbol{\lambda}'$  at some  $\boldsymbol{\lambda}'$ , all occupation numbers along the line containing  $\boldsymbol{\lambda}'$  solve (4.27), and thus correspond to some ground state configuration. This is illustrated in fig. 5.4c. In fact, the gradient of  $\mathcal{F}$  is

$$\frac{\partial\mathcal{F}}{\partial\boldsymbol{\lambda}'} = w_1 \left( \begin{array}{c} \sqrt{\frac{\lambda_3}{1-\lambda_1-\lambda_3}} \\ \sqrt{\frac{\lambda_3}{1-\lambda_1-\lambda_3}} - \sqrt{\frac{1-\lambda_1-\lambda_3}{\lambda_3}} \end{array} \right), \quad (5.26)$$

and a necessary condition for  $\tilde{\mathbf{h}}'$  to be antiparallel is

$$\tilde{h}'_3 = \tilde{h}'_1 + \frac{w_1^2}{\tilde{h}'_1}. \quad (5.27)$$

Using (5.23), this condition is equivalent to (5.10), wherein we established the degeneracy structure without computing the functional first.

## 5.3 Analytical Description of $\mathcal{R}$

Recall from section 2.9 that  $\mathcal{F}$  is affine on a subset  $\mathcal{R} \subset \mathcal{P}_N^1$  if and only if there exists a one-body operator  $\hat{h}$  such that  $\mathcal{R} = \text{Tr}_{N-1}[\mathcal{H}_{\text{GS}}]$ , where  $\mathcal{H}_{\text{GS}}$  is the ground-state manifold of the Hamiltonian  $\hat{W} + \hat{h}$ .

Characterizing such regions is important to understand how the analytic structure of the functional relates to the Hamiltonian's physical properties. Moreover, the appearance of such regions is of interest for computational purposes. Within such a region  $\mathcal{R}$ , the Levy and Valone functionals may differ. Since the Levy functional contains no more information about the system's physics than the Valone functional, the Levy functional can be chosen arbitrarily on  $\mathcal{R}$ , so long as it maintains the Valone functional as its lower convex envelope. However, as we shall see later, some choices of Levy functional are markedly easier to compute than others.

In figure 5.2, the region  $\mathcal{R}$  where  $\mathcal{F}$  is affine appears to have an elliptical shape. This, however, is not apparent from the analytical expression (5.16) of the functional. In this section, we derive the form of  $\mathcal{R}$  differently. The approach we will be taking is to start with an arbitrary  $n$ -dimensional subspace  $\mathcal{H}_{\text{GS}}$  and then compute the image  $\text{Tr}_{N-1}[\mathcal{H}_{\text{GS}}]$ .

We continue to work within a symmetry subspace  $\mathcal{H}_G$  and assume that the symmetries are restrictive enough so that the 1RDMs may be assumed diagonal (c.f. section 2.4). Let  $\{|\varphi\rangle\}_{\varphi=1,2,\dots}$  be the canonical (symmetry-adapted) basis for  $\mathcal{H}_G$ . We here continue to assume a linear one-to-one correspondence between the wave function coefficients' magnitudes and the occupation numbers, as is the case for the particle number  $N = 2$  treated extensively in the previous chapter. One may thus think of  $|\varphi\rangle$  as the states  $|\alpha \uparrow, \alpha \downarrow\rangle$  and  $\frac{1}{\sqrt{2}}(|P - \alpha \uparrow, \alpha \downarrow\rangle + |\alpha \uparrow, P - \alpha \downarrow\rangle)$  constituting the basis (4.5).

Let  $\{|\psi_i\rangle\}_{i=1}^n$  be a basis for the ground-state manifold  $\mathcal{H}_{\text{GS}}$ . We define  $\psi_{i,\varphi}$  to be the expansion coefficients of the ground-state basis in terms of the canonical basis:

$$|\psi_i\rangle = \sum_{\varphi=1}^D \psi_{i,\varphi} |\varphi\rangle; \quad i = 1, \dots, n. \quad (5.28)$$

An arbitrary ground state  $|\psi\rangle$  may then be expressed as

$$|\psi\rangle = \sum_{i=1}^n u_i |\psi_i\rangle; \quad \sum_i |u_i|^2 = 1; \quad \arg u_1 = 0. \quad (5.29)$$

We have written the last constraint, usually left implicit, to emphasize that  $|\psi\rangle$  represents the same physical state independent of any global phase shift. Mathematically, an  $n$ -component complex vector  $|\psi\rangle \sim \mathbf{u}$  subject to the constraints in (5.29) is an element of the  $(n-1)$ -dimensional complex projective space  $\mathbb{C}\mathbb{P}^{n-1}$ . (Strictly speaking, any quantum mechanical Hilbert space is a projective Hilbert space for this reason.) The subset  $\mathcal{R}$  is, therefore, the image of the map

$$\begin{aligned} \text{Tr}_{N-1} : \mathbb{C}\mathbb{P}^{n-1} &\rightarrow \mathcal{P}_N^1 \\ |\psi\rangle &\mapsto \boldsymbol{\lambda}; \quad \lambda_\varphi = |\langle \psi | \varphi \rangle|^2. \end{aligned} \quad (5.30)$$

Hence,  $\mathcal{R}$  is the image of the complex projective space under a quadratic map.

### 5.3.1 Two-Fold Degeneracies

Characterizing the image of (5.30) is nontrivial in general. This is evidenced, for example, by the recursion relation (4.22) valid in the case of the Hubbard model. As an instructive example, we consider here the set in the  $D = 3$  case seen in Fig. 5.2, which is the particular case of  $\dim[\mathcal{H}_{\text{GS}}] = 2$  and  $\dim[\mathcal{H}_G] = 3$ . To make things more interesting, we generalize to the case of arbitrary dimension  $\dim[\mathcal{H}_G] = D$ . The  $\dim[\mathcal{H}_{\text{GS}}] = 2$  case is highly relevant because it indicates a crossing of the two lowest energy levels as  $\hat{h}$  is varied, which is expected to occur for most interactions  $\hat{W}$ .

The object of interest for this section is, therefore, the image of the two-dimensional subspace  $\mathcal{H}_{\text{GS}} \subset \mathcal{H}_G$ ,  $\mathcal{H}_{\text{GS}} \simeq \mathbb{C}\mathbb{P}^1$  under the occupation number mapping  $|\psi\rangle \mapsto \boldsymbol{\lambda}$ . The subspace is spanned by the orthogonal states (5.28); it is well-known that quantum states in a two-dimensional Hilbert space can be parameterized by

$$|\psi(\theta, \phi)\rangle = \cos \theta |\psi_1\rangle + \sin \theta e^{i\phi} |\psi_2\rangle. \quad (5.31)$$

The partial trace map (5.30) yields

$$\begin{aligned}
\lambda_\varphi(\theta, \phi) &= |\langle \psi(\theta, \phi) | \varphi \rangle|^2 = |\langle \varphi | \psi_1 \rangle \cos \theta + \langle \varphi | \psi_2 \rangle e^{i\phi} \sin \theta|^2 \\
&= |\psi_{1,\varphi}|^2 \cos^2 \theta + |\psi_{2,\varphi}|^2 \sin^2 \theta + 2\operatorname{Re}(\psi_{1,\varphi}^* \psi_{2,\varphi} e^{i\phi}) \cos \theta \sin \theta \\
&= \frac{1}{2}(|\psi_{1,\varphi}|^2 + |\psi_{2,\varphi}|^2) + \frac{1}{2}(|\psi_{1,\varphi}|^2 - |\psi_{2,\varphi}|^2) \cos 2\theta \\
&\quad + |\psi_{1,\varphi}| |\psi_{2,\varphi}| \cos(\phi - \arg(\psi_{2,\varphi} \psi_{1,\varphi}^*)) \sin 2\theta
\end{aligned} \tag{5.32}$$

**Proposition 5.3.1:** *The image  $\mathcal{R}$  of  $(\theta, \phi) \mapsto \lambda(\theta, \phi)$  is either an ellipsoid, the convex hull of an ellipse, or a line segment.*

*Proof:* It suffices to show that the set is, respectively, the image of a 3-sphere, a disk, or a line segment. Define the vectors  $\mathbf{a}_i$  with  $i \in \{0, 1, 2, 3\}$  as follows:

$$\begin{aligned}
\mathbf{a}_{0,\varphi} &= \frac{1}{2}(|\psi_{1,\varphi}|^2 + |\psi_{2,\varphi}|^2) \\
\mathbf{a}_{1,\varphi} &= \frac{1}{2}(|\psi_{1,\varphi}|^2 - |\psi_{2,\varphi}|^2) \\
\mathbf{a}_{2,\varphi} &= |\psi_{1,\varphi}| |\psi_{2,\varphi}| \cos(\arg(\psi_{2,\varphi} \psi_{1,\varphi}^*)) \\
\mathbf{a}_{3,\varphi} &= |\psi_{1,\varphi}| |\psi_{2,\varphi}| \sin(\arg(\psi_{2,\varphi} \psi_{1,\varphi}^*))
\end{aligned} \tag{5.33}$$

Then we can rewrite (5.32) as

$$|\psi(\theta, \phi)\rangle \mapsto \lambda(\theta, \phi) = \mathbf{a}_0 + \mathbf{a}_1 \cos 2\theta + \mathbf{a}_2 \sin 2\theta \cos \phi + \mathbf{a}_3 \sin 2\theta \sin \phi. \tag{5.34}$$

Clearly,  $\lambda$  is the image of a 3-sphere under an affine map  $\mathbb{R}^3 \rightarrow \Delta^{D-1}$  sending the three unit vectors to  $\mathbf{a}_i$ , and the origin to  $\mathbf{a}_0$ . If this map has full rank, its image will be an ellipsoid. If, on the other hand, the mapping has rank 1 or 2, the image will be the convex hull of an ellipse or a line segment, respectively.  $\square$

We can make a stronger statement if the wave function coefficients are real:

**Corollary 5.3.2:** *If  $\psi_{i,\varphi} \in \mathbb{R}$ , then  $\mathcal{R}$  is either the convex hull of an ellipse or a line segment.*

*Proof:* Clearly  $\arg(\psi_{2,\varphi} \psi_{1,\varphi}^*) = 0$ , so  $\mathbf{a}_3 = 0$ , and the above-described map can have rank at most 2.  $\square$

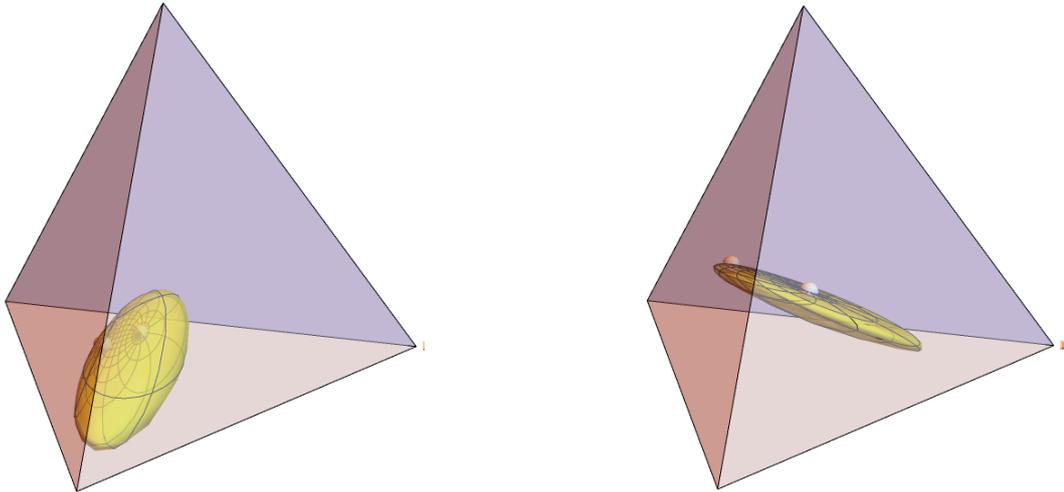
Another interesting property of  $\mathcal{R}$  is that it touches every face of the simplex.

**Corollary 5.3.3:** *For every  $\varphi \in \{1, \dots, L\}$ , there exists a  $|\psi\rangle \in S$  such that  $\lambda_\varphi(|\psi\rangle) = 0$ . In other words,  $\mathcal{R}$  touches every facet of the simplex.*

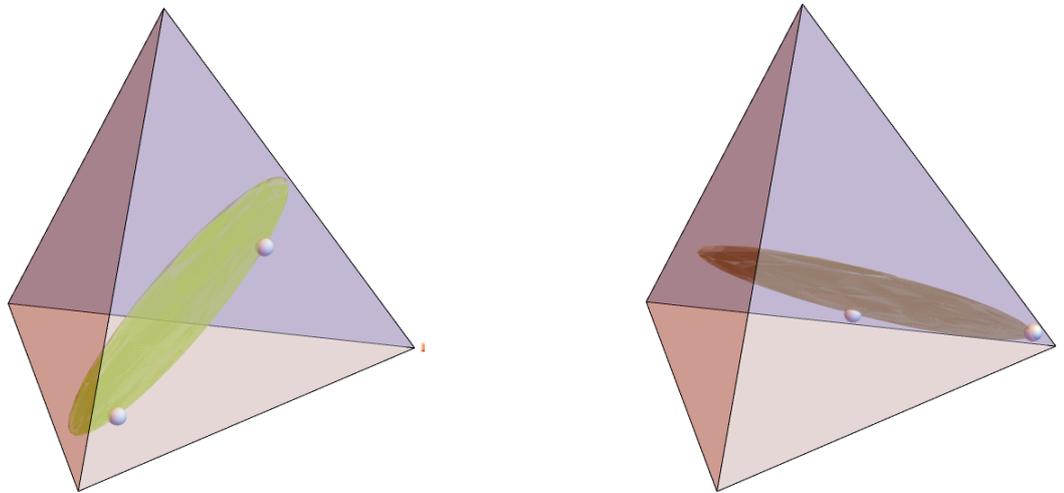
*Proof:* Referring to the first line in (5.32), it suffices to set  $\phi = \arg(\langle \varphi | \psi_1 \rangle / \langle \varphi | \psi_2 \rangle)$  and  $\theta = \arctan(\langle \varphi | \psi_1 \rangle / \langle \varphi | \psi_2 \rangle)$  to make the term vanish.  $\square$

The last corollary is rather striking: Even though the simplex has arbitrary dimension  $D$ , and  $\mathcal{R}$  can at most be a three-dimensional object,  $\mathcal{R}$  still intersects every facet of the simplex. The way this works for  $D = 3$  is shown in Fig. 5.5.

Analyses of such images of  $\mathbb{C}\mathbb{P}^{n-1}$  under the partial trace map have been carried out in the past without the restrictive assumptions we have made herein and have revealed intriguing differential-geometric behavior. We will not delve deeper into this topic but refer to the works of Penz and van Leeuwen [14, 15].



(a) complex wave function coefficients



(b) real wave function coefficients

Figure 5.5: Images of random two-dimensional subspaces  $S \subset \mathbb{C}^4$  under the mapping (5.32) from  $\mathbb{C}^4$  to  $\Delta^{D-1}$ . In general, this is a nondegenerate ellipsoid. For real wave function coefficients, however, the image is at most a (two-dimensional) ellipse. The beads mark the occupation vectors corresponding to  $|\psi_1\rangle$  and  $|\psi_2\rangle$  that span the subspace.



## Chapter 6

# Cells and Complexity of $\tilde{\mathcal{F}}$

In the previous section, we dedicated our attention to the true Levy functional  $\mathcal{F}$ , where the minimization in (2.29) or (2.60) is carried out for *all* allowed wave functions  $|\psi\rangle$ . We have noted in section 4.5 that, to obtain the correct Legendre transform, it is sufficient to restrict the minimization to wave functions with real expansion coefficients in the canonical basis. In particular, we investigated a functional of the form

$$\tilde{\mathcal{F}}[\boldsymbol{\lambda}] = \min_{\sigma_\alpha \in \{\pm 1\}^D} \langle \psi^{\sigma;\boldsymbol{\lambda}} | \hat{W} | \psi^{\sigma;\boldsymbol{\lambda}} \rangle$$

and how it compares to

$$\mathcal{F}[\boldsymbol{\lambda}] = \min_{\eta_\alpha \in U(1)^D} \langle \psi^{\eta;\boldsymbol{\lambda}} | \hat{W} | \psi^{\eta;\boldsymbol{\lambda}} \rangle.$$

In this section, we will further investigate  $\tilde{\mathcal{F}}$ . As seen in fig. 4.6a, the analytic structure of  $\tilde{\mathcal{F}}$  in the  $D = 3$  case divides the domain of  $\boldsymbol{\lambda}$  into three distinct cells. One might ask how this behavior generalizes in the case of  $L$  sites. The answer to this question has important implications for the computational viability of (4.59).

### 6.0.1 Chapter Outline:

This chapter is structured as follows: The first four sections deal with the Hubbard model. In the first section, we prove that the number of cells grows exponentially with  $D$ , and in section 6.2, we express the task of computing  $\tilde{\mathcal{F}}$  as an NP-complete computational problem. Sections 6.3 and 6.4 deal with the symmetry properties of the cells and their decomposition into further sub-cells. Finally, section 6.5 is dedicated to generalizing the results of section 6.1 to arbitrary two-fermion interactions for arbitrary system size  $L$ .

### 6.0.2 Motivation and Setup

We have seen that each cell corresponds to a different  $\boldsymbol{\sigma} \in \{-1, 1\}^D$  being the minimizer for the functional. The set corresponding to  $\boldsymbol{\sigma}$ . In general, there are  $2^D$  choices for  $\boldsymbol{\sigma}$ . If, however,  $\boldsymbol{\sigma}$  does not minimize (4.59) for *any* choice of  $\boldsymbol{\lambda}$ , it will not manifest as a distinct cell in the graph of  $\tilde{\mathcal{F}}$  (more precisely,  $\mathcal{D}_\sigma = \emptyset$ ). One may thus ask how many choices of  $\boldsymbol{\sigma} \in \{-1, 1\}^D$  are relevant for computing (4.59). Let us denote this number by  $\mathcal{N}_{\tilde{W}}$ .  $\mathcal{N}_{\tilde{W}}$  implicitly depends on  $D$  because  $\hat{W}$  is defined on a  $D$ -dimensional Hilbert space. If  $\mathcal{N}_{\tilde{W}}$  only grows polynomially with  $D$ , one may expect also to compute (4.59) in polynomial time. We will show that this is not the case.

Let us define  $\mathcal{N}_{\tilde{W}}$  more concretely. First, we must define what is meant by the region associated with  $\boldsymbol{\sigma}$ . One possible definition is

$$\mathcal{D}_\sigma = \{\boldsymbol{\lambda} \in \mathcal{D}(\tilde{\mathcal{F}}) \mid \tilde{\mathcal{F}}(\boldsymbol{\lambda}) = \tilde{\mathcal{F}}_\sigma(\boldsymbol{\lambda})\}. \quad (6.1)$$

Then  $\mathcal{D}_\sigma \neq \emptyset$  if and only if  $\boldsymbol{\sigma}$  minimizes  $\tilde{\mathcal{F}}_\sigma(\boldsymbol{\lambda})$  for *some*  $\boldsymbol{\lambda} \in \mathcal{D}(\tilde{\mathcal{F}})$ . However, this condition is insufficient. Indeed, if for all  $\boldsymbol{\lambda} \in \mathcal{D}_\sigma$ , there exists  $\boldsymbol{\sigma}' \neq \boldsymbol{\sigma}$  such that  $\tilde{\mathcal{F}}_{\boldsymbol{\sigma}'}(\boldsymbol{\lambda}) = \tilde{\mathcal{F}}_\sigma(\boldsymbol{\lambda})$ , i.e.  $\boldsymbol{\sigma}$  is a nowhere-unique minimizer, one may discard  $\boldsymbol{\sigma}'$  in the minimization and still obtain  $\tilde{\mathcal{F}}$ . We therefore define

$$\mathcal{D}_\sigma^u = \{\boldsymbol{\lambda} \in \mathcal{D}(\tilde{\mathcal{F}}) \mid \tilde{\mathcal{F}}_\sigma(\boldsymbol{\lambda}) < \tilde{\mathcal{F}}_{\boldsymbol{\sigma}'}(\boldsymbol{\lambda}) \forall \boldsymbol{\sigma}' \neq \pm\boldsymbol{\sigma}\}, \quad (6.2)$$

the set of occupation number vectors where  $\boldsymbol{\sigma}$  is a unique minimizer. The condition  $\boldsymbol{\sigma}' \neq \pm\boldsymbol{\sigma}$  is needed because  $\tilde{\mathcal{F}}_\sigma = \tilde{\mathcal{F}}_{-\sigma}$  holds everywhere. Note that the sets  $\mathcal{D}_\sigma^u$  are disjoint while the  $\mathcal{D}_\sigma$  are not. The number of

$\sigma \in \{\pm 1\}^D$  that are relevant in (4.59) is

$$\mathcal{N}_{\hat{W}} = \frac{1}{2} |\{\sigma : \mathcal{D}_{\sigma}^u(\hat{W}) \neq \emptyset\}|. \quad (6.3)$$

The factor of  $1/2$  compensates for the over-counting due to the symmetry  $\mathcal{D}_{\sigma}^u = \mathcal{D}_{-\sigma}^u$ . The remainder of this chapter will primarily be dedicated to lower-bounding  $\mathcal{N}_{\hat{W}}$  and determining the implications thereof.

## 6.1 Number of $\sigma$ -regions (Hubbard Interaction)

We now calculate  $\mathcal{N}_{\hat{W}}$  for the two-fermion Hubbard interaction, with the functional given in (4.59). For simplicity, we assume  $\zeta_{\alpha} = 1$  for all  $\alpha \in \Omega_P$ . We will show after the fact that our analysis for  $\zeta_{\alpha} = 1$  extends to general  $\zeta$ .

We will show that  $\mathcal{D}_{\sigma}^u$  is always nonempty, except when  $\sigma = (1, \dots, 1)$  and  $\sigma = (-1, \dots, -1)$ . It will be convenient to be able to reference the conditions for membership in  $\mathcal{D}_{\sigma}^u$  explicitly, so we shall state them here:

$$0 < \lambda_p < 1 \text{ for all } p \in \{1, \dots, K\}, \quad (6.4)$$

$$\sum_{\alpha \in \Omega_P} \lambda_p = 1, \quad (6.5)$$

$$\tilde{\mathcal{F}}_{\sigma}(\lambda) < \tilde{\mathcal{F}}_{\sigma'}(\lambda) \text{ for all } \sigma' \neq \pm \sigma \quad (6.6)$$

The first two conditions merely define  $\mathcal{D}(\tilde{\mathcal{F}})$ .

To begin, note that we can reorder the entries of  $\sigma$  by relabeling the basis elements  $|\alpha\rangle \equiv |\alpha \uparrow, P - \alpha \downarrow\rangle$  of the subspace of interest. To simplify notation, we choose an ordering where

$$\sigma = (\underbrace{+1, \dots, +1}_{n \text{ entries}}, \underbrace{-1, \dots, -1}_{m \text{ entries}}). \quad (6.7)$$

Let us for simplicity assume  $n, m > 0$ . It is then easy to see that there exists a  $\lambda$  satisfying (6.4) and (6.5), as well as  $\tilde{\mathcal{F}}_{\sigma}(\lambda) = 0$ . In fact, choose

$$\lambda = (\underbrace{\lambda_1, \dots, \lambda_1}_{n \text{ entries}}, \underbrace{\lambda_2, \dots, \lambda_2}_{m \text{ entries}}); \quad (6.8)$$

To achieve (6.5) and  $\tilde{\mathcal{F}}_{\sigma}(\lambda) = 0$ , one need only solve

$$\begin{aligned} n\lambda_1 + m\lambda_2 &= 1, \\ n\sqrt{\lambda_1} - m\sqrt{\lambda_2} &= 0, \end{aligned}$$

the result of which also satisfies (6.4) for all integers  $m, n > 1$ . Since  $\tilde{\mathcal{F}}_{\sigma} > 0$  by (4.63), we see that our choice (6.8) yields a minimum. In other words, we have shown (6.6) with the strict inequality relaxed to a non-strict inequality. However, we require strict inequality to ensure that  $\sigma$  is a unique minimizer; hence, we have to work a little harder.

In place of (6.8), let us try

$$\lambda = (\underbrace{\lambda_1, \dots, \lambda_1}_{k \text{ entries}}, \underbrace{\lambda_2, \dots, \lambda_2}_{n-k \text{ entries}}, \underbrace{\lambda_3, \dots, \lambda_3}_{m \text{ entries}}) \quad (6.9)$$

for some integer  $0 < k < n$  which we are free to pick. Moreover, let us fix the ratio

$$a \equiv \frac{\lambda_2}{\lambda_1}. \quad (6.10)$$

To satisfy  $\tilde{\mathcal{F}}_{\sigma}(\lambda) = 0$  and (6.5), we need

$$(k + a(n - k))\lambda_1 + m\lambda_3 = 1 \quad (6.11)$$

$$(k + (n - k)\sqrt{a})\sqrt{\lambda_1} - m\sqrt{\lambda_3} = 0, \quad (6.12)$$

which is solved by

$$\lambda_1 = \left( k + a(n - k) + \left( \frac{k + \sqrt{a}(n - k)}{m} \right)^2 \right)^{-1} \quad (6.13)$$

$$\lambda_3 = \left( \frac{k + \sqrt{a}(n - k)}{m} \right)^2 \lambda_1; \quad (6.14)$$

in particular, there exists a solution for any real  $a > 0$ . On the other hand, consider some  $\sigma' \neq \pm\sigma$ . Define new integers  $k', n', m'$  by

$$\begin{aligned} k' &= \frac{1}{2} \left( k + \sum_{p=0}^k \sigma'_p \right) \\ n' &= \frac{1}{2} \left( n + \sum_{p=0}^n \sigma'_p \right) \\ m' &= \frac{1}{2} \left( m - \sum_{p=n+1}^{D-1} \sigma'_p \right) \end{aligned}$$

In words, we obtain  $\sigma'$  by starting out with  $\sigma$  and flipping  $k'$  of the first  $k$  entries from  $+1$  to  $-1$  (and similarly for  $n'$  and  $m'$ ). We would like to derive conditions under which

$$0 = \sum_{\alpha} \sqrt{\lambda_{\alpha}} \sigma'_{\alpha} = \sum_{\alpha} \sqrt{\lambda_{\alpha}} (\sigma'_{\alpha} - \sigma_{\alpha}) = 2[(-k' - (n' - k')\sqrt{a})\sqrt{\lambda_1} + m'\sqrt{\lambda_3}]. \quad (6.15)$$

Substituting (6.14) yields

$$0 = (-k' - (n' - k')\sqrt{a})m + (k + (n - k)\sqrt{a})m'. \quad (6.16)$$

This linear equation will yield at most one solution  $a > 0$ , unless

$$(n' - k')m = (n - k)m' \quad (6.17)$$

Remember that we can freely choose  $0 < k < n$ , so there exists in particular a choice such that  $n - k$  and  $m$  are coprime, i.e. their prime number decompositions contain no nontrivial common factors. In this case the only solutions to (6.17) are  $(k', n', m') = (k, n, m)$  and  $(k', n', m') = (0, 0, 0)$ , corresponding to  $\sigma' = \pm\sigma$  which contradicts the original choice of  $\sigma'$ . Hence, every choice of  $\sigma' \neq \pm\sigma$  yields (at most) a unique solution of (6.16) and thus (6.15). Since there only exist finitely many  $\sigma'$ , we conclude there only exist finitely many choices of  $a$  for which  $\sum_{\alpha} \sqrt{\lambda_{\alpha}} \sigma'_{\alpha} = 0 = \sum_{\alpha} \sqrt{\lambda_{\alpha}} \sigma_{\alpha}$  for some  $\sigma' \neq \pm\sigma$ . Remember, however, that by (6.13) and (6.14), the condition  $\sum_{\alpha} \sqrt{\lambda_{\alpha}} \sigma_{\alpha} = 0$  alone may be satisfied for any  $a > 0$ . Therefore, there exist (in fact, uncountably many) choices of  $a$  for which  $\sqrt{\lambda_{\alpha}} \sigma_{\alpha} = 0$  but  $\sqrt{\lambda_{\alpha}} \sigma'_{\alpha} \neq 0$  for all  $\sigma' \neq \pm\sigma$ .

We have thus shown that any  $\sigma$  corresponding to (6.7) with  $n, m > 0$  is a unique minimizer of  $\tilde{\mathcal{F}}_{\sigma}(\lambda)$  for *some*  $\lambda$ , up to multiplication by  $-1$ . The number of disjoint, nonempty subdomains  $\mathcal{D}_{\sigma}$  is thus

$$\mathcal{N}_{\text{Hubbard}} = (2^D - 2)/2 = 2^{D-1} - 1. \quad (6.18)$$

The division by 2 results from the overall sign ambiguity; the subtraction corresponds to the cases  $n = 0$  and  $m = 0$ .

We still owe an explanation for why setting  $\zeta_{\alpha} = 1$  is sufficient in all of the above, which we now provide. Note that  $\zeta \rightarrow \zeta'$  amounts to a re-scaling of all the occupation numbers; if there exists  $\lambda \in \Delta^{D-1}$  such that  $\sum_{\alpha \in \Omega_P} \sqrt{\lambda_{\alpha}} \sigma_{\alpha} = 0$ , then  $\lambda'_{\alpha} = (\sum_{\beta \in \Omega_P} \lambda_{\beta} / \zeta_{\beta})^{-1} \lambda_{\alpha} / \zeta_{\alpha}$  lies in  $\Delta^{D-1}$  and satisfies  $\sum_{\alpha \in \Omega_P} \sqrt{\zeta_{\alpha}} \lambda'_{\alpha} \sigma_{\alpha} = 0$ . Similarly,  $\sum_{\alpha \in \Omega} \sqrt{\lambda_{\alpha}} \sigma'_{\alpha} > 0$  implies  $\sum_{\alpha \in \Omega_P} \sqrt{\zeta_{\alpha}} \lambda'_{\alpha} \sigma'_{\alpha} = 0$  for all  $\sigma' \neq \sigma$ .

### 6.1.1 Error Bounds for Approximations

We note that it is unfeasible to pick one particular phase combination  $\sigma$  and approximate  $\tilde{\mathcal{F}} \simeq \tilde{\mathcal{F}}_{\sigma}$ . In particular, assume we let  $\tilde{\mathcal{F}}^{\text{approx}}(\lambda) = \tilde{\mathcal{F}}_{\sigma}(\lambda)$  where

$$\sigma = \underbrace{(+1, \dots, +1)}_{n \text{ entries}} \underbrace{(-1, \dots, -1)}_{m \text{ entries}}. \quad (6.19)$$

The error made in this approximation can be rigorously lower-bounded:

**Proposition 6.1.1:**

$$\begin{aligned} \sup_{\lambda \in \Delta^{D-1}} \left| \tilde{\mathcal{F}}_{\sigma}(\lambda) - \tilde{\mathcal{F}}(\lambda) \right| &\geq \frac{V}{2L} \left( \max(n, m) - \frac{1}{\max(n, m)} \right) \geq \frac{V}{2L} \left( \frac{D}{2} - \frac{2}{D} \right) \\ &\geq \frac{V}{4} \left( 1 - \frac{16}{L^2} \right) \end{aligned} \quad (6.20)$$

*Proof:* We only need to show the first inequality since the remaining two trivially use that  $\max(n, m) \geq D/2$  and  $D \geq L/2$ . It suffices to find one  $\boldsymbol{\lambda} \in \Delta^{D-1}$  that saturates the first inequality. Since  $\tilde{\mathcal{F}}_{\boldsymbol{\sigma}} = \tilde{\mathcal{F}}_{-\boldsymbol{\sigma}}$ , we can assume without loss of generality that  $n \geq m$ . We pick

$$\boldsymbol{\lambda} = (\underbrace{1/n, \dots, 1/n}_{n \text{ entries}}, \underbrace{0, \dots, 0}_{m \text{ entries}}), \quad (6.21)$$

yielding  $\tilde{\mathcal{F}}_{\boldsymbol{\sigma}}(\boldsymbol{\lambda}) = Vn/2L$  using (4.63). On the other hand, consider the sign combination

$$\boldsymbol{\sigma}' = (\underbrace{+1, \dots, +1}_{\lceil \frac{n}{2} \rceil \text{ entries}}, \underbrace{-1, \dots, -1}_{\lfloor \frac{n}{2} \rfloor \text{ entries}}, \underbrace{\pm 1, \dots, \pm 1}_{m \text{ entries}}). \quad (6.22)$$

The last  $m$  entries are arbitrary. Again using (4.63), we have  $\tilde{\mathcal{F}}_{\boldsymbol{\sigma}'}(\boldsymbol{\lambda}) \leq V/2Ln$ . Using (4.59), we then have

$$\left| \tilde{\mathcal{F}}_{\boldsymbol{\sigma}}(\boldsymbol{\lambda}) - \tilde{\mathcal{F}}(\boldsymbol{\lambda}) \right| \geq \left| \tilde{\mathcal{F}}_{\boldsymbol{\sigma}}(\boldsymbol{\lambda}) - \tilde{\mathcal{F}}_{\boldsymbol{\sigma}'}(\boldsymbol{\lambda}) \right| \geq \frac{V}{2L} \left( n - \frac{1}{n} \right) \quad (6.23)$$

Since we have assumed  $n \geq m$ , this establishes (6.20).  $\square$

The bound established in Proposition 6.1.1 is uniform over all  $\boldsymbol{\sigma}$ , so every  $\tilde{\mathcal{F}}^{\text{approx}}$  will deviate from the true  $\tilde{\mathcal{F}}$  by at least this quantity. This makes approximations of this form effectively unfeasible and supports the claim that the complexity of  $\tilde{\mathcal{F}}$  cannot be ignored for practical purposes.

## 6.1.2 Monte Carlo Sampling

We confirm the above result numerically by uniformly sampling  $N_{\Delta}$  points in the simplex  $\Delta^{D-1}$ . This is achieved using algorithm 2, which we now explain. We begin by slightly generalizing the problem and defining the simplex

$$\Delta_a^{D-1} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}^D \mid 0 \leq \lambda_j \leq a, \sum_{j=1}^D \lambda_j = a \right\}; \quad (6.24)$$

the previous definition is, of course recovered by the special case  $\Delta^{D-1} = \Delta_{a=1}^{D-1}$ . We want to generate a random point  $(\lambda_1, \dots, \lambda_D)$  in the  $(D-1)$ -dimensional simplex. Let  $f(\boldsymbol{\lambda})$  be a probability density function. In our case, it is a uniform distribution  $f(\boldsymbol{\lambda}) = 1/|\Delta_a^{D-1}|$ , where  $|\Delta_a^{D-1}| = a^{D-1}/(D-1)!$  is the simplex's volume. The probability density function for  $\lambda_1$  is then

$$f_1(\lambda_1) = \int_0^{a-\lambda_1} d\lambda_2 \cdots \int_0^{a-\sum_{j=1}^D \lambda_j} d\lambda_D \delta(a - \lambda_1 - \cdots - \lambda_D) f(\boldsymbol{\lambda}) = \frac{|\Delta_{a-\lambda_1}^{D-2}|}{|\Delta_a^{D-1}|} = \frac{(D-1)(a-\lambda_1)^{D-2}}{a^{(D-1)}} \quad (6.25)$$

In the second expression, the denominator is the constant value of  $f(\boldsymbol{\lambda})$ , while the numerator is the volume of the domain covered by the  $d$ -fold integral. The cumulative distribution function is thus

$$F_1(\lambda_1) = \int_0^{\lambda_1} d\lambda f_1(\lambda) = 1 - \frac{(a-\lambda_1)^{D-1}}{a^{D-1}}. \quad (6.26)$$

$\lambda_1$  can then be chosen stochastically by first sampling  $x$  uniformly over some interval  $[x_0, x_1]$  and setting

$$\lambda_1 = F_1^{-1}(x) = a - a(1-x)^{1/(D-1)}. \quad (6.27)$$

The interval that  $x$  is sampled from is implicitly determined by  $F^{-1}([x_0, x_1]) = [0, a]$ , the latter being the range of  $\lambda_1$ . It is not difficult to show using (6.27) that  $x_0 = 0$  and  $x_1 = 1$ .

To sample  $\lambda_j, j = 2, \dots, d$ , we begin by assuming that the preceding entries  $\{\lambda_i\}_{i < j}$  have already been picked. It remains to sample the entries  $\{\lambda_i\}_{i \geq j}$  uniformly from the simplex  $\Delta_{a'}^{d+1-j}$  where  $a' = a - \lambda_1 - \cdots - \lambda_{j-1}$ . Applying the reasoning of eq. (6.26) then yields the conditional CDF

$$F_j(\lambda_j | \{\lambda_i\}_{i < j}) = 1 - \frac{(a' - \lambda_j)^{D-j-1}}{a'^{D-j}} \quad (6.28)$$

and the formula

$$\lambda_j = F_j^{-1}(x) = a' - a'(1-x)^{1/(D-j)}, \quad (6.29)$$

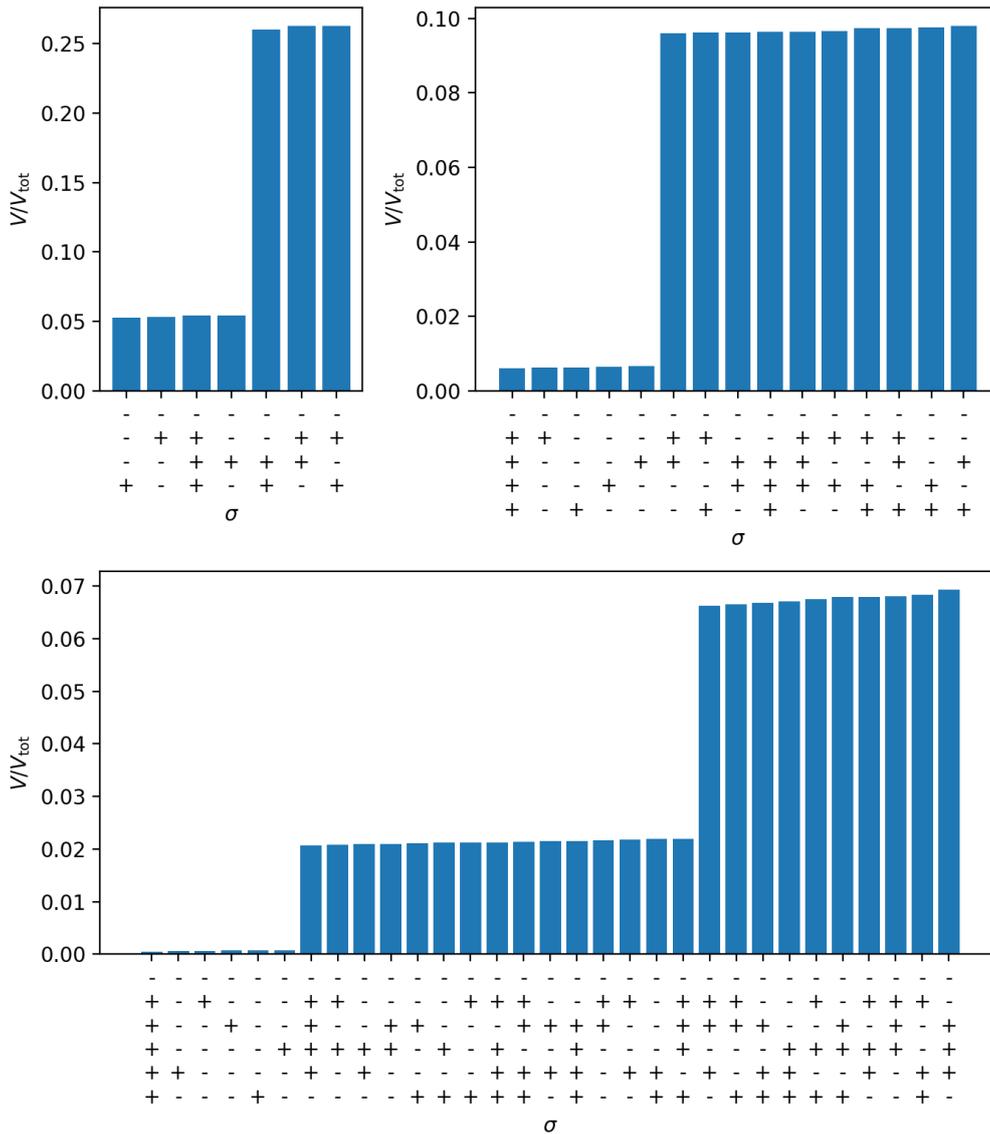


Figure 6.1: Relative volume  $V$  of  $\mathcal{D}_\sigma$  plotted over  $\sigma \in \{-1, 1\}^D$ , with  $\hat{W}$  the Hubbard on-site interaction. The number  $N_\Delta$  sampled from  $\Delta^{D-1}$  was  $10^5$ .

where again,  $x$  is a uniform random variable between 0 and 1.

After sampling  $N_\Delta$  points  $\lambda$  using the above-described method, we compute  $\tilde{\mathcal{F}}_\sigma(\lambda)$  for all  $\sigma \in \{-1, 1\}^D$  and compare the values. For each  $\sigma$ , we count the number of  $\lambda \in \Delta^{D-1}$  that fall inside  $\mathcal{D}_\sigma$ . Owing to the symmetry  $\tilde{\mathcal{F}}_\sigma(\lambda) = \tilde{\mathcal{F}}_{-\sigma}(\lambda)$ , we need only account for those  $\sigma$  where  $\sigma_1 = -1$ . As  $\lambda$  is being sampled uniformly, the expected number found in  $\mathcal{D}_\sigma \subset \Delta^{D-1}$  is proportional to the volume  $|\mathcal{D}_\sigma|$ .

The relative volumes of the cells are shown in Fig. 6.1. Observe that all cells  $\mathcal{D}_\sigma$  where  $\sigma$  has the same number of plus and minus signs have the same volume and that the volume shrinks drastically with greater sign imbalance. Also note  $\sigma = (-1, \dots, -1)$  is excluded since  $\mathcal{D}_\sigma = \emptyset$  (as is evident from (4.59)). Crucially, all other cells have finite volume as predicted by (6.18).

## 6.2 Complexity (Hubbard Interaction)

In the previous section, we showed that the number of  $\sigma$ -regions,  $|\mathcal{D}_\sigma|$ , grows exponentially with the physical system size  $L$ . However, this does not preclude that one can calculate  $\tilde{\mathcal{F}}[\lambda]$  in less than exponential time using an efficient algorithm. Here, we use the theory reviewed in chapter 3 to argue against the existence of such an algorithm.

Recall the Subset Sum problem defined in section 3.0.3 and the corresponding optimization problem from

---

**Algorithm 2** The function `RandomPoint` uniformly samples a random point from  $\Delta^{D-1}(a)$ , the  $(D-1)$ -dimensional simplex with  $\sum_{j=1}^D \lambda_j = a$ . The entries of  $\lambda$  are sampled sequentially according to (6.29).

---

```

1: function RANDOMPOINT( $D, a$ ):
2:   Let  $\lambda$  be array with  $D$  entries
3:   for  $j = 1 \dots D - 1$  do
4:      $x \leftarrow$  uniformly sampled random number between 0 and 1
5:      $a' \leftarrow a - \sum_{i=1}^{j-1} \lambda_i$ 
6:      $\lambda[j] \leftarrow a' - a'(1-x)^{1/(D-j)}$ 
7:   end for
8: end function

```

---

section 3.3. Let us start by rewriting the optimization problem in a more suitable form. We define the variables  $x_j \in \{0, 1\}$ , where  $j \in \{1, \dots, n\}$ . We use  $\{x_j\}$  to define the subsets  $U_{\mathbf{x}} = \{j \in I \mid x_j = 1\} \subset I$ . The solution to the Subset Sum optimization problem from section 3.3 can now be rewritten as

$$\max \left\{ \sum_{j \in U} a_j \mid U \subset I, \sum_{j \in U} a_j < B \right\} = \max \left\{ \sum_{j=1}^n a_j x_j \mid \mathbf{x} \in \{0, 1\}^n, \sum_{j=1}^n a_j x_j < B \right\}. \quad (6.30)$$

Now consider the  $\{\pm 1\}$ -functional (4.59), and define  $a_\alpha = \sqrt{\zeta_\alpha \lambda_\alpha}$ . Let us also use the shorthand

$$\sum_{\alpha \in \Omega_P} a_\alpha \sigma_\alpha \equiv \mathbf{a} \cdot \boldsymbol{\sigma}. \quad (6.31)$$

The minimization

$$M \equiv \frac{2L}{V} \tilde{\mathcal{F}}[\lambda] = \min_{\boldsymbol{\sigma} \in \{\pm 1\}^D} |\mathbf{a} \cdot \boldsymbol{\sigma}|^2 \quad (6.32)$$

does not yet have the form (6.30) because the norm square of the sum is taken and because the optimization is unconstrained. However, we can use the identity

$$M = \min_{\boldsymbol{\sigma} \in \{\pm 1\}^D} |\mathbf{a} \cdot \boldsymbol{\sigma}|^2 = \left[ \min_{\boldsymbol{\sigma} \in \{\pm 1\}^D} |\mathbf{a} \cdot \boldsymbol{\sigma}| \right]^2 \quad (6.33)$$

and define  $x_\alpha = \frac{1}{2}(1 - \sigma_\alpha)$  to rewrite

$$\begin{aligned} \sqrt{M} &= \min_{\boldsymbol{\sigma} \in \{\pm 1\}^D} |\mathbf{a} \cdot \boldsymbol{\sigma}| = \min \left\{ |\mathbf{a} \cdot \boldsymbol{\sigma}| \mid \boldsymbol{\sigma} \in \{-1, 1\}^D \right\} \\ &= \min \left\{ \mathbf{a} \cdot \boldsymbol{\sigma} \mid \boldsymbol{\sigma} \in \{-1, 1\}^D, \mathbf{a} \cdot \boldsymbol{\sigma} > 0 \right\} \\ &= -\max \left\{ 2\mathbf{a} \cdot \mathbf{x} - \sum_{\alpha \in \Omega_P} a_\alpha \mid \mathbf{x} \in \{0, 1\}^D, -2\mathbf{a} \cdot \mathbf{x} + \sum_{\alpha \in \Omega_P} a_\alpha > 0 \right\} \\ &= -2 \max \left\{ \mathbf{a} \cdot \mathbf{x} \mid \mathbf{x} \in \{0, 1\}^D, \mathbf{a} \cdot \mathbf{x} < \frac{1}{2} \sum_{\alpha \in \Omega_P} a_\alpha \right\} - \sum_{\alpha \in \Omega_P} a_\alpha. \end{aligned} \quad (6.34)$$

The first line is just a formal rewriting. In the second line, we remove the absolute value around the sum and introduce in its place the constraint  $\sum_{\alpha \in \Omega_P} a_\alpha \sigma_\alpha > 0$ . One can convince oneself that this tradeoff indeed yields the same minimum  $\sqrt{M}$ . In the third line, we express  $\boldsymbol{\sigma}$  in terms of the above-defined  $\mathbf{x}$ , and the fourth line is a straightforward rearrangement of terms. This optimization problem is identical in form to (6.30), with the index set  $I$  replaced by  $\Omega_P$ ,  $n$  replaced by  $D$  and  $B = \frac{1}{2} \sum_{\alpha \in \Omega_P} a_\alpha$ . We have thus established the phase dilemma (4.59) as an instance of the subset-sum problem. More importantly, we have shown that the phase dilemma reduces to the special case mentioned at the end of section 3.0.3, which is equivalent to the 2-partition problem. The only remaining constraints in (6.34) that are not explicit in 2-partition are the restrictions  $a_\alpha \in [0, \sqrt{\zeta_\alpha}]$  and  $\sum_{\alpha=1}^n a_\alpha^2 / \zeta_\alpha = 1$ , stemming from the normalization of the occupation numbers  $\lambda_\alpha = a_\alpha^2 / \zeta_\alpha$ .

### 6.2.1 NP-Completeness of the Phase Dilemma

We can now show that the Hubbard-model phase dilemma (6.34) is NP-complete. It is easy to see that the phase dilemma is in NP since it is a special case of 2-partition as explained above. Suppose now that there

exists a polynomial-time (in input size) algorithm that solves the phase dilemma. Given an arbitrary instance  $\{a_\alpha\}_{\alpha=1}^n$  of 2-partition, we can define

$$a'_\alpha = \frac{a_\alpha}{\sqrt{\sum_{j=1}^n a_j^2 / \zeta_\alpha}},$$

so that  $\sum_{\alpha=1}^n (a'_\alpha)^2 / \zeta_\alpha = 1$ . Note that this automatically guarantees  $a'_\alpha \in [0, \sqrt{\zeta_\alpha}]$ . It is trivial to show that 2-partition is invariant under global multiplication by a constant, so the instances  $\{a'_\alpha\}_{\alpha=1}^n$  and  $\{a_\alpha\}_{\alpha=1}^n$  are equivalent.  $\{a'_\alpha\}_{\alpha=1}^n$  is also an instance of the phase dilemma since it satisfies the constraints, and we can use the supposed algorithm to solve it in polynomial time. By the equivalence, we have then also solved the arbitrary 2-partition instance  $\{a_\alpha\}_{\alpha=1}^n$  in polynomial time. This shows that the phase dilemma is NP-hard and, therefore, NP-complete.

Let us discuss the implications of the above conclusion. NP-completeness precludes the existence of an algorithm that is polynomial in input size. The input size, as mentioned in section 3.0.3, is  $n \log_2(N) + \log_2(nN)$ , where  $n = D$  is the number of variables and  $N$  is the maximum value of the individual numbers  $a_\alpha$  when represented as integers. In section 3.2.1, we mentioned that NP-hardness does not preclude the existence of algorithms with complexity polynomial in  $n$  and that, indeed, dynamic programming can be used to construct an algorithm that runs with  $\mathcal{O}(nB)$  arithmetic operations. For the phase dilemma,  $B = \mathcal{O}(nN)$ , so its running time will be  $\mathcal{O}(n^2N)$ . However, the phase dilemma is not an integer problem, so to apply any such algorithm, one must employ a decimal approximation. Note that the magnitude of the real numbers  $a_\alpha$  is of order unity (since they are square roots of occupation numbers). If we encode them with base 10 and demand keeping track of  $d$  decimals, then the corresponding integers will have magnitude  $N \sim 10^d$ , and the running time of the pseudo-polynomial algorithm will be  $\mathcal{O}(n^2 10^d)$ .

For sufficiently large  $d$ , the problem becomes unfeasible. In practice, the lookup table  $X(S_j, B)$ , as defined in section 3.2.1, will be so big that most entries are never referenced, and the advantage gained by dynamically storing partial results evaporates. The pseudo-polynomial algorithm then amounts to trying out all possible subsets  $U \subset \{1, \dots, n\}$ , which takes  $\sim 2^n$  operations.

A more comprehensive study of phase dilemma solution algorithms, accounting for the typically required precision and containing a more comprehensive account of pseudo-polynomial algorithms, is beyond the scope of this thesis. The result we have found is nevertheless intriguing: Even for a two-particle lattice model with maximally local interactions, solving for the functional with conventional methods is a (weakly) NP-complete problem.

### 6.3 Permutation Symmetry (Hubbard Interaction)

This section deals with the behavior of  $\tilde{\mathcal{F}}(\boldsymbol{\lambda})$  when certain linear transformations are applied to  $\boldsymbol{\lambda}$ . We will ignore the factors of  $\zeta_\alpha$  in the Levy functional (4.15). If we were to take the factors into account, we would have to modify the action  $\boldsymbol{\lambda} \mapsto P\boldsymbol{\lambda}$  (defined below) by multiplying the appropriate entries of  $\boldsymbol{\lambda}$  by  $\zeta_\alpha$  or  $\zeta_\alpha^{-1}$ .

Our goal is to show that the  $\sigma$ -regions defined in the previous section are symmetric under certain subgroups of the permutation group. To do so, we need a few additional definitions. We denote by  $S_D$  the group of permutations acting on  $D$ -tuples. The action of  $P \in S_D$  on phase or occupation number vectors is

$$\begin{aligned} (P\boldsymbol{\lambda})_\alpha &= \boldsymbol{\lambda}_{P(\alpha)} \\ (P\boldsymbol{\sigma})_\alpha &= \boldsymbol{\sigma}_{P(\alpha)}. \end{aligned} \tag{6.35}$$

We also define the action of permutations on  $\mathcal{D}_\sigma^u$ ,

$$P(\mathcal{D}_\sigma^u) = \{P\boldsymbol{\lambda} \mid \boldsymbol{\lambda} \in \mathcal{D}_\sigma^u\}, \tag{6.36}$$

along with the stabilizer subgroup,

$$S(\mathcal{D}_\sigma^u) = \{P \in S_D \mid P(\mathcal{D}_\sigma^u) = \mathcal{D}_\sigma^u\}. \tag{6.37}$$

**Lemma 6.3.1:** (6.36) is a well-defined group action, and  $P(\mathcal{D}_\sigma^u) = \mathcal{D}_{P\sigma}^u$ .

*Proof:* We claim that

$$\boldsymbol{\lambda} \in \mathcal{D}_\sigma^u \text{ if and only if } \tilde{\mathcal{F}}_{P\sigma}(P\boldsymbol{\lambda}) < \tilde{\mathcal{F}}_{\sigma''}(P\boldsymbol{\lambda}) \text{ for all } \sigma'' \neq \pm P\sigma.$$

Upon pondering the definition (6.2), one sees that the above statement implies  $P(\mathcal{D}_\sigma^u) = \mathcal{D}_{P\sigma}^u$ . To see why the statement is true, note that  $\tilde{\mathcal{F}}_\sigma(\lambda) = \tilde{\mathcal{F}}_{P\sigma}(P\lambda)$  holds for any permutation  $P$ , so that

$$\tilde{\mathcal{F}}_{P\sigma}(P\lambda) = \tilde{\mathcal{F}}_\sigma(\lambda) < \tilde{\mathcal{F}}_{P^{-1}\sigma''}(\lambda) = \tilde{\mathcal{F}}_{\sigma''}(P\lambda) \quad (6.38)$$

holds for all  $\sigma'' \neq \pm P\sigma$  if and only if  $\lambda \in \mathcal{D}_\sigma^u$  (the inequality follows from the definition (6.2)). This shows that indeed  $P(\mathcal{D}_\sigma^u) = \mathcal{D}_{P\sigma}^u$ , which implies that (6.36) maps the collection of sets  $\{\mathcal{D}_\sigma^u\}_\sigma$  to itself, and that it is a group action.  $\square$

**Corrolary 6.3.2:**  $P \in S(\mathcal{D}_\sigma^u)$  if and only if  $P\sigma = \pm\sigma$ .

*Proof:* This follows from the definitions (6.2) and (6.36) together with the previous lemma.  $\square$

We can classify the  $\sigma$ 's using the group action of  $S_D$ . We denote by  $\Sigma^n$  the set of  $\sigma$ -vectors with  $n$  plus signs and  $D - n$  minus signs, or vice-versa:

$$\Sigma^n = \left\{ \sigma \in \{\pm 1\}^D \mid \left| \sum_\alpha \sigma_\alpha \right| = D - 2n \right\}; \quad n \leq \frac{K}{2} \quad (6.39)$$

For  $n = K/2$ ,  $\Sigma^n$  is a single  $S_D$ -orbit, while for  $n \neq K/2$ ,  $\Sigma^n$  is the union of two  $S_D$ -orbits.

**Lemma 6.3.3:** Let  $S(\sigma)$  be the stabilizer of  $\sigma \in \Sigma^n$ . Then  $S(\sigma) \simeq S_n \times S_{D-n}$ .

*Proof:*  $S_D$  acts transitively on  $\Sigma^n$ , so all stabilizers are isomorphic, and it suffices to show the claim for  $\sigma = (+, \dots, +, -, \dots, -)$ . It is clear that  $P\sigma = \sigma$  if and only if it permutes the first  $n$  and last  $D - n$  entries among each other. The subgroup of such permutations is isomorphic to  $S_n \times S_{D-n}$ .  $\square$

Now, we can identify the stabilizer of  $\mathcal{D}_\sigma^u$ . Clearly,  $\sigma$  must be in  $\Sigma^n$  for some  $n \leq K/2$ . If  $n \neq K/2$ , we cannot have  $P\sigma = -\sigma$ , since permutations preserve the number of minus signs. Corollary 6.3.2 then implies that the stabilizer of the  $\mathcal{D}_\sigma^u$  is exactly the stabilizer of  $\sigma$ . If, on the other hand,  $n = K/2$ , permutations inverting the sign of  $\sigma$  must also be accounted for. By transitivity, however, we can pick out one sign-inverting permutation  $P_-$  and express every other sign-inverting permutation  $P'_-$  as a product of  $P_-$  and a sign-preserving permutation:  $P'_- = P_-(P_-^{-1}P'_-)$ .  $P_-$  together with the identity forms the group  $\mathbb{Z}_2$ , so the group of permutations satisfying  $P\sigma = \pm\sigma$  can be expressed as the subgroup product  $\mathbb{Z}_2(S_n \times S_{D-n})$ . In summary, we have established

**Lemma 6.3.4:**

$$S(\mathcal{D}_\sigma^u) \simeq \begin{cases} \mathbb{Z}_2(S_n \times S_{D-n}) & \text{if } \sigma \in \Sigma^n \text{ where } n = \frac{K}{2} \\ S_n \times S_{D-n} & \text{otherwise. } \square \end{cases} \quad (6.40)$$

In this section, we have derived the symmetry group of  $\mathcal{D}_\sigma^u$ . In the next section, we will use this symmetry to lower-bound the connected components of  $\mathcal{D}_\sigma^u$  based on the properties of  $\sigma$ .

## 6.4 Connectedness (Hubbard Interaction)

We established at the beginning of this chapter that  $\mathcal{D}(\mathcal{F})$  is covered by (the closure of)  $\mathcal{N}_{\text{Hubbard}}$  disjoint subregions,  $\tilde{\mathcal{F}}$  being analytic in each subregion and non-analytic at region boundaries. We did not, however, establish that  $\mathcal{D}_\sigma^u$  is connected (in the sense of topology), leaving the possibility that each  $\mathcal{D}_\sigma^u$  decomposes into even more connected cells. In this section, we show that such behavior does indeed occur. This implies, in particular that the result of section 6.1 underestimates the complexity of the functional's analytic structure.

**Lemma 6.3.5:** Consider some  $\sigma \in \{\pm 1\}^D$  and  $\lambda \in \mathcal{D}(\mathcal{F})$ . If for some indices  $\alpha, \beta \in \Omega_P$ ,  $\sigma_\alpha \neq \sigma_\beta$  and  $\lambda_\alpha = \lambda_\beta$ , then  $\lambda \notin \mathcal{D}_\sigma^u$ .

*Proof:* Denote by  $P_{\alpha\beta} \in S_D$  the transposition that swaps entries  $\alpha$  and  $\beta$ . By familiar arguments, we have

$$\mathcal{F}_\sigma(\lambda) = \mathcal{F}_{P_{\alpha\beta}\sigma}(P_{\alpha\beta}\lambda) = \mathcal{F}_{P_{\alpha\beta}\sigma}(\lambda),$$

and  $P_{\alpha\beta}\sigma \neq \pm\sigma$ , so the claim directly follows from the definition of  $\mathcal{D}_\sigma^u$ .  $\square$

By (6.40), the stabilizer of  $\mathcal{D}_\sigma^u$  can be expressed in terms of permutation groups, which in turn are generated by transpositions  $P_{\alpha\beta}$ . For the moment, we will assume that  $n \neq K/2$  and consider the permutation  $(P_{\alpha\beta}, e) \in S_n \times S_{D-n} = S(\mathcal{D}_\sigma^u)$  where  $e \in S_{D-n}$  is the identity element; the reasoning for  $(e, P_{\alpha\beta})$  is completely analogous.

**Lemma 6.3.6:** *Consider  $\lambda \in \mathcal{D}_\sigma^u$  and label its entries  $\lambda_\alpha$  such that they are decreasingly ordered. Let  $P_{\alpha\beta} \in S_D$  be a permutation, and assume WLOG that  $\alpha < \beta$ . If there exists  $\delta \in \Omega_P$  such that  $\alpha < \delta < \beta$  and  $\sigma_\delta \neq \sigma_\alpha = \sigma_\beta$ , then  $\lambda$  and  $P_{\alpha\beta}\lambda$  are both in  $\mathcal{D}_\sigma^u$  and belong to different connected components of  $\mathcal{D}_\sigma^u$ .*

*Proof:* It is clear that  $P_{\alpha\beta}\lambda \in \mathcal{D}_\sigma^u$ , because  $P_{\alpha\beta}(\mathcal{D}_\sigma^u) = \mathcal{D}_{P_{\alpha\beta}\sigma}^u = \mathcal{D}_\sigma^u$ . To show they cannot be in the same connected component, consider a continuous path  $\gamma : [0, 1] \rightarrow \mathcal{D}(\mathcal{F})$  such that  $\gamma(0) = \lambda$  and  $\gamma(1) = P_{\alpha\beta}\lambda$ . Then  $\gamma_\alpha(0) < \gamma_\delta(0) < \gamma_\beta(0)$  and  $\gamma_\beta(1) < \gamma_\delta(1) < \gamma_\alpha(1)$ , implying there is  $0 < r < 1$  such that  $\gamma_\beta(r) = \gamma_\delta(r)$  by continuity. By the previous claim, it follows that  $\gamma(r) \notin \mathcal{D}_\sigma^u$ . We have thus shown that any continuous curve connecting  $\lambda$  to  $P_{\alpha\beta}\lambda$  contains a point outside  $\mathcal{D}_\sigma^u$ , so  $\lambda$  and  $P_{\alpha\beta}\lambda$  cannot lie in the same connected component.  $\square$

We will now prove that the constellations of  $\sigma \in \{\pm 1\}^D$  and  $\lambda \in \mathcal{D}_\sigma^u$  as in the above lemma materialize. However, as seen in Figs 6.2 and 6.4,  $\mathcal{D}_\sigma^u$  do indeed decompose into disconnected components, and Lemma 6.3.6 explains this behavior.

As an example, take  $D = 5$ ,  $n = 2$  and suppose that  $\lambda \in \mathcal{D}_\sigma^u$  is decreasingly ordered with

$$\sigma = (+, -, +, -, +).$$

Then  $P_{13}$ ,  $P_{35}$ ,  $P_{15}$ , and  $P_{24}$  are distinct permutations, each of which satisfies the above claim. They generate a subgroup of order eight, so we can expect  $\mathcal{D}_\sigma^u$  to have at least eight connected components. Again, we must remark that we did not prove that a decreasingly ordered  $\lambda$  exists such that  $\lambda \in \mathcal{D}_\sigma^u$  for  $\sigma = (+, -, +, -, +)$ , and the lemma only implies that if they do exist, then there are at least eight connected components.

Finally, let us look at the case  $\sigma \in \Sigma^n$  where  $n = K/2$ . In addition to the permutations in  $S(\sigma)$ , which are covered by the above lemma, we also have the permutations satisfying  $P\sigma = -\sigma$ .

**Lemma 6.3.7:** *Let  $\lambda \in \mathcal{D}_\sigma^u$  and  $P\sigma = -\sigma$ . Then  $\lambda$  and  $P\lambda$  are not in the same connected component of  $\mathcal{D}_\sigma^u$ .*

*Proof:* Pick  $\alpha, \beta$  such that  $P(\alpha) = \beta$ . Since  $P$  is sign-reversing, we must have  $\sigma_\alpha \neq \sigma_\beta$ . Also assume WLOG that  $\lambda_\alpha < \lambda_\beta$ . Again considering a continuous curve  $\gamma$  from  $\lambda$  to  $P\lambda$ , we have  $\gamma_\alpha(0) < \gamma_\beta(0)$  and  $\gamma_\alpha(1) > \gamma_\beta(1)$ . Therefore, by continuity, there is  $0 < r < 1$  such that  $\gamma_\alpha(r) = \gamma_\beta(r)$  and hence  $\gamma(r) \notin \mathcal{D}_\sigma^u$  as in the previous claim.  $\square$

Note that this last claim did not require  $\lambda$  to be ordered in any specific way, so without any additional assumptions, we can assert:

**Corollary 6.3.8:** *If  $\sigma \in \Sigma^n$  with  $n = K/2$ , then  $\mathcal{D}_\sigma^u$  has at least two connected components.*  $\square$

In reality, however, Corollary 6.3.8 leads to a doubling of the components arising from Lemma 6.3.6, corresponding to the  $\mathbb{Z}_2$  in (6.40). The true number of connected components is expected to be greater than two. For example, consider

$$\sigma = (+, -, -, +)$$

and assume  $\lambda \in \mathcal{D}_\sigma^u$  is decreasingly ordered. By Lemma 6.3.6,  $P_{14}$  produces two connected components, and Corollary 6.3.8 doubles the number to four. The case is depicted in Fig. 6.2.

Since the domain (4.11) is a  $D$ -dimensional simplex, plotting its decomposition into the  $\sigma$ -regions (6.1) is only possible up to  $D = 4$  as in fig. 6.2. Nevertheless, as the system size increases, the vast growth in the number of  $\sigma$ -regions can be observed by looking at intersections of the domain with two-dimensional subspaces. We select a random point  $\lambda_0 \in \mathcal{D}(\mathcal{F})$  and two random orthogonal vectors  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^D$  satisfying  $\mathbf{a} \cdot \mathbf{n} = \mathbf{b} \cdot \mathbf{n} = 0$  and  $\|\mathbf{a}\| = \|\mathbf{b}\| = 1$  with  $\mathbf{n}$  the normal vector to the embedded hypersimplex (4.11). Then, we intersect the

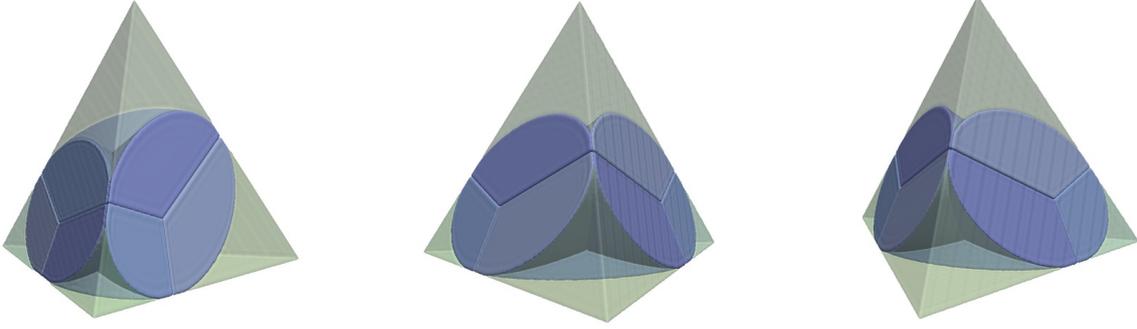


Figure 6.2:  $\sigma$ -regions in the case  $D = 4$  and  $\zeta = (1, 1, 2, 2)$ . Three different viewing angles are shown for ease of visualization. The green-shaded corner regions are  $\mathcal{D}_\sigma^u$  for  $\sigma = (+, +, +, -), (+, +, -, +), (+, -, +, +)$  and  $(-, +, +, +)$ . The blue-shaded regions correspond to  $\sigma = (+, +, -, -), (+, -, +, -)$  and  $(+, -, -, +)$ . Each of the three latter regions has four connected components; components with the same shade of blue are part of the same  $\sigma$ -region. This decomposition is explained by Lemma 6.3.6.

two-dimensional subspace parameterized by  $\lambda(x, y) = \lambda_0 + xa + yb$  with the simplex (4.11) and its subdomains (6.1). This procedure is illustrated in fig 6.3. The resulting region plots are shown in the lower half of Fig. 6.3 for  $D = 4$  and in Fig. 6.4.

## 6.5 Number of $\sigma$ -regions (2 Fermions, Arbitrary Finite-Range Interaction)

So far, we have only considered the real-phase functional (4.59) for the Hubbard on-site interaction. For the particular case of a two-fermion system, the Hubbard interaction is peculiar — looking at (4.13), we can see that  $\langle \psi | \hat{W} | \psi \rangle = V |\langle \phi | \psi \rangle|^2$ , where

$$|\phi\rangle = \frac{1}{\sqrt{L}} \sum_{p=0}^{L-1} |p \uparrow, P - p \downarrow\rangle, \quad (6.41)$$

so interaction operator  $\hat{W}$  has rank 1. This arouses the suspicion that the lower bounds on the number of regions obtained in section 6.1 is not sufficiently general and may collapse in the case of a generic interaction. In this section, we generalize the result to interaction operators of rank finite but greater than one. We find that for  $r = \text{rank} \hat{W}$  is held fixed, the number of regions still grows exponentially in  $L$ . On the other hand, our lower bounds do not generalize for infinite-range interactions where we have to resort to heuristic arguments.

We continue to work with two fermions on a finite one-dimensional lattice, so the one-to-one correspondence between wave function coefficients and occupation numbers obeys the relationship  $|\psi_\alpha|^2 = \zeta_\alpha \lambda_\alpha$ . As in previous sections and chapters, we assume that  $\zeta_\alpha = 1$ , all the results obtained herein continue to hold after conjugating  $\hat{W}$  with an invertible transformation. We label the basis indices of the Hilbert space  $\alpha \in \{1, \dots, D\}$ , keeping the momentum-subspace setting of chapter 4 and the basis set  $\Omega_P$  in mind as an example where the one-to-one correspondence between  $\lambda_\alpha$  and  $|\psi_\alpha|^2$  arises, but keeping the discussion without reference to specific physics. This section's central object of study will be the Levy functional (2.29) applied to a real and positive-semidefinite, but otherwise arbitrary interaction operator  $\hat{W}$ . As explained previously, the assumption of real coefficients allows us to restrict the minimization therein to wave functions with real coefficients, reducing (2.29) to a discrete optimization problem:

$$\tilde{\mathcal{F}}(\lambda) = \min_{|\psi_\alpha|^2 = \lambda_\alpha, \psi_\alpha \in \mathbb{R}} \langle \psi | \hat{W} | \psi \rangle = \min_{\sigma \in \{\pm 1\}} \tilde{\mathcal{F}}_\sigma(\lambda); \quad \tilde{\mathcal{F}}_\sigma(\lambda) = \langle \psi^{\sigma; \lambda} | \hat{W} | \psi^{\sigma; \lambda} \rangle; \quad \psi_\alpha^{\sigma; \lambda} = \sigma_\alpha \sqrt{\lambda_\alpha}. \quad (6.42)$$

For the case of  $\text{rank}[\hat{W}] = 1$ , we have already characterized the functional's complexity in sections 6.1 and 6.2. To generalize these results to  $\text{rank}[\hat{W}] > 1$ , we begin with a definition. Let

$$\mathcal{W}^r = \{\hat{W} \in M(n \times n, \mathbb{R}) \mid \hat{W}^\dagger = W, \hat{W} \geq 0, \text{rank}(\hat{W}) = r\} \quad (6.43)$$

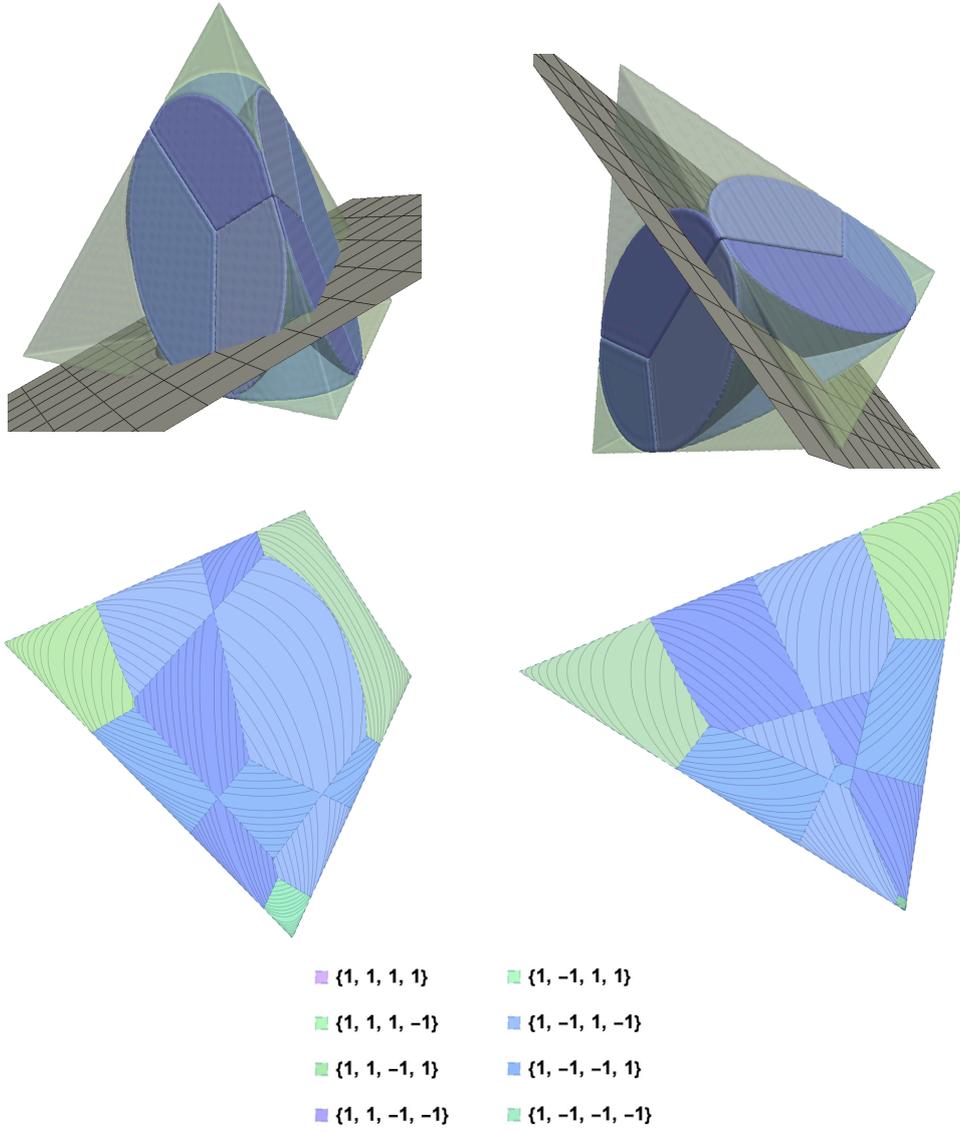


Figure 6.3: The intersection of two-dimensional subspaces with the functional's domain (4.11) for  $D = 4$ . The subspace is the plane defined by  $S = \{\boldsymbol{\lambda}_0 + x\mathbf{a} + y\mathbf{b} \mid x, y \in \mathbb{R}\}$ , and is illustrated in the top figures. The cross-sections  $S \cap \mathcal{D}(\mathcal{F})$  are shown on the bottom. The differently-colored regions correspond to the intersections  $S \cap \mathcal{D}_\sigma$  for different  $\sigma$ . Different shades of the same color (blue or green) correspond to  $\sigma$ -regions related by permutations  $\lambda_i \mapsto \lambda_{P(i)}$ . Note that each blue  $\sigma$ -region, each illustrated a different shade of blue, decomposes into multiple connected components ("cells"), as Lemma 6.3.6 and Corollary 6.3.8 suggest. Green shades indicate the regions corresponding to  $n = 1$ , having only a single connected component each.

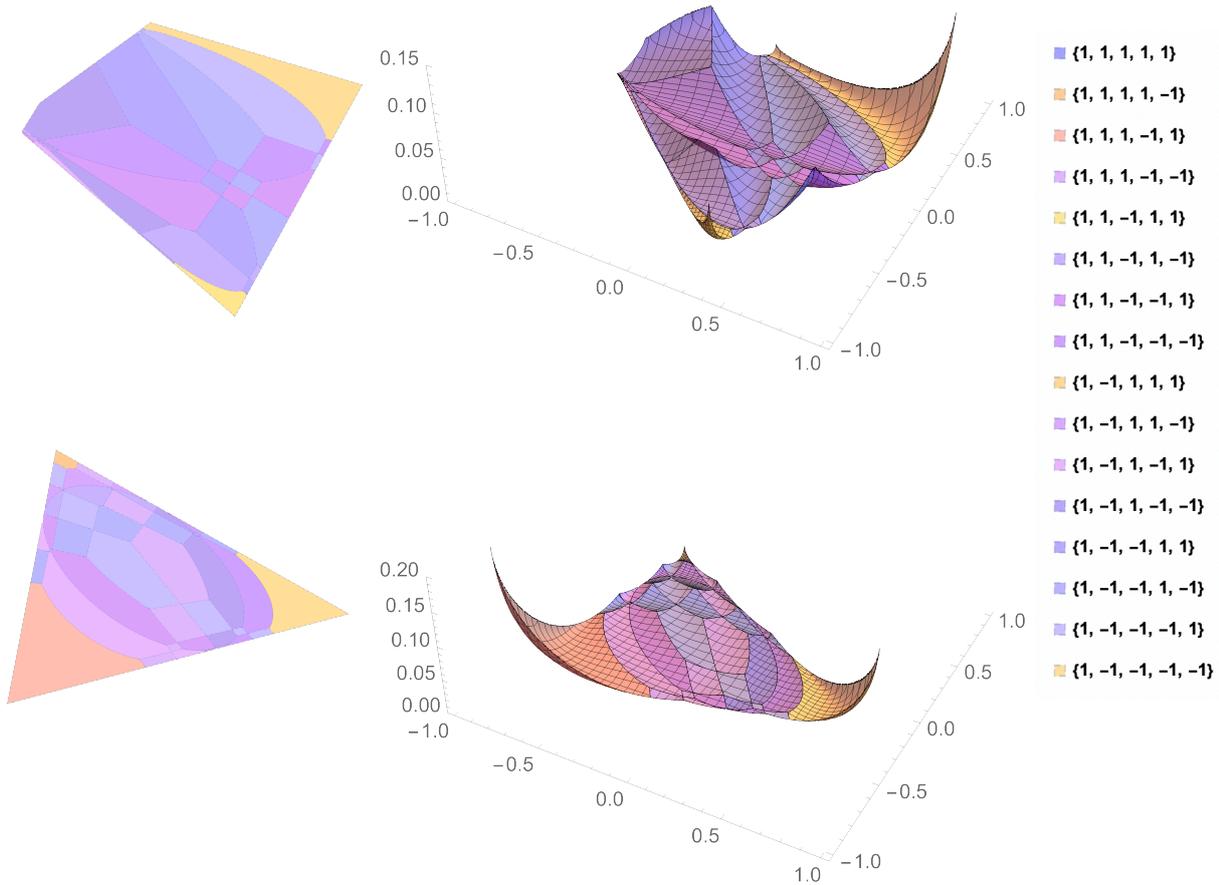


Figure 6.4: Intersections  $S \cap \mathcal{D}(\mathcal{F})$  for two randomly selected two-dimensional subspaces  $S$  and  $D = 5$ .  $\hat{W}$  corresponds to the Hubbard on-site interaction so  $\text{rank}(\hat{W}) = 1$ . The left-hand figures were obtained by applying the method illustrated in figure 6.3 to the four-dimensional hypersimplex. The right-hand figures plot the functional over the domains in the leftmost figure, showing explicitly how the functional splits the domain into subdomains. Note that each purple  $\sigma$ -region, corresponding to some  $\sigma \in \Sigma^n$  with  $n = 2$ , and illustrated one shade of purple, decomposes into multiple connected components ("cells"), as explained by Lemma 6.3.6 and Corollary 6.3.8. The regions corresponding to  $n = 1$  are indicated by orange shades and have only one connected component each.

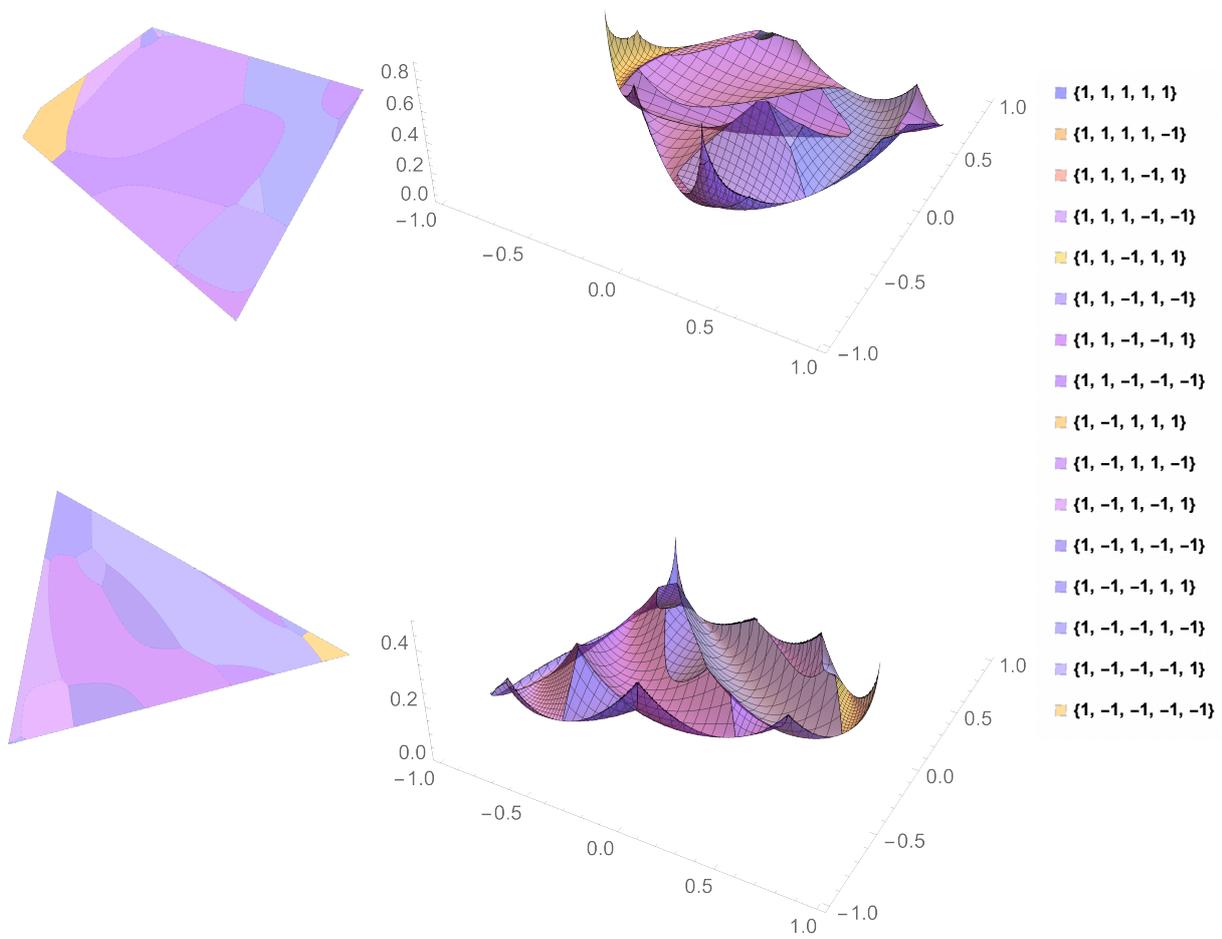


Figure 6.5: Intersections  $S \cap \mathcal{D}(\mathcal{F})$  for two randomly selected two-dimensional subspaces  $S$  and  $D = 5$ .  $\hat{W}$  is a random  $D \times D$  matrix, so generically, it has full rank.

be the set of positive-semidefinite interaction Hamiltonians of rank  $r$ . (Throughout this section, we assume  $r < D$ , so every interaction Hamiltonian has at least one null vector.) Any such hermitian operator can be decomposed into a sum of projectors,

$$\hat{W} = \sum_{i=1}^r w_i |\phi_i\rangle\langle\phi_i|. \quad (6.44)$$

Our goal is to find a (rough) lower bound on the number of  $\sigma$ -regions. Let us define

$$\Sigma_{\hat{W}} = \{\sigma \mid \exists \lambda : \tilde{\mathcal{F}}_{\sigma}(\lambda) < \min_{\sigma' \neq \pm\sigma} \tilde{\mathcal{F}}_{\sigma'}(\lambda)\}. \quad (6.45)$$

This is nothing other than the set of nonempty  $\sigma$ -regions defined in (6.2), so

$$\mathcal{N}_{\hat{W}} = \frac{1}{2} |\Sigma_{\hat{W}}|. \quad (6.46)$$

The lower bound on  $|\Sigma_{\hat{W}}|$ , which we are about to demonstrate, will not apply for all  $\hat{W} \in \mathcal{W}^r$ , but only a certain subset thereof. There are special interaction Hamiltonians where the bound does not apply, but they form a subset of measure zero in the space of positive-semidefinite rank- $r$  interactions.

For brevity, we will say that a statement is true for *almost all*  $\hat{W} \in \mathcal{W}^r$  if it holds on a dense and open subset of  $\mathcal{W}^r$ .

Recall that the states  $|\psi\rangle \in \mathcal{H}_G$  are, in general, members of a symmetry subspace labeled by a tuple of eigenvalues  $G$ ; in our case, these are the eigenvalues of the total-momentum operator, the spin- $z$  operator, and the total spin operator. However, this context did not have any relevance in this chapter thus far, nor will it in this section. For clarity, we will therefore always write  $|\psi\rangle \in \mathbb{R}^D$ , with the understanding that  $|\psi\rangle$  is a quantum state and the restriction from  $\mathcal{H}_G \simeq \mathbb{C}^D$  to  $\mathbb{R}^D$  is permitted due to the assumption that  $\hat{W}$  and  $\hat{h}$  are real matrices in the canonical basis of  $\mathcal{H}_G$ . Denoting the canonical basis  $\{|\alpha\rangle\}_{\alpha=1}^D$ , the states under consideration are the vectors in the real-linear span of  $\{|\alpha\rangle\}_{\alpha=1}^D$  and the latter takes the role of a basis for  $\mathbb{R}^D$ .

We define the  $\sigma$ -hyperoctant

$$\text{HO}_{\sigma} = \{|\psi\rangle \in \mathbb{R}^D \mid \text{sign}(\psi_{\alpha}) = \sigma_{\alpha}\}. \quad (6.47)$$

$\text{HO}_{\sigma}$  is the set of vectors such that each entry has the same sign as the corresponding entry of sigma, e.g.,  $(2, 7, -5)^T \in \text{HO}_{(1,1,-1)}$ . Let us also define

$$\Sigma_{\hat{W}}^0 = \{\sigma \mid \text{HO}_{\sigma} \cap \ker(\hat{W}) \neq \emptyset\}, \quad (6.48)$$

the set of  $\pm$  sequences  $\sigma$  such that there exists a state  $|\psi\rangle \in \text{HO}_{\sigma}$  with  $\langle\psi|\hat{W}|\psi\rangle = 0$ . The latter is equivalent to  $\hat{W}|\psi\rangle = 0$  since  $\hat{W}$  is assumed positive-semidefinite. We will, for now, try to find a lower bound on  $\Sigma_{\hat{W}}^0$  and later extend it to  $\Sigma_{\hat{W}}$ .

### 6.5.1 Hyperplane arrangements

Let  $H_{\alpha} = \{|\psi\rangle \in \mathbb{R}^D : \psi_{\alpha} = 0\}$  be the  $\alpha$ -th coordinate hyperplane. We define the following collection of linear subspaces of  $\ker[\hat{W}]$ :

$$\mathcal{A}_{\hat{W}} = \left\{ H_{\alpha} \cap \ker(\hat{W}) \right\}_{\alpha=1}^D. \quad (6.49)$$

Any such collection of linear subspaces embedded in an ambient space is called a *hyperplane arrangement*. In the literature [51], this term typically refers to collections of affine hyperplanes. We will reserve the term for collections of linear hyperplanes, such as  $\mathcal{A}_{\hat{W}}$ . Note that the set  $\mathcal{A} = \{H_{\alpha}\}_{\alpha=1}^D$  of coordinate hyperplanes in  $\mathbb{R}^D$  is also a hyperplane arrangement with  $\mathbb{R}^D$  the embedding space instead of  $\ker[\hat{W}]$ .

A remark on notation: We will use  $D$  to denote the number of hyperplanes in a hyperplane arrangement and  $n$  the dimension of the ambient space. We will use  $\alpha$  to index the hyperplane and denote subsets of the index set by  $\{\alpha_1, \dots, \alpha_d\} \subset \{1, \dots, D\}$ .

We say that an  $n$ -dimensional hyperplane arrangement  $\mathcal{A} = \{H_{\alpha}\}_{\alpha=1}^D$  is in *general position* if for any permutation  $P$ ,

$$\begin{aligned} d \leq n &\implies \dim(H_{\alpha_1} \cap \dots \cap H_{\alpha_d}) = n - d \\ d > n &\implies \dim(H_{\alpha_1} \cap \dots \cap H_{\alpha_d}) = 0. \end{aligned} \quad (6.50)$$

In other words, the intersection of any sub-arrangement of  $d$  hyperplanes has the lowest possible dimension achievable by intersection  $d$  hyperplanes. This definition has been slightly adapted from [51] since the latter also deals with affine hyperplanes.

Any arrangement  $\mathcal{A} = \{H_\alpha\}_{\alpha=1}^D$  partitions its ambient space  $\simeq \mathbb{R}^n$  into disjoint sets called *sectors*. More precisely, the sectors are the connected components of  $\mathbb{R}^n \setminus \bigcup_\alpha H_\alpha$ . If the number of hyperplanes is finite, so is the number of sectors. Let  $a(\mathcal{A})$  be the number of sectors, and denote the collection of sectors by  $\mathcal{S}(\mathcal{A})$ . By construction,  $S \in \mathcal{S}$  is open,  $S \cap S' = \emptyset$  for  $S \neq S'$ , and  $\bigcup_{S \in \mathcal{S}} S = \mathbb{R}^n \setminus \bigcup_\alpha H_\alpha$ . For example, the sectors generated by the coordinate hyperplanes  $\mathcal{A} = \{H_\alpha\}_{\alpha=1}^D$  are just the hyperoctants of  $\mathbb{R}^D$ :  $\mathcal{S}(\mathcal{A}) = \{\text{HO}_\sigma\}_{\sigma \in \{\pm 1\}^D}$

The number of sectors  $a(\mathcal{A})$  depends on the details of the arrangement, e.g., if all hyperplanes are identical, there are only two sectors. For a sufficiently generic arrangement (i.e., one satisfying (6.50)), however, the situation is better controlled:

**Proposition 6.5.1:** *Let  $\mathcal{A}$  be an arrangement of  $D$  hyperplanes in  $V \simeq \mathbb{R}^n$ , in general position. Then*

$$a(\mathcal{A}) = a_{n-1}^D \equiv \begin{cases} 2^D & D \leq n \\ 2 \sum_{j=0}^{n-1} \binom{D-1}{j} & D > n \end{cases} \quad (6.51)$$

We defer the proof to Appendix A.

### 6.5.2 Lower bound on $|\Sigma_{\hat{W}}^0|$

Let us explain how the preceding section is useful for lower-bounding  $|\Sigma_{\hat{W}}^0|$ . Recalling the definition (6.48), it is clear that we are interested in the number of hyperoctants that intersect  $\ker[\hat{W}]$ . But note that the intersections  $\text{HO}_\sigma \cap \ker(\hat{W})$  are nothing other than the sectors of the hyperplane arrangement defined in (6.49):

$$|\Sigma_{\hat{W}}^0| = a(\mathcal{A}_{\hat{W}}). \quad (6.52)$$

By Proposition 6.5.1, the number of nonempty intersections (i.e., elements of  $\Sigma_{\hat{W}}^0$ ) can be determined from the number of hyperplanes  $D$  and the dimension of the ambient space  $\ker[\hat{W}]$  alone, so long as  $\mathcal{A}_{\hat{W}}$  is in general position. Whether this is the case depends on the specifics of  $\hat{W}$ . As we want to lower-bound  $|\Sigma_{\hat{W}}^0|$  for almost all  $\hat{W}$ , we seek to establish that  $\mathcal{A}_{\hat{W}}$  will be in general position for almost all  $\hat{W}$ .

Let  $\mathcal{A} = \{H_\alpha\}_{\alpha=1}^D$  be the hyperplane arrangement containing the coordinate hyperplanes of  $\mathbb{R}^D$ . Then clearly,  $\mathcal{A}$  is in general position. To show that  $\mathcal{A}_{\hat{W}} = \mathcal{A} \cap \ker(\hat{W})$  is also in general position, we must show that  $\dim(\ker(\hat{W}) \cap H_{\alpha_1} \cap \dots \cap H_{\alpha_d}) = \dim(\ker[\hat{W}]) - d$  for all  $I = \{\alpha_1, \dots, \alpha_d\} \subset \{1, \dots, D\}$ . Define

$$\begin{aligned} \mathcal{W}_I^r &= \{\hat{W} \in \mathcal{W}^r \mid \dim(\ker[\hat{W}] \cap H_{\alpha_1} \cap \dots \cap H_{\alpha_d}) = \max(\dim(\ker[\hat{W}]) - d, 0)\} \\ &= \{\hat{W} \in \mathcal{W}^r \mid \mathcal{A}_{\hat{W}} \text{ is in general position}\} \end{aligned} \quad (6.53)$$

Then  $\mathcal{A} \cap \ker(\hat{W})$  is in general position if and only if  $\hat{W} \in \bigcap_{I \subset \{1, \dots, D\}} \mathcal{W}_I^r$ . Clearly, if  $\mathcal{W}_I^r$  is dense in  $\mathcal{W}^r$  for all  $I$ , then  $\bigcap_{I \subset \{1, \dots, D\}} \mathcal{W}_I^r$  is, too. Since one expects a generic hyperplane arrangement to be in general position, one also expects that for a randomly picked  $\hat{W} \in \mathcal{W}^r$ , the arrangement  $\mathcal{A}_{\hat{W}}$  should be in general position. We can show that this is indeed the case:

**Lemma 6.5.2:**  $\mathcal{W}_I^r$  is dense and open in  $\mathcal{W}^r$ .

**Corollary 6.5.3:**  $\bigcap_{I \subset \{1, \dots, D\}} \mathcal{W}_I^r$  is dense and open in  $\mathcal{W}^r$ .  $\square$

The Proof of Lemma 6.5.2 is somewhat technical and shown in App. B. We are now able to formulate a lower bound on  $|\Sigma_{\hat{W}}^0|$  for a dense subset of  $\mathcal{W}^r$ :

**Proposition 6.5.4:**  $|\Sigma_{\hat{W}}^0| = a_{D-r-1}^D$  for almost all  $\hat{W} \in \mathcal{W}^r$ .

*Proof:* From the definitions (6.50), (6.53) and (6.49) follows that  $\mathcal{A}_{\hat{W}}$  is in general position if  $\hat{W} \in \bigcap_{I \subset \{1, \dots, D\}} \mathcal{W}_I^r$ . By Corollary 6.5.3, this is the case for almost all  $\hat{W} \in \mathcal{W}^r$ . Proposition 6.5.1 then implies that the  $a(\mathcal{A}_{\hat{W}}) = a_{D-r-1}^D$  for almost all  $\hat{W}$ , so the proposition follows from (6.52).  $\square$

### 6.5.3 Lower bound on $|\Sigma_{\hat{W}}|$

Having characterized  $\Sigma_{\hat{W}}^0$ , we turn to  $\Sigma_{\hat{W}}$ . Note that  $\Sigma_{\hat{W}}^0$  is the set of sign combinations  $\sigma$  such that there exists a  $\lambda \in \Delta^{D-1}$  satisfying  $\mathcal{F}_\sigma(\lambda) = 0$ . This follows from the definition (6.48) because, picking any  $|\psi\rangle \in \text{HO}_\sigma \cap \ker[\hat{W}]$ , there exist  $\lambda_\alpha \in [0, 1]$  such that  $|\psi\rangle = \sum_\alpha \sigma_\alpha \sqrt{\lambda_\alpha}$ .

The assertion that  $\sigma \in \Sigma_{\hat{W}}^0$  does not imply  $\sigma \in \Sigma_{\hat{W}}$ , unless also  $\mathcal{F}_{\sigma'}(\lambda) > 0$  holds for all  $\sigma' \neq \sigma$ . To fulfill this last requirement, it is useful to define the orthogonal transformation  $U_\sigma$ ,

$$\langle \alpha | U_\sigma | \psi \rangle = \sigma_\alpha \langle \alpha | \psi \rangle. \quad (6.54)$$

The group of such operators shall be written

$$\mathcal{U}_\Sigma = \{U_\sigma \mid \sigma \in \{-1, 1\}^D\}. \quad (6.55)$$

Let  $U \in U(D)$  be an orthogonal matrix acting on  $\mathbb{R}^D$ . For  $V \subset \mathbb{R}^D$  a subspace, we denote  $U(V)$  the image of  $V$  under  $U$ . Let us define yet another set of positive-semidefinite hermitian operators:

$$\mathcal{W}_U^r = \{\hat{W} \in \mathcal{W}^r \mid U(\ker[\hat{W}]) \neq \ker[\hat{W}]\}. \quad (6.56)$$

The condition for belonging to this set is that the kernel of  $\hat{W}$  not be invariant under the action of  $U$ . To this end, we use a slightly more general statement:

**Lemma 6.5.5:** *Let  $U : \mathbb{C}^D \rightarrow \mathbb{C}^D$  be a unitary matrix with  $U \neq \pm I$ . Then  $\mathcal{W}_U^r \subset \mathcal{W}^r$  is dense and open.*

We prove the lemma in App. C.

**Corollary 6.5.6:** *For any  $U \in \mathcal{U}_\Sigma \setminus \{I, -I\}$ , the subset  $\mathcal{W}_U^r$  is dense and open in  $\mathcal{W}^r$ . Since  $|\mathcal{U}_\Sigma|$  is finite,  $\bigcap_{U \in \mathcal{U}_\Sigma \setminus \{I, -I\}} \mathcal{W}_U^r$  is also dense and open in  $\mathcal{W}^r$ .  $\square$*

**Lemma 6.5.7:** *Assume  $\hat{W} \in \bigcap_{U \in \mathcal{U}_\Sigma \setminus \{I, -I\}} \mathcal{W}_U^r$ . Then, for all  $\sigma \in \Sigma_{\hat{W}}^0$ , there exists  $|\psi\rangle \in \text{HO}_\sigma$  such that  $|\psi\rangle \in \ker[\hat{W}]$  and  $U|\psi\rangle \notin \ker[\hat{W}]$  for all  $U \in \mathcal{U}_\Sigma \setminus \{I, -I\}$ . Furthermore,  $\sigma \in \Sigma_{\hat{W}}$ .*

*Proof:* Note that  $\hat{W} \in \mathcal{W}_U^r$  implies that  $\dim[U(\ker[\hat{W}]) \cap \ker[\hat{W}]] < \dim[\ker[\hat{W}]]$ , so its complement,  $\ker[\hat{W}] \setminus U(\ker[\hat{W}])$ , is dense in  $\ker[\hat{W}]$ . By the assumption of the lemma, this is the case for all  $U \in \mathcal{U}_\Sigma \setminus \{I, -I\}$ . Since  $\mathcal{U}_\Sigma$  is a finite set, the intersection of complements,  $\ker[\hat{W}] \setminus \left(\bigcup_{U \in \mathcal{U}_\Sigma \setminus \{I, -I\}} U(\ker[\hat{W}])\right)$  must also be dense.

$\text{HO}_\sigma \cap \ker[\hat{W}]$  is open in  $\ker[\hat{W}]$ , and the intersection of a dense subset with an open subset is nonempty, so there exists

$$|\psi\rangle \in \text{HO}_\sigma \cap \left( \ker[\hat{W}] \setminus \left( \bigcup_{U \in \mathcal{U}_\Sigma \setminus \{I, -I\}} U(\ker[\hat{W}]) \right) \right). \quad (6.57)$$

It is clear that  $|\psi\rangle$  fulfills the properties claimed in the lemma; In particular,  $U|\psi\rangle \notin \ker[\hat{W}]$  for all  $U \in \mathcal{U}_\Sigma \setminus \{I, -I\}$ : If we had  $U|\psi\rangle \in \ker[\hat{W}]$ , this would imply  $|\psi\rangle \in U^{-1}(\ker[\hat{W}])$ , which contradicts (6.57) because  $U^{-1}$  also an element of  $\mathcal{U}_\Sigma \setminus \{I, -I\}$ .

To prove that  $\sigma \in \Sigma_{\hat{W}}$ , we must show that there exists  $\lambda \in \Delta^{D-1}$  such that  $\tilde{\mathcal{F}}_{\sigma'}(\lambda) > \tilde{\mathcal{F}}_\sigma(\lambda)$  for all  $\sigma' \neq \sigma$ . We pick  $\lambda_\alpha = |\psi_\alpha|^2$  with  $|\psi\rangle$  as in the first part of the lemma. Then, because  $|\psi\rangle \in \text{HO}_\sigma$ , we have  $|\psi\rangle = |\psi^{\sigma;\lambda}\rangle$  as defined in (6.42). But  $\hat{W}|\psi^{\sigma;\lambda}\rangle = 0$  and for every  $\sigma' \neq \pm\sigma$  we have  $|\psi^{\sigma';\lambda}\rangle = U|\psi^{\sigma;\lambda}\rangle$  for some  $U \in \mathcal{U}_\Sigma \setminus \{I, -I\}$ , implying  $\hat{W}|\psi^{\sigma';\lambda}\rangle \neq 0$ , again by the first part of the lemma. Because  $\hat{W}$  is positive-semidefinite, this implies  $\tilde{\mathcal{F}}_{\sigma'}(\lambda) = \langle \psi^{\sigma';\lambda} | \hat{W} | \psi^{\sigma';\lambda} \rangle > 0$ . This holds for all  $\sigma' \neq \pm\sigma$ , and  $\tilde{\mathcal{F}}_\sigma(\lambda) = \langle \psi^{\sigma;\lambda} | \hat{W} | \psi^{\sigma;\lambda} \rangle = 0$ , so  $\sigma \in \Sigma_{\hat{W}}$ .  $\square$

Finally, we can show that the cardinality of  $\Sigma_{\hat{W}}$  grows exponentially with  $D$ :

**Proposition 6.5.8:**  $|\Sigma_{\hat{W}}| = a_{D-r-1}^D$  for almost all  $\hat{W} \in \mathcal{W}^r$ .

*Proof:* The subset

$$\mathcal{V} = \left( \bigcap_{U \in \mathcal{U}_\Sigma \setminus \{I, -I\}} \mathcal{W}_U^r \right) \cap \left( \bigcap_{I \subset \{1, \dots, D\}} \mathcal{W}_I^r \right) \quad (6.58)$$

is dense and open in  $\mathcal{W}^r$  by corollaries 6.5.3 and 6.5.6, and if  $\hat{W} \in \mathcal{V}$ , then the definitions (6.53) and (6.49) together with Proposition 6.5.1 imply that  $|\Sigma_{\hat{W}}^0| = a_{D-r-1}^D$ . Lemma 6.5.7 then implies that all  $\sigma \in \Sigma_{\hat{W}}^0$  are also in  $\Sigma_{\hat{W}}$ .  $\square$

### 6.5.4 Numerical Analysis

Proposition 6.5.8 shows that, as long as the interaction operator has fixed rank, the number of  $\sigma$ -regions will grow exponentially with the Hilbert-space dimension  $D$ . Recalling the relation between dimension and chain length  $L$  as stated in table 4.1, the functional consequently becomes exponentially more involved as the chain length is increased.

Physically, the finite-rank requirement translates into a finite interaction range. The Hubbard interaction, being strictly local, has rank 1 on the two-particle, total-momentum- $P$ -subspace. On the other hand, choosing an interaction that ranges over  $r$  lattice sites will yield an operator of rank  $r + 1$  in this subspace. Many interactions of physical interest, however, have infinite range, which implies that  $\hat{W}$  has full rank ( $r = D$ ) and voids the results of subsections 6.5.2 and 6.5.3. To have some idea of how the  $\sigma$ -regions behave for such interactions, we once again resort to Monte Carlo sampling, as already done for the Hubbard interaction in section 6.1.2. As a first primitive attempt, we return the sampling experiment explained in section 6.1.2, the results of which were shown for the Hubbard model in Fig. 6.1.

We investigate two cases. Firstly, we consider a rank-1 interaction; the functional will have the form (4.59). For the Hubbard model,  $\zeta_\alpha$  could take the value 1 or 2; generalizing to an arbitrary rank-1 interaction amounts to allowing any  $\zeta > 0$ . Note that maintaining the form (4.59) restricts the interaction Hamiltonian to positive matrix entries. However, it is easy to see in (4.59) that changing  $\sqrt{\zeta_\alpha} \rightarrow -\sqrt{\zeta_\alpha}$  leaves the functional unchanged because this sign flip is equivalent to a permutation of the  $\sigma$ 's. Such a permutation does not change the number of nonempty  $\mathcal{D}_\sigma$ 's, merely relabeling them. The result of the Monte Carlo experiment is shown in Fig. 6.6. The plots show that some regions appear to be missing, but in light of the results seen in Fig. 6.1, we can anticipate that regions with a greater imbalance of plus and minus signs are expected to be very small and it is conceivable that the regions are nonempty but weren't sampled. As explained at the end of section 6.1, changing  $\zeta$  is only a coordinate transformation in  $\lambda$ -space, so it is expected that all but one of the  $\sigma$ -regions continue to manifest.

Secondly, we consider a full-rank operator  $\hat{W}$  with polynomially decaying spectrum  $w_j = 1/j$ . This case is particularly interesting because the eigenvalues of the Coulomb interaction operator also behave like this. The Monte Carlo results are shown in Fig. 6.7. It is apparent that  $\mathcal{D}_\sigma$  failed to manifest for a greater number of the  $\sigma \in \{\pm 1\}^D$  compared to the rank-1 case, and that the cell volumes are less evenly distributed. This is consistent with the analytical observation that increasing the rank of  $\hat{W}$  leads to a decrease in the number of nonempty  $\sigma$ -regions.

Figs. 6.6 and 6.7 represent randomly picked interaction Hamiltonians, i.e., the spectra are fixed while the eigenvectors are chosen arbitrarily. The routine that samples the eigenvectors is shown in algorithm 4. The figures shown are illustrative but do not provide much insight into how  $\mathcal{N}_{\text{Coulomb}}$  scales with the system size  $L$ . They are also not guaranteed to be representative due to the arbitrary choice of diagonal basis for  $\hat{W}$ .

To obtain a more meaningful result, we instead sample  $n$  different interaction operators  $\{\hat{W}_j\}_{j=1}^n$  — sampling diagonal bases while fixing the spectrum — and run the Monte Carlo routine for each  $\hat{W}$ . We also extend the simplex dimension up to  $D = 17$ . There is no hope of producing plots like Figs. 6.6 and 6.7, as the number of inequivalent  $\sigma$ 's is  $2^{D-1} = 2^{16}$ . Rather than plotting the volume of each  $\mathcal{D}_\sigma$ , we will only try to determine whether a given region exists, i.e., we only distinguish between the number of  $\lambda$  found in  $\mathcal{D}_\sigma$  being zero and non-zero.

Sampling the eigenbasis of  $\hat{W}$  is shown in algorithm 4. For each of the  $n$  operators  $\hat{W}$ , we define the functionals  $\widetilde{\mathcal{F}}_{G\sigma}$  as in (6.42) and sample  $N$  random occupation number vectors  $\lambda \in \Delta^{D-1}$  as before. If  $\widetilde{\mathcal{F}}_{G\sigma}(\lambda) < \widetilde{\mathcal{F}}_{G\sigma'}(\lambda)$  for all  $\sigma' \neq \sigma$ , we can conclude that the region  $\mathcal{D}_\sigma$  is nonempty. After performing this test for each  $\sigma$  and each  $\lambda$ , we obtain a lower bound on  $x_{\hat{W}}$ , the number of regions  $\mathcal{D}_\sigma$  that are nonempty. Finally, we average  $x_{\hat{W}_j}$  over  $j = 1, \dots, n$  to estimate

$$\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}^{\text{Coulomb}} = \int \cdots \int \mathcal{N}_{\hat{W}} d\mu_{\mathbb{S}^0}(|\phi_D\rangle) \cdots d\mu_{\mathbb{S}^{D-1}}(|\phi_1\rangle). \quad (6.59)$$

$\mu_{\mathbb{S}^{D-j}}(|\phi_j\rangle)$  is the uniform measure on the  $(D-1)$ -dimensional sphere, and samples the  $j$ -th eigenvector of  $\hat{W}$  after  $j-1$  eigenvectors have been fixed by  $j-1$  outermost integrals. The innermost integral,  $d\mu_{\mathbb{S}^0}$ , is a discrete sum with two summands.  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}^{\text{Coulomb}}$  is the average number of nonempty  $\sigma$ -regions for interactions with spectrum  $(1, 1/2, 1/3, \dots)$ .

We repeat the procedure for different values of  $D$  ranging from 3 to 17. The results are presented in Fig. 6.8. In subplot (a), we see that the estimate  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}^{\text{Coulomb}}$  grows exponentially for the first few values of the Hilbert space dimension  $D$ . For  $D \gtrsim 7$ , however, the convergence of  $\mathcal{N}_{\hat{W}}$  is very poor in the number of samples  $N_\lambda$ , owing to the dimensionality of  $\Delta^{D-1} \ni \lambda$ . This becomes apparent in subplot (d), where  $\mathcal{N}_{\hat{W}}$  is plotted over

$N_\lambda$ . Our results are far from converged in this regime, and the challenge of determining  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}^{\text{Coulomb}}$  remains unmet.

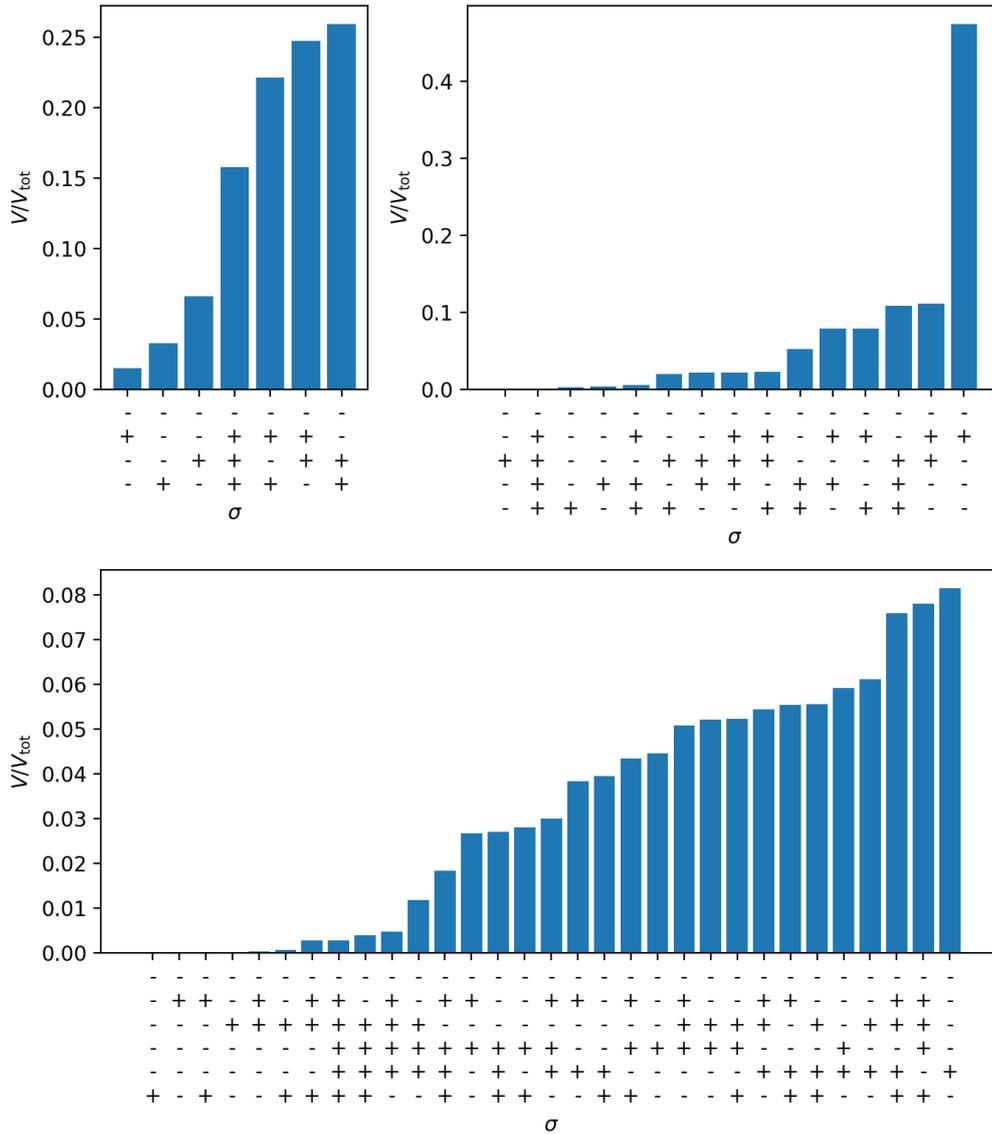


Figure 6.6: Relative volume  $V$  of  $\mathcal{D}_\sigma$  plotted over  $\sigma \in \{-1, 1\}^D$ , with  $\hat{W} = |\phi\rangle\langle\phi|$  a projector where  $|\phi\rangle$  is picket at random.  $N_\Delta = 10^6$  points were sampled from the domain for each value of  $D$ .

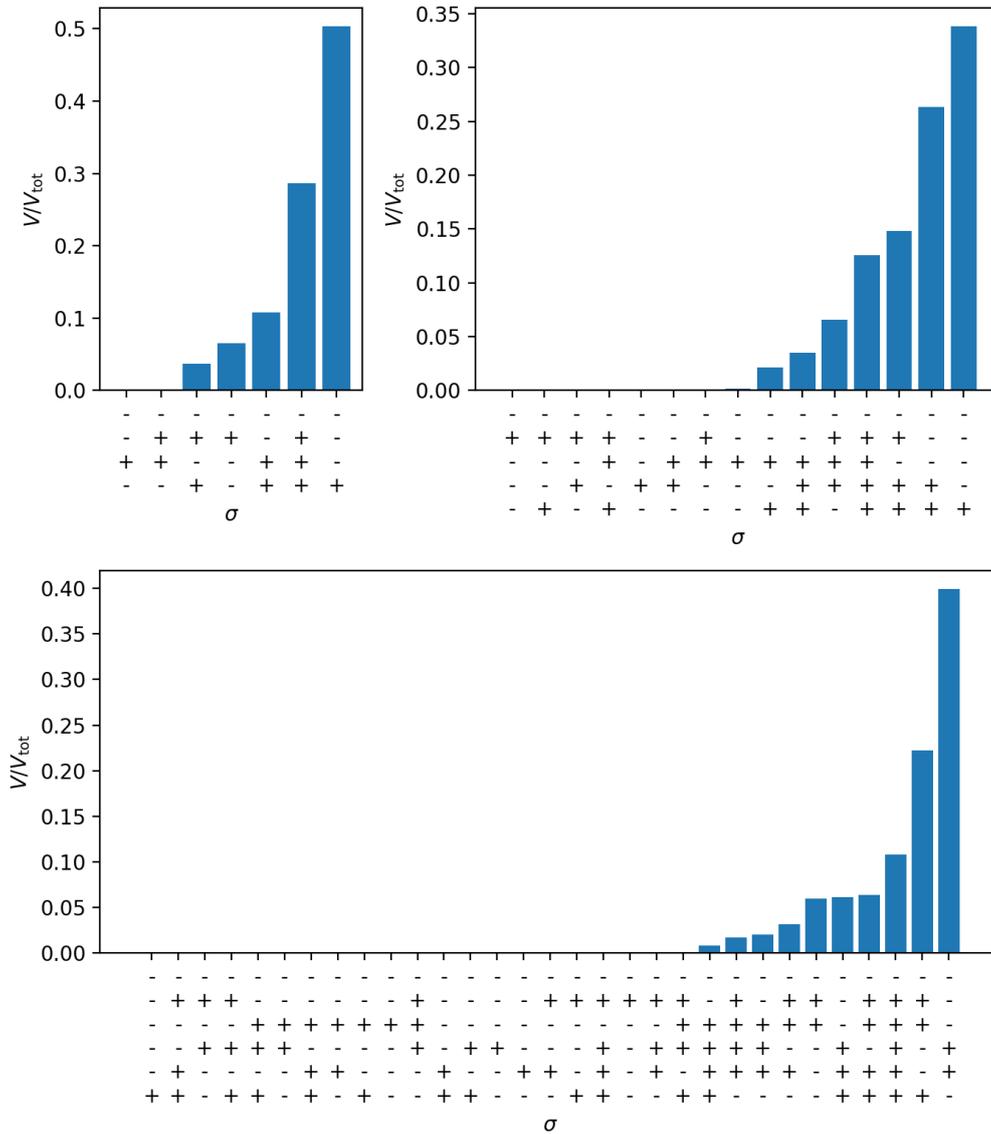
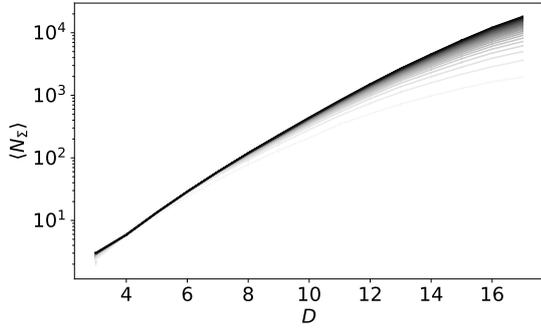
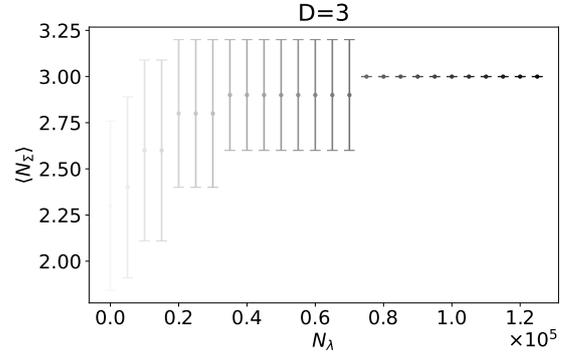


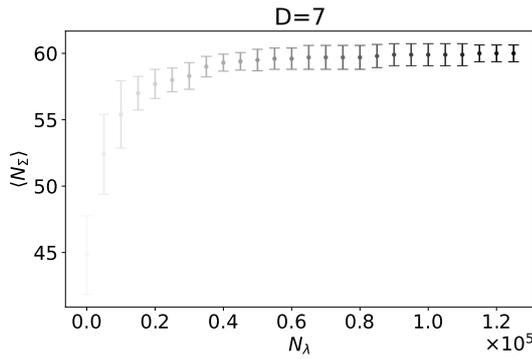
Figure 6.7: Relative volume  $V$  of  $\mathcal{D}_\sigma$  plotted over  $\sigma \in \{-1, 1\}^D$ , with  $\hat{W}$  a randomly sampled interaction operator with spectrum  $(1, 1/2, 1/3, \dots)$ .



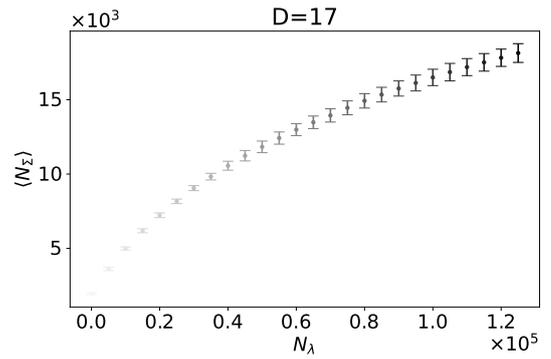
(a) The variable on the  $x$ -axis is the dimension of the Hilbert space. Each curve corresponds to a different number  $N_\lambda$  of points that were sampled from the simplex. The darkest-shaded curve is for  $N_\lambda = 1.3 \times 10^5$ ; the lightest-shaded curve is for  $N_\lambda = 5 \times 10^3$



(b) The number of  $\sigma$ -regions found by the Monte-Carlo routine, plotted over the total number of samples taken from the simplex. The error bars indicate the standard deviation of  $N_\Sigma$  over the  $n = 10$  different choices of eigenbasis for  $\hat{W}$ .



(c) Same as (b), but for Hilbert space dimension  $D = 7$ . The simplex is still sufficiently low-dimensional to effectively sample the  $\sigma$ -regions, and the lower bounds converge.



(d) Same as (b), but for Hilbert space dimension  $D = 17$ . The number of samples is too small for the lower bounds  $N_\Sigma$  to converge adequately, and only a rough lower bound can be obtained.

Figure 6.8: Monte Carlo estimates for the number of  $\sigma$ -regions. For every simplex dimension  $D$ , the results were obtained by sampling  $n$  interaction Hamiltonians  $\{\hat{W}_i\}_{i=1}^n$  using algorithm 4. All data shown are for  $n = 10$ . For each  $\hat{W}_i$ , the simplex  $\Delta^{D-1}$  is sampled  $N_\lambda$  times to find as many  $\sigma$ -regions as possible and obtain as tight a lower-bound  $\mathcal{N}^{\hat{W}_i}$  as possible. The individual lower bounds are averaged over  $i = 1, \dots, n$  to obtain an approximate (due to statistical fluctuations over  $i$ ) lower bound for  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}$ . This average of lower bounds is denoted  $\langle N_\Sigma \rangle$ . In the limit  $n \rightarrow \infty$ ,  $\langle N_\Sigma \rangle$  converges to an exact lower bound on  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}$ . In the limit  $N_\lambda \rightarrow \infty$ , the individual lower bounds, and therefore, the  $\langle N_\Sigma \rangle$  become tighter and eventually converge to  $\mathcal{N}_{\hat{W}}$ , as more and more points are sampled from the simplex, and more of the  $\sigma$  regions are found.  $\langle N_\Sigma \rangle$  then converges to an approximation of  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}$ , or the exact value in the  $n \rightarrow \infty$  limit. This is the convergence observed in panels (b) and (c); also note that the  $D = 3$  case indicates that the simplex will always decompose into three cells as in Fig. 4.6a However, as seen in panel (d), this convergence is out of reach for higher-dimensional cases, and only a very rough lower bound can be obtained.

---

**Algorithm 3** Numerical estimation of  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}}$ , the expected number of nonempty regions  $\mathcal{D}_{\sigma}$  when sampling random Hamiltonians with a fixed spectrum.  $c$  is a dictionary where  $c[\sigma]$  is the number of  $\lambda \in \mathcal{D}(\sigma)$  that were found.  $x[j]$  is the number of  $\sigma$ -regions where at least one such  $\lambda$  was found. Only  $\sigma$  satisfying  $\sigma_1 = -1$  are considered, as combinations differing by a global sign flip are equivalent

---

```

1: Let  $\mathbf{x}$  be an array of  $n$  real numbers
2: for  $j = 1 \dots n$  do
3:   Generate random  $\hat{W}$  with spectrum  $(1, 1/2, 1/3, \dots)$ 
4:   for  $\sigma \in \{-1, 1\}^D$  with  $\sigma_0 = -1$  do
5:      $c[\sigma] \leftarrow 0$ 
6:   end for
7:   for  $l = 1 \dots N$  do
8:     Let  $\lambda$  be a uniformly sampled random point in  $\Delta^{D-1}$ 
9:     if  $\tilde{\mathcal{F}}_{G_{\sigma}}(\lambda) < \tilde{\mathcal{F}}_{G_{\sigma'}}(\lambda)$  for all  $\sigma' \neq \sigma$  then
10:       $c[\sigma] \leftarrow c[\sigma] + 1$ 
11:    end if
12:  end for
13:   $x[j] \leftarrow 0$ 
14:  for  $\sigma \in \{-1, 1\}^D$  with  $\sigma_0 = 1$  do
15:    if  $c[\sigma] > 0$  then
16:       $x[j] \leftarrow x[j] + 1$ 
17:    end if
18:  end for
19: end for
20:  $\langle \mathcal{N}_{\hat{W}} \rangle_{\hat{W}} \leftarrow \text{mean}[\mathbf{x}]$ 

```

---



---

**Algorithm 4** Generation of a random interaction Hamiltonian  $\hat{W}$  with spectrum  $\mathbf{w}$ .  $\Phi$  is a  $D \times D$  matrix that contains the orthonormal eigenvectors of  $\hat{W}$ . Each iteration of the  $j$ -loop generates a new eigenvector. During each iteration, the vector  $\phi$  is first sampled from the hypersphere  $\mathbb{S}^{D-1}$ . The  $l$ -loop is a Gram-Schmidt process ensuring that the new eigenvector  $\Phi[j, :]$  is orthogonal to all previous ones.

---

```

1: Let  $W$  be a  $D \times D$  array
2: Let  $\Phi$  be a  $D \times D$  array
3: for  $j = 1, \dots, r$  do
4:   Let  $\lambda$  be a uniformly sampled random point in  $\Delta^{D-1}$ 
5:   Let  $\sigma$  be a uniformly sampled random element of  $\{-1, 1\}^D$ 
6:   Let  $\phi$  be an array with entries  $\phi_l = \sqrt{\lambda_l} \sigma_l$ 
7:   for  $i = 1, \dots, j-1$  do
8:      $\phi \leftarrow \phi - \Phi[i, :] \sum_{l=1}^D \Phi[i, l] \phi[l]$ 
9:   end for
10:   $\phi \leftarrow \phi / |\phi|$ 
11:   $\Phi[j, :] \leftarrow \phi$ 
12:   $W \leftarrow W + w_j \phi \phi^T$ 
13: end for

```

---

# Chapter 7

## Summary and Conclusion

In this thesis, we have lower-bounded the complexity of the Levy functional of a two-particle system in various settings. Strikingly, we have established exponential complexity in several instances despite the Hilbert space dimension being linear in the system size.

We have introduced the general formalism of functional theory in chapter 2. Therein, we first explained the foundation of density functional theory and the Hohenberg-Kohn functional [5]. Then, we defined the concepts of one-particle reduced density matrices and one-body operators and how they behave under symmetry transformation. This laid the basis for presenting reduced density matrix functional theory, or RDMFT, as a generalization of DFT. We defined the Gilbert, Levy, and Valone functionals [39, 40, 43] and explained the problem of  $\hat{h}$ -representability and  $N$ -representability.

Chapter 3 was dedicated to establishing a basic background in the theory of computational complexity, without assuming any prior knowledge of computer science. The complexity classes P and NP were explained, as well as the concepts of NP-hardness and NP-completeness. These abstract notions were illustrated with examples such as the subset-sum problem and the 3-satisfiability problem. We explained how an optimization problem can be mapped to a decision problem, and how the complexity classes of the two relate. We kept the discussion general in order to provide an intuitive introduction, but mainly limited our aim to establishing the concepts needed to put the later results in chapter 6 into proper context.

In chapter 4, we applied the theory of chapter 2 to a concrete model, namely two fermions subjected to a repulsive on-site interaction on a one-dimensional finite periodic lattice. We found firstly that the Levy functional  $\mathcal{F}$  can, in principle, be computed analytically by performing a minimization over  $U(1)^D$  where  $D$  is a positive integer and  $U(1)$  is the set of complex numbers of unit magnitude. We introduced a second candidate, the Levy functional  $\tilde{\mathcal{F}}$ , which is known to be equivalent to  $\mathcal{F}$  but only requires minimization over the smaller set  $\{\pm 1\}^D$ . Owing to the discreteness of  $\{\pm 1\}^D$ , we found that this second functional is markedly more difficult to determine, with no immediate analytical solution.

In chapter 5, we studied the analytical structure of  $\mathcal{F}$  for an arbitrary interaction Hamiltonian on a three-dimensional Hilbert space. In particular, we identified all possible analytical forms the functional may take and classified them according to the Hamiltonian's degeneracy structure.

Chapter 6 was entirely dedicated to studying  $\tilde{\mathcal{F}}$  for the Hubbard on-site interaction. Despite the lack of a closed-form expression, we were able to determine that  $\tilde{\mathcal{F}}$  is non-analytic on the interior of its domain and that the manifolds of non-analyticity partition the domain into  $\sim 2^L$  subdomains, with  $L$  the system size. We found that computing  $\tilde{\mathcal{F}}$  is equivalent to solving the partition problem, thereby demonstrating NP-completeness. We then made progress towards generalizing this result to finite-range interaction Hamiltonians. Using tools from the theory of hyperplane arrangements [51], we established that the number of subdomains still behaves as  $\sim 2^L$  as long as no infinite-range interactions are introduced. Finally, we attempted to extrapolate the results to infinite-range interactions using Monte Carlo techniques. Still, we did not obtain any definitive bounds for Hilbert space dimension  $\geq 8$  due to the difficulties inherent in higher-dimensional sampling.

Several exciting potential avenues of future research emerge from this study. Firstly, the analytical complexity for a generic interaction between two particles on a one-dimensional lattice could be analyzed for the Levy functional without assuming time-reversal invariance, having found that using the latter only complicates evaluation. Secondly, attempts could be made to extrapolate the analytical results from finite-range interactions to infinite-range interactions with sufficiently rapid decay. Finally, more sophisticated numerical studies could be undertaken to gain insights along those lines.



# Appendix A

## Number of Sectors of Hyperplane Arrangements

Here, we prove Proposition 6.5.1 on the number of sectors generated by a hyperplane arrangement that is in general position. We begin with a Lemma that follows directly from the definition (6.50):

**Lemma A.1:** *Assume without loss of generality that  $\{\alpha_1, \dots, \alpha_d\} \subset \{1, \dots, m\}$ , and define the sub-arrangement  $\mathcal{A}_d = \{H_1, \dots, H_d\}$ . Assume that  $\mathcal{A}$  is in general position. Then  $\mathcal{A}_d$  is also in general position. Moreover, denote*

$$\mathcal{A}_d \cap H_{d+1} \equiv \{H_1 \cap H_{d+1}, \dots, H_d \cap H_{d+1}\}.$$

*Then  $\mathcal{A}_d \cap H_{d+1}$  is an  $(n-1)$ -dimensional hyperplane arrangement in  $H_{d+1}$ , also in general position.  $\square$*

Next, we prove the intuitive fact that the intersections of the sectors with a hyperplane equal the sectors of the intersected hyperplane arrangement.

**Lemma A.2:** *Consider the arrangements  $\mathcal{A}_d$  and  $\mathcal{A}_d \cap H_{d+1}$  defined in Lemma A.1. Then we have*

$$\mathcal{S}(\mathcal{A}_d \cap H_{d+1}) = \{S \cap H_{d+1} \mid S \in \mathcal{S}(\mathcal{A}_d)\}.$$

*Proof:* Let  $\mathcal{S}^\cap = \{S \cap H_{d+1} \mid S \in \mathcal{S}(\mathcal{A}_d)\}$ . To show that  $\mathcal{S}^\cap = \mathcal{S}(\mathcal{A}_d \cap H_{d+1})$ , it suffices to show that its sets are disjoint and open and that their union equals  $H_{d+1} \setminus (\cup_{\alpha=1}^d (H_\alpha \cap H_{d+1}))$ . The first two are obvious, since  $(S' \cap H_{d+1}) \cap (S \cap H_{d+1}) = \emptyset$  if  $S \cap S' = \emptyset$ , and since  $S \cap H_{d+1}$  is open in  $H_{d+1}$  if  $S$  is open in  $\mathbb{R}^n$ . The third property amounts to unwrapping the definitions:

$$\begin{aligned} H_{d+1} \setminus \cup_{\alpha=1}^d (H_\alpha \cap H_{d+1}) &= H_{d+1} \cap (\mathbb{R}^n \setminus \cup_{\alpha=1}^d H_\alpha) \\ &= H_{d+1} \cap (\cup_{S \in \mathcal{S}(\mathcal{A}_d)} S) \\ &= \cup_{S \in \mathcal{S}(\mathcal{A}_d)} (H_{d+1} \cap S), \end{aligned} \tag{A.1}$$

which is the union of sets in  $\mathcal{S}^\cap$ , concluding the Lemma.  $\square$

Now that we have established the necessary technical details, we can prove Proposition 6.5.1 on the number of sectors to a generally positioned arrangement. We will use induction over  $n$ , the dimension of the ambient space. In particular, we will derive the increase in the number of sectors when another Hyperplane  $H_{m+1}$  is added to an arrangement of  $m$  hyperplanes. This will allow us to write down a recursion relation for  $a(\mathcal{A}_m)$  that can be solved using the combinatorial expression (6.51).

*Proof of Proposition 6.5.1:* We proceed by induction over  $n$ . For  $n = 1$ , the statement is true because  $\{0\}$  is the only linear hyperplane, and it divides the real line into two sectors. For the inductive step, assume the statement is true for dimension  $n-1$ , and consider  $\mathcal{A}_d$  and  $\mathcal{A}_d \cap H_{d+1}$  as defined in Lemmas 6.5.1 and 6.5.2.  $\mathcal{A}_d$  divides  $\mathbb{R}^n$  into  $a(\mathcal{A}_d)$  sectors. Let  $S \in \mathcal{S}(\mathcal{A}_d)$  be such a sector. If  $H_{d+1} \cap S = \emptyset$ , then  $S$  is still a connected component of  $\mathbb{R}^n \setminus (\cup_{\alpha=1}^{d+1} H_\alpha)$  and thus  $S \in \mathcal{S}(\mathcal{A}_{d+1})$ . On the other hand, suppose  $H_{d+1} \cap S \neq \emptyset$ .  $S$  is open while  $H_{d+1}$  is closed, so  $S \setminus (H_{d+1} \cap S)$  is open. Moreover,  $S$  is convex, and slicing a convex open set with

a hyperplane leaves two connected components, so  $S \setminus (H_{d+1} \cap S)$  must have two connected components. In summary, either  $H_{d+1} \cap S = \emptyset$ , implying  $S \in \mathcal{S}(\mathcal{A}_{d+1})$ , or  $S$  decomposes into two connected components both of which are in  $\mathcal{S}(\mathcal{A}_{d+1})$ . Following Lemma A.2, this implies

$$a(\mathcal{A}_{d+1}) = a(\mathcal{A}_d) + a(\mathcal{A}_d \cap H_{d+1}) \quad (\text{A.2})$$

Eq. (A.2), using induction over  $d$ , gives

$$a(\mathcal{A}) = a(\mathcal{A}_D) = a(\mathcal{A}_1) + \sum_{d=1}^{D-1} a(H_{d+1} \cap \mathcal{A}_d) \quad (\text{A.3})$$

But  $H_{d+1} \cap \mathcal{A}_d$  is a  $(n-1)$ -dimensional hyperplane arrangement, so by the inductive assumption, which is in general position by Lemma A.1. Therefore, we can apply the inductive assumption to conclude  $a(H_{d+1} \cap \mathcal{A}_d) = a_{n-1}^d$ . Using that  $\mathcal{A}_1$  consists of one hyperplane, i.e.,  $a(\mathcal{A}_1) = 2$ , eq. (A.2) becomes

$$a(\mathcal{A}) = 2 + \sum_{d=1}^{D-1} a(H_{d+1} \cap \mathcal{A}_d) = 2 + \sum_{d=1}^{D-1} a_{n-1}^d. \quad (\text{A.4})$$

Substituting the explicit expression (6.51) in (A.4), we get

$$\begin{aligned} a(\mathcal{A}) &= 2 + \sum_{d=1}^{n-1} 2^d + 2 \sum_{d=n-1}^{D-2} \sum_{j=0}^{n-2} \binom{d}{j} \\ &= 2^n + 2 \sum_{d=n}^{D-1} \sum_{j=0}^{n-2} \binom{d-1}{j} = 2^n + 2 \sum_{d=n}^{D-1} \sum_{j=0}^{n-2} \binom{d-1}{j}. \end{aligned} \quad (\text{A.5})$$

Now expand  $2^n = 2 \sum_{j=0}^{n-1} \binom{n-1}{j}$  and use the identity  $\binom{n-1}{j} = \sum_{d=0}^{n-2} \binom{d}{j-1}$  for  $j > 0$  to rewrite

$$2^n = 2 \left( 1 + \sum_{j=1}^{n-1} \sum_{d=0}^{n-2} \binom{d}{j-1} \right) = 2 \left( 1 + \sum_{j=0}^{n-2} \sum_{d=0}^{n-2} \binom{d}{j} \right) \quad (\text{A.6})$$

Substituting in (A.5), we have

$$\begin{aligned} a(\mathcal{A}) &= 2 \left( 1 + \sum_{j=0}^{n-2} \sum_{d=0}^{D-2} \binom{d}{j} \right) = 2 \left( D + \sum_{j=1}^{n-2} \sum_{d=0}^{D-2} \binom{d}{j} \right) \\ &= 2 \left( D + \sum_{j=1}^{n-2} \binom{D-1}{j+1} \right) = 2 \sum_{j=0}^{n-1} \binom{D-1}{j}, \end{aligned} \quad (\text{A.7})$$

where in the last step we used  $\binom{D-1}{1} = D-1$  and  $\binom{D-1}{0} = 1$ . This confirms that (6.51) fulfills the recursion (A.4), completing the proof.  $\square$

# Appendix B

## Density of $\mathcal{W}_I^r$

Here, we prove that  $\mathcal{W}_I^r$  is a dense and open subset of  $\mathcal{W}_r$ .

**Lemma B.1:**  $\mathcal{W}_I^r$  is dense.

*Proof:* Let  $W$  be an arbitrary element of  $\mathcal{W}^r$ . Our strategy will be as follows: We will first prove that for some basis  $\{u_i\}_{i=1}^{D-r}$  of  $N \equiv \ker \hat{W}$ , there exists a set of vectors  $\{\tilde{u}_i\}_{i=1}^{D-r}$  such that  $\max_i \|\tilde{u}_i - u_i\| < \epsilon$  for any  $\epsilon > 0$ , and such that  $N' = \text{span}\{\tilde{u}_i\}$  satisfies

$$\dim(N' \cap H_{\alpha_1} \cap \cdots \cap H_{\alpha_d}) = \max(\dim N' - d, 0) = \max(\dim N - d, 0) \quad (\text{B.1})$$

In words,  $N'$  has the same dimension as  $N$ , and yields a subspace of the lowest possible dimension when intersected with all the  $H_{\alpha_j}$ . Once we have constructed this subspace, the vectors  $\{\tilde{u}_i\}_{i=1}^{D-r}$  can be orthogonalized using a Gram-Schmidt procedure, yielding an orthonormal basis  $\{u'_i\}_{i=1}^{D-r}$  for  $N'$ :

$$u'_i = \tilde{u}_i - \sum_{j < i} u'_j \langle \tilde{u}_i, u'_j \rangle. \quad (\text{B.2})$$

One can use induction over  $i$  together with  $\|\tilde{u}_i - u_i\| < \epsilon$  and the orthogonality of the  $u_i$ 's to prove that  $\|u'_i - \tilde{u}_i\| = \mathcal{O}(\epsilon)$ . The triangle inequality then gives

$$\|u'_i - u_i\| \leq \|u'_i - \tilde{u}_i\| + \|\tilde{u}_i - u_i\| = \mathcal{O}(\epsilon). \quad (\text{B.3})$$

We now construct a new interaction operator  $\hat{W}'$ . Denote with  $\{\phi_i\}_{i=1}^r$  the eigenvectors of  $\hat{W}$ . We define a new set of eigenvectors  $\{\phi'_i\}$  with another application of Gram-Schmidt:

$$\phi'_i = \phi_i - \sum_{j=1}^{D-r} u'_j \langle u'_j, \phi_i \rangle - \sum_{j < i} \phi'_j \langle \phi_j, \phi_i \rangle \quad (\text{B.4})$$

We define  $\hat{W}'$  as the operator with eigenvectors  $\phi'_i$  and corresponding eigenvalues identical to those of  $\hat{W}$ . Applying the same argument as for the  $u'_i$  above, we see that the new basis states fulfill  $\|\phi'_i - \phi_i\| = \mathcal{O}(\epsilon)$ , so  $\|\hat{W} - \hat{W}'\| = \mathcal{O}(\epsilon)$ . Moreover,  $\ker[\hat{W}'] = N'$  and  $N'$  has the same dimension as  $N$ , so  $\hat{W}' \in \mathcal{W}^r$ . To show that  $\hat{W}' \in \mathcal{W}_I^r$ , it remains to construct  $\{\tilde{u}_i\}$  such that  $N'$  fulfills (B.1).

Let  $\{v_i\}_{i=1}^{D-d}$  be a basis for  $H_{\alpha_1} \cap \cdots \cap H_{\alpha_d}$  (e.g., the appropriate coordinate vectors). Arrange the  $u_i$ 's and  $v_i$ 's as column vectors of a  $D \times (2D - d - r)$  rectangular matrix

$$M \equiv [u_1, \dots, v_1, \dots] \quad (\text{B.5})$$

The well-known formula  $\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V)$  gives

$$\dim(N \cap H_{\alpha_1} \cap \cdots \cap H_{\alpha_d}) = \dim(N) + D - d - \text{rank}[M], \quad (\text{B.6})$$

where we have used that  $H_{\alpha_j}$  are the coordinate hyperplanes in  $\mathbb{R}^D$ , so  $\dim[H_{\alpha_1} \cap \cdots \cap H_{\alpha_d}] = D - d$ . Our task is to construct  $\{\tilde{u}_i\}$  such that (B.1) and  $\|\tilde{u}_i - u_i\| < \epsilon$  hold. Equivalently, we can construct a matrix

$$M' \equiv [\tilde{u}_1, \dots, v_1, \dots] \quad (\text{B.7})$$

that shares its  $D - d$  last columns with  $M$ . Recall that  $N' \equiv \text{span}\{\tilde{u}_1, \dots, \tilde{u}_{D-r}\}$ , so (B.1) is equivalent to

$$\dim(N') + D - d - \text{rank}[M'] = \max(\dim N' - d, 0). \quad (\text{B.8})$$

This will be achieved if

$$\text{rank}[M'] = \min(\dim(N') + D - d, D) = \min(2D - d - r, D),$$

i.e., if  $M'$  has full rank. In other words, it suffices to show that for every  $D \times (2D - d - r)$  rectangular matrix  $M$  with the last  $D - d$  columns fixed to  $(v_i)_{i=1}^{D-d}$ , there exists a sequence  $M'^{(k)}$  of  $D \times (2D - d - r)$  rectangular matrices that converge to  $M$  and also have  $(v_i)_{i=1}^{D-d}$  as their rightmost columns. To do so, we perform a singular value decomposition of  $M$ :

$$M = UAV^T, \quad (\text{B.9})$$

where  $U$  and  $V$  are orthogonal square matrices and  $A$  is a  $D \times (2D - d - r)$  diagonal matrix. If  $A$  has full rank, so does  $M$ , and we are done. If it does not have full rank, there trivially exists a sequence of full-rank diagonal  $D \times (2D - d - r)$  matrices  $(A^{(k)})_{k=1}^\infty$  that converge to  $A$  (e.g., with the missing diagonal entries of  $A$  set to  $1/k$ ). The sequence  $M^{(k)} = UA^{(k)}V^T$  then consists of full-rank matrices that converge to  $M$ , but the last  $D - d$  columns  $(v_i^{(k)})_{i=1}^{D-d}$  do not necessarily agree with  $(v_i)_{i=1}^{D-d}$ . However, they are orthogonal and  $\lim_{k \rightarrow \infty} v_i^{(k)} = v_i \in \mathbb{R}^D$ , so there exists a sequence of orthogonal transformations  $R^{(k)} \in SO(D)$  such that  $R^{(k)}v_i^{(k)} = v_i$  and  $\lim_{k \rightarrow \infty} R^{(k)} = I$ . The sequence  $M'^{(k)} = R^{(k)}UA^{(k)}V^T$  still converges to  $M$  and has full rank because  $R^{(k)}$  is orthogonal, but it also has  $(v_i)_{i=1}^{D-d}$  for its last  $D - d$  columns.

We have shown that the set of full-rank matrices is dense in the set of  $D \times (2D - d - r)$  matrices with  $(v_i)_{i=1}^{D-d}$  as their rightmost columns. Therefore, there exists a matrix  $M'$  in this set that satisfies  $\|M - M'\| < \epsilon$ . Letting  $(\tilde{u}_i)_{i=1}^{D-r}$  be the first  $D - r$  columns of  $M'$ , it follows that  $\max_i \|\tilde{u}_i - u_i\| < \epsilon$  and that  $N' = \text{span}\{\{u_i\}_{i=1}^{D-r}\}$  satisfies (B.1), completing the proof.  $\square$

**Lemma B.2:**  $\mathcal{W}_I^r$  is open.

*Proof:* We will show that the complement is closed, and begin by recalling that  $\hat{W} \notin \mathcal{W}_I^r$  is equivalent to

$$\dim(N \cap H_{\alpha_1} \cap \dots \cap H_{\alpha_d}) > \max(\dim N - d, 0). \quad (\text{B.10})$$

We also define  $N = \ker[\hat{W}]$  as before.

Consider an arbitrary operator  $\hat{W} \notin \mathcal{W}_I^r$ . Constructing  $M$  as in (6.32) and using  $\dim(U \cap V) = \dim(U) + \dim(V) - \dim(U + V)$  reveals that (B.10) is equivalent to

$$\dim(N) + D - d - \text{rank}[M] > \max(\dim N - d, 0), \quad (\text{B.11})$$

which in turn is equivalent to

$$\text{rank}[M] < \min(\dim(N) + D - d, D) = \min(2D - d - r, D),$$

i.e., if  $M$  does not have maximal rank.

We have shown that  $\hat{W} \notin \mathcal{W}_I^r$  is equivalent to  $M$  not having full rank. Let  $(\hat{W}^{(k)})_{k=1}^\infty$  be a sequence such that  $\hat{W}^{(k)} \notin \mathcal{W}_I^r$  and  $\lim_{k \rightarrow \infty} \hat{W}^{(k)} = \hat{W}$  for some  $\hat{W} \in \mathcal{W}^r$ . If we construct  $M^{(k)}$  for each of the  $\hat{W}^{(k)}$ , then  $M^{(k)}$  has less than full rank for all  $k$ . But the subset of matrices with zero determinant is closed, and  $M$  in (6.32) is a continuous function of  $\hat{W}$ , so the matrix  $M$  constructed from the accumulation point  $\hat{W}$  must also have determinant zero and therefore  $\hat{W} \notin \mathcal{W}_I^r$ . We conclude that the complement of  $\mathcal{W}_I^r$  is closed.  $\square$

*Proof of Lemma 6.5.2:* Trivially follows from lemmas B.1 and B.2.

# Appendix C

## Density of $\mathcal{W}_{\hat{U}}^r$

Here we prove lemma 6.5.5 in two steps.

**Lemma C.1:**  $\mathcal{W}_{\hat{U}}^r$  is dense.

*Proof:* Assume that  $N = \ker[\hat{W}]$  satisfies  $U(N) = N$ , and let  $\{u_i\}_{i=1}^{D-r}$  be a basis for  $N$ . We will show that for any  $\epsilon > 0$ , there exists a vector  $\tilde{u}_1 \in \mathbb{C}^D$  with  $\|\tilde{u}_1 - u_1\| < \epsilon$  such that

$$N' = \text{span} [\{\tilde{u}_1\} \cup \{u_i\}_{i=2}^{D-r}] \quad (\text{C.1})$$

is not invariant under  $U$ . Note first that since  $U$  is unitary,  $U(N) = N$  implies  $U(N^\perp) = N^\perp$  where  $N^\perp$  is the orthogonal complement with respect to the canonical scalar product. Moreover, since  $U \neq \pm I$ , we either have  $U_{N^\perp} \neq I_{N^\perp}$  or  $U_N \neq I_N$ . We deal with these two cases in turn.

Case 1: Assume that  $U_{N^\perp} \neq I_{N^\perp}$ . Then there exists  $v \in N^\perp$  with  $|v| = 1$  such that  $|\langle v, Uv \rangle| < 1$ . Since  $U_N$  is a unitary on  $N$ , we can pick the basis for  $N$  such that  $Uu_1 = e^{i\phi}u_1$  for some  $\phi$ . We define  $\tilde{u}_1 = u_1 + \delta v$ . Define the subspace  $\tilde{N} = \text{span} [\{u_i\}_{i=1}^{D-r} \cup \{v\}]$  (note that the basis includes  $u_1$ , so  $\tilde{N}$  has dimension  $D - r + 1$ ). Let us apply  $U$  to  $\tilde{u}_1$  and compute its component along  $\tilde{N}$ :

$$\begin{aligned} \|P_{\tilde{N}}(U\tilde{u}_1)\|^2 &= |\langle v, U\tilde{u}_1 \rangle|^2 + \sum_{j=1}^{D-r} |\langle u_j, U\tilde{u}_1 \rangle|^2 \\ &= |\langle v, U(u_1 + \delta v) \rangle|^2 + \sum_{j=1}^{D-r} |\langle u_j, U(u_1 + \delta v) \rangle|^2 \\ &= \delta^2 |\langle v, Uv \rangle|^2 + \sum_{j=1}^{D-r} |\langle u_j, Uu_1 \rangle|^2 \\ &= \delta^2 |\langle v, Uv \rangle|^2 + \|P_N(Uu_1)\|^2 \\ &< \|u_1 + \delta v\|^2 \\ &= \|\tilde{u}_1\|^2 = \|U\tilde{u}_1\|^2 \end{aligned} \quad (\text{C.2})$$

In the second line, we have substituted the definition of  $\tilde{u}_1$ . In the third line, we used that  $U$  leaves both  $N$  and  $N^\perp$  invariant. In the inequality, we used  $|\langle v, Uv \rangle| < 1$  as well as  $\|P_N(Uu_1)\|^2 = 1$ , and the final line again substitutes the definition of  $\tilde{u}_1$ . The conclusion  $\|P_{\tilde{N}}(U\tilde{u}_1)\|^2 < \|U\tilde{u}_1\|^2$  implies that  $U\tilde{u}_1 \notin \tilde{N}$ . But  $N' \subset \tilde{N}$ , so in particular  $U\tilde{u}_1 \notin N'$ .

Case 2: Assume that  $U_N \neq I_N$ . Then there exists a basis for  $N$  where  $|\langle u_1, Uu_1 \rangle| < 1$ . This time, because  $U_{N^\perp}$  is unitary on  $N^\perp$ , there exists  $v \in N^\perp$  such that  $Uv = e^{i\phi}v$ . As for case 1, we define  $\tilde{u}_1 = u_1 + \delta v$ .

This time, we must use the projection into  $N'$ :

$$\begin{aligned}
\|P_{N'}(U\tilde{u}_1)\|^2 &= \frac{|\langle \tilde{u}_1, U\tilde{u}_1 \rangle|^2}{\|\tilde{u}_1\|^2} + \sum_{j=2}^{D-r} |\langle u_j, U\tilde{u}_1 \rangle|^2 \\
&= \frac{|\langle u_1 + \delta v, U(u_1 + \delta v) \rangle|^2}{1 + \delta^2} + \sum_{j=2}^{D-r} |\langle u_j, U(u_1 + \delta v) \rangle|^2 \\
&= \frac{|\langle u_1, Uu_1 \rangle + \delta^2 \langle v, Uv \rangle|^2}{1 + \delta^2} + \sum_{j=2}^{D-r} |\langle u_j, Uu_1 \rangle|^2 \\
&\leq \frac{(|\langle u_1, Uu_1 \rangle| + \delta^2)^2}{1 + \delta^2} + \sum_{j=2}^{D-r} |\langle u_j, Uu_1 \rangle|^2 \\
&= \delta^2 \frac{\delta^2 + 2|\langle u_1, Uu_1 \rangle| - |\langle u_1, Uu_1 \rangle|^2}{1 + \delta^2} + |\langle u_1, Uu_1 \rangle|^2 + \sum_{j=2}^{D-r} |\langle u_j, Uu_1 \rangle|^2 \\
&= \delta^2 \frac{\delta^2 + 2|\langle u_1, Uu_1 \rangle| - |\langle u_1, Uu_1 \rangle|^2}{1 + \delta^2} + \sum_{j=1}^{D-r} |\langle u_j, Uu_1 \rangle|^2 \\
&< \delta^2 + \sum_{j=1}^{D-r} |\langle u_j, Uu_1 \rangle|^2 \\
&= \delta^2 + 1 \\
&= \|U\tilde{u}_1\|^2
\end{aligned} \tag{C.3}$$

In the second line, we have substituted the definition of  $\tilde{u}_1$ . In the third line, we used that  $N$  and  $N^\perp$  are both invariant under  $U$ . In the fourth line, we used that  $|\langle v, Uv \rangle| = 1$ , and  $|x + y|^2 \leq (|x| + |y|)^2$  for  $x, y \in \mathbb{C}$ . The fifth line is a straightforward algebraic manipulation. The sixth line absorbs the penultimate term into the sum, and the strict inequality uses that  $|\langle u_1, Uu_1 \rangle| < 1$ , and that  $x(1 - x) < 1$  for  $0 < x < 1$ . The penultimate line uses invariance of  $N$  under  $U$ , and the final line uses the definition of  $\tilde{u}_1$  and the orthogonality of  $u_1$  and  $v$ . We have thus shown that  $\|P_{N'}(U\tilde{u}_1)\|^2 < \|U\tilde{u}_1\|^2$  and therefore,  $U\tilde{u}_1 \notin N'$ .

We learn that in both cases,  $U\tilde{u}_1 \notin N'$ , which by the definition (C.1) implies that  $U(N') \neq N'$ . Moreover, in both cases, this assertion holds for  $\tilde{u}_1 = u_1 + \delta v$  for arbitrarily small  $v$  so that we can satisfy the constraint  $\|\tilde{u}_1 - u_1\|$  stipulated in the beginning. It remains to construct an operator  $\hat{W}'$  that is arbitrarily close to  $\hat{W}$  and has  $N'$  as its kernel. This can be done with the same procedure as in App. B.  $\square$

**Lemma C.2:**  $\mathcal{W}_{\hat{U}}^r$  is open

*Proof:* We show that the complement is closed. Suppose that  $(\hat{W}^{(k)})_{k=1}^\infty$  is a sequence converging to some  $\hat{W} \in \mathcal{W}^r$  and such that  $\hat{W}^{(k)} \notin \mathcal{W}_{\hat{U}}^r$ . By definition of  $\mathcal{W}_{\hat{U}}^r$ , we have

$$\hat{W}^{(k)}v = 0 \implies \hat{W}^{(k)}\hat{U}v = 0 \tag{C.4}$$

for all  $v \in \mathbb{R}^D$ . Now pick an arbitrary vector  $v$  and suppose that  $\hat{W}v = 0$ . This implies  $\hat{W}^{(k)}v \rightarrow 0$  by continuity, and therefore  $\hat{P}_{(N^{(k)})^\perp}v \rightarrow 0$ , where  $\hat{P}_{(N^{(k)})^\perp}$  is the projector into the subspace orthogonal to  $N^{(k)} = \ker \hat{W}^{(k)}$ . This, together with  $\hat{W}^{(k)} \notin \mathcal{W}_{\hat{U}}^r$  implies that  $\hat{W}^{(k)}\hat{U}v \rightarrow 0$  and therefore  $\hat{W}\hat{U}v = 0$  by continuity. Since this is true for arbitrary  $v \in \ker[\hat{W}]$ , we conclude that  $\hat{W} \notin \mathcal{W}_{\hat{U}}^r$ .  $\square$

*Proof of Lemma 6.5.5:* Follows trivially from lemmas C.1 and C.2

# Bibliography

- [1] U. Schollwöck. “The density-matrix renormalization group”. In: *Rev. Mod. Phys.* 77 (1 2005), pp. 259–315. DOI: [10.1103/RevModPhys.77.259](https://doi.org/10.1103/RevModPhys.77.259). URL: <https://link.aps.org/doi/10.1103/RevModPhys.77.259>.
- [2] U. Schollwöck. “The density-matrix renormalization group in the age of matrix product states”. In: *Ann. of Phys.* 326 (Aug. 2010). DOI: [10.1016/j.aop.2010.09.012](https://doi.org/10.1016/j.aop.2010.09.012).
- [3] A. Georges et al. “Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions”. In: *Rev. Mod. Phys.* 68 (1 1996), pp. 13–125. DOI: [10.1103/RevModPhys.68.13](https://doi.org/10.1103/RevModPhys.68.13). URL: <https://link.aps.org/doi/10.1103/RevModPhys.68.13>.
- [4] D. Vollhardt. “Dynamical mean-field theory for correlated electrons”. In: *Annalen der Physik* 524.1 (2012), pp. 1–19. DOI: <https://doi.org/10.1002/andp.201100250>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/andp.201100250>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.201100250>.
- [5] P. Hohenberg and W. Kohn. “Inhomogeneous Electron Gas”. In: *Phys. Rev.* 136 (3B 1964), B864. DOI: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864). URL: <https://link.aps.org/doi/10.1103/PhysRev.136.B864>.
- [6] W. Kohn and L. J. Sham. “Self-Consistent Equations Including Exchange and Correlation Effects”. In: *Phys. Rev.* 140 (4A 1965), A1133–A1138. DOI: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133). URL: <https://link.aps.org/doi/10.1103/PhysRev.140.A1133>.
- [7] Julia Liebert. “Reduced Density Matrix Functional Theory for Bosons: Foundations and Applications”. Master’s Thesis. Ludwig-Maximilians-Universität München, 2021.
- [8] Yeong E. Kim and Alexander L. Zubarev. “Density-functional theory of bosons in a trap”. In: *Phys. Rev. A* 67 (1 2003), p. 015602. DOI: [10.1103/PhysRevA.67.015602](https://doi.org/10.1103/PhysRevA.67.015602). URL: <https://link.aps.org/doi/10.1103/PhysRevA.67.015602>.
- [9] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness (Series of Books in the Mathematical Sciences)*. First Edition. W. H. Freeman, 1979. ISBN: 0716710455. URL: <http://www.amazon.com/Computers-Intractability-NP-Completeness-Mathematical-Sciences/dp/0716710455>.
- [10] Tomasz Maciążek. “Repulsively diverging gradient of the density functional in the reduced density matrix functional theory”. In: *New Journal of Physics* 23.11 (2021), p. 113006. DOI: [10.1088/1367-2630/ac309c](https://doi.org/10.1088/1367-2630/ac309c). URL: <https://dx.doi.org/10.1088/1367-2630/ac309c>.
- [11] C. L. Benavides-Riveros et al. “Reduced Density Matrix Functional Theory for Bosons”. In: *Phys. Rev. Lett.* 124 (18 2020), p. 180603. DOI: [10.1103/PhysRevLett.124.180603](https://doi.org/10.1103/PhysRevLett.124.180603). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.124.180603>.
- [12] Julia Liebert and Christian Schilling. “Functional theory for Bose-Einstein condensates”. In: *Phys. Rev. Res.* 3 (1 2021), p. 013282. DOI: [10.1103/PhysRevResearch.3.013282](https://doi.org/10.1103/PhysRevResearch.3.013282). URL: <https://link.aps.org/doi/10.1103/PhysRevResearch.3.013282>.
- [13] C. Kapelle, C. A. Ullrich, and Vignale G. “Degenerate ground states and nonunique potentials: Breakdown and restoration of density functionals”. In: *Phys. Rev. A* 76 (2007), p. 012508. DOI: [10.1103/PhysRevA.76.012508](https://doi.org/10.1103/PhysRevA.76.012508). URL: <https://journals.aps.org/prapdf/10.1103/PhysRevA.76.012508>.
- [14] M. Penz and R. Van Leeuwen. “Density-functional theory on graphs”. In: *Journal of Chemical Physics* 155 (24 2021), p. 244111. DOI: [10.1063/5.0074249](https://doi.org/10.1063/5.0074249). URL: <https://arxiv.org/abs/2106.15370>.
- [15] Markus Penz and Robert van Leeuwen. “Geometry of Degeneracy in Potential and Density Space”. In: *Quantum* 7 (Feb. 2023), p. 918. ISSN: 2521-327X. DOI: [10.22331/q-2023-02-09-918](https://doi.org/10.22331/q-2023-02-09-918). URL: <https://doi.org/10.22331/q-2023-02-09-918>.

- [16] Jerzy Cioslowski and Katarzyna Pernal. “Size versus volume extensivity of a new class of density matrix functionals”. In: *The Journal of Chemical Physics* 120.22 (May 2004), pp. 10364–10367. ISSN: 0021-9606. DOI: [10.1063/1.1738411](https://doi.org/10.1063/1.1738411). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/120/22/10364/10857228/10364\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/120/22/10364/10857228/10364_1_online.pdf). URL: <https://doi.org/10.1063/1.1738411>.
- [17] Jerzy Cioslowski. “New constraints upon the electron-electron repulsion energy functional of the one-electron reduced density matrix”. In: *The Journal of Chemical Physics* 123.16 (Oct. 2005), p. 164106. ISSN: 0021-9606. DOI: [10.1063/1.2074527](https://doi.org/10.1063/1.2074527). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2074527/15372885/164106\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.2074527/15372885/164106_1_online.pdf). URL: <https://doi.org/10.1063/1.2074527>.
- [18] Daniel R. Rohr and Katarzyna Pernal. “Open-shell reduced density matrix functional theory”. In: *The Journal of Chemical Physics* 135.7 (Aug. 2011), p. 074104. ISSN: 0021-9606. DOI: [10.1063/1.3624609](https://doi.org/10.1063/1.3624609). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.3624609/15440403/074104\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.3624609/15440403/074104_1_online.pdf). URL: <https://doi.org/10.1063/1.3624609>.
- [19] Katarzyna Pernal. “Excitation energies from range-separated time-dependent density and density matrix functional theory”. In: *The Journal of Chemical Physics* 136.18 (May 2012), p. 184105. ISSN: 0021-9606. DOI: [10.1063/1.4712019](https://doi.org/10.1063/1.4712019). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4712019/15450623/184105\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/1.4712019/15450623/184105_1_online.pdf). URL: <https://doi.org/10.1063/1.4712019>.
- [20] Klaas J. H. Giesbertz, Anna-Maija Uimonen, and Robert van Leeuwen. “Approximate energy functionals for one-body reduced density matrix functional theory from many-body perturbation theory”. In: *The European Physical Journal B* 91.11 (2018), p. 282. ISSN: 1434-6036. DOI: [10.1140/epjb/e2018-90279-1](https://doi.org/10.1140/epjb/e2018-90279-1). URL: <https://doi.org/10.1140/epjb/e2018-90279-1>.
- [21] Jian Wang and Peter J. Knowles. “Nonuniqueness of algebraic first-order density-matrix functionals”. In: *Phys. Rev. A* 92 (1 2015), p. 012520. DOI: [10.1103/PhysRevA.92.012520](https://doi.org/10.1103/PhysRevA.92.012520). URL: <https://link.aps.org/doi/10.1103/PhysRevA.92.012520>.
- [22] Katarzyna Pernal and Jerzy Cioslowski. “Phase dilemma in density matrix functional theory”. In: *The Journal of Chemical Physics* 120.13 (Apr. 2004), pp. 5987–5992. ISSN: 0021-9606. DOI: [10.1063/1.1651059](https://doi.org/10.1063/1.1651059). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/120/13/5987/19266414/5987\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/120/13/5987/19266414/5987_1_online.pdf). URL: <https://doi.org/10.1063/1.1651059>.
- [23] I. Mitxelena, M. Mayorga, and M. Piris. “Phase Dilemma in Natural Orbital Functional Theory from the N-representability Perspective”. In: *The European Physical Journal B* 91 (May 2018), p. 109. DOI: [10.1140/epjb/e2018-90078-8](https://doi.org/10.1140/epjb/e2018-90078-8).
- [24] Tim Baldsiefen, Attila Cangi, and E. K. U. Gross. “Reduced-density-matrix-functional theory at finite temperature: Theoretical foundations”. In: *Phys. Rev. A* 92 (5 2015), p. 052514. DOI: [10.1103/PhysRevA.92.052514](https://doi.org/10.1103/PhysRevA.92.052514). URL: <https://link.aps.org/doi/10.1103/PhysRevA.92.052514>.
- [25] C. Schilling and R. Schilling. “Diverging Exchange Force and Form of the Exact Density Matrix Functional”. In: *Phys. Rev. Lett.* 122 (1 2019), p. 013001. DOI: [10.1103/PhysRevLett.122.013001](https://doi.org/10.1103/PhysRevLett.122.013001). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.122.013001>.
- [26] Oleg V. Gritsenko, Jian Wang, and Peter J. Knowles. “Symmetry dependence and universality of practical algebraic functionals in density-matrix-functional theory”. In: *Phys. Rev. A* 99 (4 2019), p. 042516. DOI: [10.1103/PhysRevA.99.042516](https://doi.org/10.1103/PhysRevA.99.042516). URL: <https://link.aps.org/doi/10.1103/PhysRevA.99.042516>.
- [27] Jerzy Cioslowski. “Off-diagonal derivative discontinuities in the reduced density matrices of electronic systems”. In: *The Journal of Chemical Physics* 153.15 (Oct. 2020), p. 154108. ISSN: 0021-9606. DOI: [10.1063/5.0023955](https://doi.org/10.1063/5.0023955). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0023955/9683658/154108\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0023955/9683658/154108_1_online.pdf). URL: <https://doi.org/10.1063/5.0023955>.
- [28] Jerzy Cioslowski. “One-Electron Reduced Density Matrix Functional Theory of Spin-Polarized Systems”. In: *Journal of Chemical Theory and Computation* 16.3 (2020), pp. 1578–1585. ISSN: 1549-9618. DOI: [10.1021/acs.jctc.9b01155](https://doi.org/10.1021/acs.jctc.9b01155). URL: <https://doi.org/10.1021/acs.jctc.9b01155>.
- [29] K. J. H. Giesbertz. “Implications of the unitary invariance and symmetry restrictions on the development of proper approximate one-body reduced-density-matrix functionals”. In: *Phys. Rev. A* 102 (5 2020), p. 052814. DOI: [10.1103/PhysRevA.102.052814](https://doi.org/10.1103/PhysRevA.102.052814). URL: <https://link.aps.org/doi/10.1103/PhysRevA.102.052814>.

- [30] Jerzy Cioslowski. “Construction of explicitly correlated one-electron reduced density matrices”. In: *The Journal of Chemical Physics* 153.22 (Dec. 2020), p. 224109. ISSN: 0021-9606. DOI: [10.1063/5.0031195](https://doi.org/10.1063/5.0031195). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0031195/15582147/224109\\_1\\_online.pdf](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0031195/15582147/224109_1_online.pdf). URL: <https://doi.org/10.1063/5.0031195>.
- [31] Jerzy Cioslowski, Zsuzsanna É. Mihálka, and Ágnes Szabados. “Bilinear Constraints upon the Correlation Contribution to the Electron–Electron Repulsion Energy as a Functional of the One-Electron Reduced Density Matrix”. In: *Journal of Chemical Theory and Computation* 15.9 (2019), pp. 4862–4872. ISSN: 1549-9618. DOI: [10.1021/acs.jctc.9b00443](https://doi.org/10.1021/acs.jctc.9b00443). URL: <https://doi.org/10.1021/acs.jctc.9b00443>.
- [32] Julia Liebert et al. “Foundation of One-Particle Reduced Density Matrix Functional Theory for Excited States”. In: *Journal of Chemical Theory and Computation* 18.1 (2022), pp. 124–140. ISSN: 1549-9618. DOI: [10.1021/acs.jctc.1c00561](https://doi.org/10.1021/acs.jctc.1c00561). URL: <https://doi.org/10.1021/acs.jctc.1c00561>.
- [33] Daniel Gibney, Jan-Niklas Boyn, and David A. Mazziotti. “Density Functional Theory Transformed into a One-Electron Reduced-Density-Matrix Functional Theory for the Capture of Static Correlation”. In: *The Journal of Physical Chemistry Letters* 13.6 (2022), pp. 1382–1388. DOI: [10.1021/acs.jpcllett.2c00083](https://doi.org/10.1021/acs.jpcllett.2c00083). URL: <https://doi.org/10.1021/acs.jpcllett.2c00083>.
- [34] C. Schilling and S. Pittalis. “Ensemble reduced density matrix functional theory for excited states and hierarchical generalization of Pauli’s exclusion principle”. In: *arXiv:2106.02560* (2021). URL: <https://arxiv.org/abs/2106.02560>.
- [35] S. Di Sabatino et al. “Introducing screening in one-body density matrix functionals: Impact on charged excitations of model systems via the extended Koopmans’ theorem”. In: *Phys. Rev. B* 105 (23 2022), p. 235123. DOI: [10.1103/PhysRevB.105.235123](https://doi.org/10.1103/PhysRevB.105.235123). URL: <https://link.aps.org/doi/10.1103/PhysRevB.105.235123>.
- [36] Bruno Senjean et al. “Reduced density matrix functional theory from an ab initio seniority-zero wave function: Exact and approximate formulations along adiabatic connection paths”. In: *Phys. Rev. A* 106 (3 2022), p. 032203. DOI: [10.1103/PhysRevA.106.032203](https://doi.org/10.1103/PhysRevA.106.032203). URL: <https://link.aps.org/doi/10.1103/PhysRevA.106.032203>.
- [37] S. M. Sutter and K. J. H. Giesbertz. “One-body reduced density-matrix functional theory for the canonical ensemble”. In: *Phys. Rev. A* 107 (2 2023), p. 022210. DOI: [10.1103/PhysRevA.107.022210](https://doi.org/10.1103/PhysRevA.107.022210). URL: <https://link.aps.org/doi/10.1103/PhysRevA.107.022210>.
- [38] Julia Liebert, Adam Yanis Chaou, and Christian Schilling. “Refining and relating fundamentals of functional theory”. In: *The Journal of Chemical Physics* 158.21 (June 2023), p. 214108. ISSN: 0021-9606. DOI: [10.1063/5.0143657](https://doi.org/10.1063/5.0143657). eprint: [https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0143657/17914418/214108\\_1\\_5.0143657.pdf](https://pubs.aip.org/aip/jcp/article-pdf/doi/10.1063/5.0143657/17914418/214108_1_5.0143657.pdf). URL: <https://doi.org/10.1063/5.0143657>.
- [39] T. L. Gilbert. “Hohenberg-Kohn theorem for nonlocal external potentials”. In: *Phys. Rev. B* 12 (6 1975), p. 2111. URL: <https://link.aps.org/doi/10.1103/PhysRevB.12.2111>.
- [40] M. Levy. “Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem”. In: *Proc. Natl. Acad. Sci. U.S.A* 76.12 (1979), p. 6062. URL: <http://www.pnas.org/content/76/12/6062>.
- [41] E. H. Lieb. “Variational Principle for Many-Fermion Systems”. In: *Phys. Rev. Lett.* 46 (7 1981), pp. 457–459. DOI: [10.1103/PhysRevLett.46.457](https://doi.org/10.1103/PhysRevLett.46.457). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.46.457>.
- [42] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. DOI: [10.1017/CB09780511804441](https://doi.org/10.1017/CB09780511804441).
- [43] S. M. Valone. “Consequences of extending 1-matrix energy functionals from pure-state representable to all ensemble representable 1-matrices”. In: *J. Chem. Phys.* 73.3 (1980), p. 1344. DOI: [10.1063/1.440249](https://doi.org/10.1063/1.440249). URL: <https://doi.org/10.1063/1.440249>.
- [44] C. Schilling. “Communication: Relating the pure and ensemble density matrix functional”. In: *J. Chem. Phys.* 149.23 (2018), p. 231102. DOI: [10.1063/1.5080088](https://doi.org/10.1063/1.5080088). URL: <https://doi.org/10.1063/1.5080088>.
- [45] Thomas H. Cormen et al. *Introduction to Algorithms*. 2nd. The MIT Press, 2001. ISBN: 0262032937. URL: <http://www.amazon.com/Introduction-Algorithms-Thomas-H-Cormen/dp/0262032937%3FSubscriptionId%3D13CT5CVB80YFWJEPWS02%26tag%3Dws%26linkCode%3Dxm2%26camp%3D2025%26creative%3D165953%26creativeASIN%3D0262032937>.

- [46] Richard M. Karp. “Reducibility among Combinatorial Problems”. In: *Complexity of Computer Computations: Proceedings of a symposium on the Complexity of Computer Computations, held March 20–22, 1972, at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, and sponsored by the Office of Naval Research, Mathematics Program, IBM World Trade Corporation, and the IBM Research Mathematical Sciences Department*. Ed. by Raymond E. Miller, James W. Thatcher, and Jean D. Bohlinger. Boston, MA: Springer US, 1972, pp. 85–103. ISBN: 978-1-4684-2001-2. DOI: [10.1007/978-1-4684-2001-2\\_9](https://doi.org/10.1007/978-1-4684-2001-2_9). URL: [https://doi.org/10.1007/978-1-4684-2001-2\\_9](https://doi.org/10.1007/978-1-4684-2001-2_9).
- [47] L. Khachiyan. “A Polynomial Algorithm for Linear Programming”. In: *Doklady Akademii Nauk SSSR* 224 (5 1979), p. 1093.
- [48] G.B. Dantzig and M.N. Thapa. *Linear Programming 2: 2: Theory and Extensions*. Linear Programming. Springer, 1997. ISBN: 9780387986135. URL: <https://books.google.de/books?id=qUvXMT00PZwC>.
- [49] N. Karmarkar. “A new polynomial-time algorithm for linear programming”. In: *Combinatorica* 4.4 (1984), pp. 373–395. ISSN: 1439-6912. DOI: [10.1007/BF02579150](https://doi.org/10.1007/BF02579150). URL: <https://doi.org/10.1007/BF02579150>.
- [50] Julia Liebert and Christian Schilling. “Deriving density-matrix functionals for excited states”. In: *SciPost Phys.* 14 (2023), p. 120. DOI: [10.21468/SciPostPhys.14.5.120](https://doi.org/10.21468/SciPostPhys.14.5.120). URL: <https://scipost.org/10.21468/SciPostPhys.14.5.120>.
- [51] Richard P. Stanley. “An Introduction to Hyperplane Arrangements”. In: (2007). URL: <http://www.cis.upenn.edu/~cis610/sp06stanley.pdf>.

# Acknowledgements

I am deeply grateful to my supervisor, Dr. Christian Schilling, and my co-supervisors, Julia Liebert and Dr. Sebastian Paeckel, for their inspiring ideas and mentorship. I would, in particular, like to acknowledge Christian for proposing this fruitful research project and Julia for providing invaluable insights throughout its duration. I am incredibly thankful to all my fellow group members for the stimulating research environment that made this work possible. Finally, I want to thank my friends and family for their patient support.