Ludwig-Maximilians-Universität München
Faculty of Physics

Master's thesis

Geometry of Generalized Density Functional Theories

Chih-Chun Wang

Supervised by
Dr. Christian Schilling

Munich, October 5, 2025

Ludwig-Maximilians-Universität München
Fakultät für Physik

Masterarbeit

Geometrie der verallgemeinerten Dichtefunktionaltheorien

Chih-Chun Wang

Betreut von
Dr. Christian Schilling

München, den 5. Oktober 2025

Abstract

Density functional theory (DFT) is an indispensable *ab initio* method in both quantum chemistry and condensed matter physics. Based on recent advancements in reduced density matrix functional theory (RDMFT), a variant of DFT that is believed to be better suited for strongly correlated systems, we construct a mathematical framework generalizing all ground state functional theories, which in particular applies to fermionic, bosonic, and spin systems. We show that within the special class of such functional theories where the space of external potentials forms the Lie algebra of a compact Lie group, the N-representability problem is readily solved by applying techniques from the study of momentum maps in symplectic geometry, an approach complementary to Klyachko's famous solution to the quantum marginal problem. The "boundary force", a diverging repulsive force from the boundary of the functional's domain observed in previous works but only qualitatively understood in isolated systems, is studied quantitatively and extensively in our work. Specifically, we present a formula capturing the exact behavior of the functional close to the boundary. In the case where the Lie algebra is abelian, a completely rigorous proof of the boundary force formula based on Levy-Lieb constrained search is given. Our formula is a first step towards developing more accurate functional approximations, with the potential of improving current RDMFT approximate functionals such as the Piris natural orbital functionals (NOFs). All key concepts and ideas of our work are demonstrated in translation invariant bosonic lattice systems.

Zusammenfassung

Die Dichtefunktionaltheorie (DFT) ist eine unverzichtbare *ab initio*-Methode sowohl in der Quantenchemie als auch in der Festkörperphysik. Basierend auf jüngsten Fortschritten in der Funktionaltheorie der reduzierten Dichtematrix (RDMFT), einer Variante der DFT, die als besser geeignet für stark korrelierte Systeme gilt, entwickeln wir einen mathematischen Rahmen, der alle Grundzustands-Funktionaltheorien verallgemeinert und insbesondere auf fermionische, bosonische und Spin-Systeme anwendbar ist. Wir zeigen, dass innerhalb der speziellen Klasse solcher Funktionaltheorien, in denen der Raum der externen Potentiale die Lie-Algebra einer kompakten Lie-Gruppe bildet, das N-Repräsentierbarkeitsproblem sich leicht durch die Anwendung von Methoden aus der Untersuchung der Impulsabbildungen in der symplektischen Geometrie lösen lässt ein Ansatz, der komplementär zu Klyachkos berühmter Lösung des Quantenmarginalproblems ist. Die "Randkraft", eine divergierende abstoSSende Kraft von dem Rand des Funktionaldefinitionsbereichs, die in früheren Arbeiten beobachtet, aber in isolierten Systemen nur qualitativ verstanden wurde, wird in unserer Arbeit quantitativ und umfassend untersucht. Genauer gesagt stellen wir eine Formel vor, die das exakte Verhalten des Funktionals nahe des Rands erfasst. Im Fall, dass die Lie-Algebra abelsch ist, wird ein vollständig strenger Beweis der Randkraftformel auf Basis der Levy-Lieb'schen eingeschränkten Suche präsentiert. Unsere Formel stellt einen ersten Schritt zur Entwicklung genauerer Funktionalapproximationen dar, mit dem Potenzial, aktuelle RDMFT-Funktionalapproximationen wie Piris' Funktionale der natürlichen Orbitale (NOFs) zu verbessern. Alle zentralen Konzepte und Ideen unserer Arbeit werden an translationsinvarianten bosonischen Gittersystemen veranschaulicht.

Contents

Notation and Symbols

$\mathcal{F}_{HK}$	Hohenberg-Kohn functional
$\mathcal{F}_p$	pure functional
$\mathcal{F}_e$	ensemble functional
$\mathcal{P}$	set of pure states
$\mathcal{E}$	set of ensemble states
$\mathbb{P}(\mathcal{H})$	projective Hilbert space
w°	annihilator of a vector or a subset w
f^*	dual map of a linear map f
$\langle \cdot, \cdot \rangle$	inner product or dual pairing
conv	convex hull
aff	affine hull
relint	relative interior
$\overrightarrow{A}$	translation space of an affine space A
Δ_n	standard n-simplex
Δ_X	standard simplex over a finite set X, i.e., $\{t \in \mathbb{R}^X_{\geq 0} \mid \sum_{i \in X} t_i = 1\}$
$C^\infty(M)$	smooth functions on a smooth manifold M
$\mathfrak{X}(M)$	vector fields on M
$T_x M$	tangent space of M at a point x
$H^\bullet_{\mathrm{dR}}(M)$	de Rham cohomology of M
Df	derivative of a smooth map f
$D_x f$	derivative of f at a point x
$\mathrm{d}\alpha$	exterior derivative of a differential form α
$\iota_X \alpha$	contraction of a vector field X with α
$L_X \alpha$	Lie derivative of α along X

Chapter 1

Introduction

One fundamental problem in the physical sciences is that of understanding systems consisting of many constituents. In quantum chemistry, one is interested in the behavior of the electrons in an atom or molecule. A large part of solid-state physics is dedicated to the study of how electrons propagate and scatter in a lattice. In the realm of ultracold atoms, one studies systems of interacting bosonic or fermionic atoms trapped in optical lattices. If there is no time dependence, the properties of all of these systems, in particular those of the ground state, are described by the Schrödinger equation

$$H\left|\Psi\right\rangle = E\left|\Psi\right\rangle. \tag{1.1}$$

In principle, then, we know everything about the system once we know the Hamiltonian H, provided that we have the computational power to solve the eigenvalue equation (1.1). The problem is, of course, that for the vast majority of interesting systems, and by "interesting" we mean, among other things, that the number of constituents is large, directly solving Eq. (1.1) is practically impossible even numerically.

Various methods have been devised to overcome this problem by employing sophisticated and physically motivated approximations, each with its own merits. Density matrix renormalization group (DMRG) [1–4] is particularly well-suited for one-dimensional systems. Dynamical mean-field theory (DMFT) is especially effective with strongly correlated electronic systems [5–7]. In density functional theory (DFT) [8, 9], the role of the electronic wave function as the basic quantity is replaced by the electron density $\rho(\mathbf{r})$, and ground state properties are encoded in the density functional $\mathcal{F}(\rho)$. The functional is *universal* in the sense that it depends only on the kinetic energy of the electrons and the Coulomb interaction, but not on the applied external potential $v(\mathbf{r})$, which, for example, could be the electron-nucleus interaction. It has been an indispensable *ab-initio* numerical method in both solid-state physics and modern chemistry [10–12]. Practical implementations of DFT are typically based on the Kohn-Sham scheme [9] together with suitable approximations of the so-called exchange-correlation functional [9, 13–16].

Reduced density matrix functional theory (RDMFT) [17–21] has established itself as an alternative to DFT. With the one-particle reduced density matrix (1RDM) treated as the basic variable, the ground state problem is similarly reduced to that of determining the density *matrix* functional $\mathcal{F}(\gamma)$. Since the kinetic energy is a simple function of the 1RDM, the density matrix functional depends solely on the particle-particle interaction, which suggests that a

fictitious noninteracting Kohn-Sham system does not need to be introduced. Moreover, it has been shown [20, 22] that RDMFT performs better for transition metal oxides as compared to DFT with the local density approximation, which tends to predict metallic ground states as opposed to Mott insulators, indicating that RDMFT might be better suited for strongly correlated systems.

Nevertheless, as with DFT, the exact density matrix functional is usually unknown apart from few small systems [23–25]. Numerous proposed approximate functionals exist in the literature [20, 26–35], most of which are based on modifying Müller's reconstruction of the two-particle reduced density matrix (2RDM) in terms of the 1RDM [36]. Since RDMFT is formally exact (as is DFT), it is important to take into account known exact properties of the functional when proposing approximations. One such possibility is the behavior of the functional near the boundary of its domain. Recently, it has been observed that the functional exhibits a diverging repulsive force from the boundary in various fermionic and bosonic systems [25, 37–39], effectively preventing the 1RDM from reaching the boundary. In fermionic systems, this has been termed the *exchange force*, whereas in bosonic systems the same phenomenon has been referred to as the *BEC force*, since previous work on bosonic systems mostly focused on investigating the functional near the Bose-Einstein condensate (BEC) vertex. To be more precise, the observed force is always repulsive, and takes the form

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* + \epsilon\eta) \approx -\mathcal{G}\frac{1}{\sqrt{\epsilon}} \tag{1.2}$$

for some prefactor $\mathcal{G}$, where ρ_* is a density on the boundary of $\operatorname{dom}\mathcal{F}$, η is an inward-pointing vector, and ϵ is small. Not only is such a force intimately related to the presence or absence of *(quasi-) pinning* [40–46], understanding quantitatively why and how the derivative of the functional diverges as the density approaches the boundary of the domain could potentially lead to more accurate approximate functionals. However, these questions have not been addressed in the literature in a systematic way; the only rigorous quantitative demonstration which is not restricted to particular small systems is Ref. [46], where a bound of the functional $\mathcal{F}$ near the boundary has been found, which in turn implies that the derivative must diverge. Yet, there is no known formula for computing the prefactor $\mathcal{G}$.

In this thesis, we show that this diverging behavior of the derivative of the functional is entirely general and does not depend on the specific details of each system. In light of this, we often refer to the phenomenon as the *boundary force* to emphasize its geometric origin. More importantly, as will become apparent in the course of this thesis, the fact that such divergence exists does not even depend on the "type" of the functional theory itself, so boundary forces should also exist in ordinary DFT. To make these claims more precise, then, we shall set up a framework of functional theories which generalizes both DFT and RDMFT, which we do in Chapter 2. It turns out, however, that generality comes at the cost of tractability: in order to obtain useful statements, it seems necessary to restrict oneself to subclasses of functional theories satisfying certain "niceness conditions". We will explore two such classes: the *abelian* functional theories and *momentum map* functional theories, for each of which we derive a concrete formula for the boundary force prefactor $\mathcal{G}$.

The rest of this thesis is organized as follows. After a brief review of DFT and RDMFT in Section 1.1, we present a general framework of functional theories in Chapter 2, defining in the abstract setting concepts and objects such as density maps, representability, and universal functionals, which will be of central importance throughout the thesis. Chapter 2 is also

where basic results like the weak Hohenberg-Kohn theorem and the relations between the universal functionals are established. In Chapter 3, the abstract framework is instantiated by a functional theory describing translation invariant bosonic lattice systems. We show that the translation symmetry can be exploited to simplify RDFMT, with momentum space occupation numbers playing the role of density in the new functional theory. Having worked out concrete examples in Chapter 3, we move again to the abstract realm in Chapter 4 to develop the theory of what we call abelian functional theories, which include both (finite-dimensional) DFT and the momentum space functional theory in Chapter 3 as special cases and constitute arguably the simplest class of generalized functional theories. The central result in this chapter is Theorem 4.8, which establishes the existence of a diverging boundary force and provides a formula to compute the prefactor $\mathcal{G}$. Roughly two thirds of Chapter 4 is devoted to proving Theorem 4.8.

The last two chapters are dedicated to the study of a class of functional theories built upon representations of compact Lie groups, to which (finite-dimensional) RDMFT, DFT, and the functional theory of Chapter 3 all belong. We will refer to these as momentum map functional theories. The rigid structure allows us to translate the representability problem to that of characterizing images of *momentum maps* from symplectic geometry, an approach heavily explored and developed in quantum information theory, in particular in the quantum marginal problem (note: the word "momentum" in "momentum map" has nothing to do, for our purpose, with the familiar physical concept, although the two are etymologically related). In Chapter 5, we present a short introduction to the theory of momentum maps. The most relevant results from this chapter are Kirwan's theorem (Theorem 5.15), which is a statement about certain convexity properties of the image of a momentum map, and Theorem 5.20, which gives a way to characterize the image. Almost all content of Chapter 5, including these two theorems, is drawn from existing work in the literature, though largely unfamiliar to the functional theory community. In Chapter 6, we apply the machinery set up in the previous chapter to discuss momentum map functional theories themselves, with primary focus on the boundary behavior of the functional. The main result is Conjecture 6.1, which is the counterpart of Theorem 4.8. We put forward an argument for the conjecture based on perturbative calculations, although we do not have a rigorous proof at the moment.

1.1 Background: DFT and RDMFT

We quickly review the fundamentals of DFT and RDMFT, starting with the former. For DFT, nice introductions can be found in Refs. [47, 48]. Although there are extensions which accommodate finite temperatures and time dependence [49, 50], we will only be concerned with the ground state theory here and throughout the thesis. Subtleties of dealing with infinite-dimensional Hilbert spaces are swept aside, since we will only consider finite-dimensional ones from Chapter 2 onwards.

Consider a system of N electrons interacting via a pair interaction U (the exact form of U will not matter, but the most common choice is the Coulomb interaction), under the influence of an external potential $v(\mathbf{r})$. The Hamiltonian is

$$H(v) = \sum_{i=1}^{N} \frac{p_i^2}{2m} + \sum_{1\leq i<j\leq N} U(\mathbf{r}_i - \mathbf{r}_j) + \sum_{i=1}^{N} v(\mathbf{r}_i). \tag{1.3}$$

For example, the external potential $v(\mathbf{r})$ could come from a collection of nuclei around the electrons, in which case the Hamiltonian (1.3) would be that of an atom or molecule after the Born-Oppenheimer approximation. Moving the nuclei then corresponds to changing the external potential $v(\mathbf{r})$, and for each $v(\mathbf{r})$ there corresponds a ground state $|\Psi(v)\rangle$ (assuming no degeneracy for simplicity), from which we can compute the electron density $\rho(\mathbf{r})$ by

$$\rho(\mathbf{r}) = N \sum_{\sigma_1,\cdots,\sigma_N \in \{\uparrow,\downarrow\}} \int \langle \mathbf{r}, \sigma_1, \mathbf{r}_2, \sigma_2, \cdots, \mathbf{r}_N, \sigma_N | \Psi \rangle \langle \Psi | \mathbf{r}, \sigma_1, \mathbf{r}_2, \sigma_2, \cdots \mathbf{r}_N, \sigma_N \rangle \, \mathrm{d}\mathbf{r}_2 \cdots \mathrm{d}\mathbf{r}_N.$$

We then have a map

$$v \mapsto |\Psi(v)\rangle \mapsto \rho.$$

Hohenberg and Kohn [8] showed that this map is in fact invertible, in the sense that given the ground state electron density $\rho(\mathbf{r})$, the external potential $v(\mathbf{r})$ can be recovered up to a constant. Consequently, the ground state $|\Psi\rangle$ is a functional of ρ, and so is every property of the ground state. The *Hohenberg-Kohn functional* is defined as

$$\mathcal{F}_{HK}(\rho) := \langle \Psi(\rho) | T + U | \Psi(\rho) \rangle \,, \tag{1.4}$$

where T is the kinetic energy. One basic relation is

$$E(v) = \min_{\rho} \left[\int v(\mathbf{r}) \rho(\mathbf{r}) \mathrm{d}\mathbf{r} + \mathcal{F}_{HK}(\rho) \right], \tag{1.5}$$

where $E(v)$ is the ground state energy corresponding to the external potential v. Indeed, if ρ is the ground state density for v, then $E(v) = \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} + \mathcal{F}_{HK}(\rho)$. For any other density ρ', we have

$$\int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r} + \mathcal{F}_{HK}(\rho) \leq \langle \Psi(\rho') | H(v) | \Psi(\rho') \rangle = \int v(\mathbf{r})\rho'(\mathbf{r})\mathrm{d}\mathbf{r} + \mathcal{F}_{HK}(\rho').$$

Instead of $\mathcal{F}_{HK}$, one can define the *constrained search functionals* [51, 52]:

$$\begin{aligned} \mathcal{F}_p(\rho) &:= \min_{|\Psi\rangle \mapsto \rho} \langle \Psi | T + U | \Psi \rangle \\ \mathcal{F}_e(\rho) &:= \min_{\Gamma \mapsto \rho} \mathrm{Tr}(\Gamma(T+U)), \end{aligned} \tag{1.6}$$

where the minimizations are over N-particle pure states $|\Psi\rangle$ and ensemble states Γ which yield ρ as the density respectively. Eq. (1.5) still holds if we replace $\mathcal{F}_{HK}$ with $\mathcal{F}_p$ or $\mathcal{F}_e$.

The domain of $\mathcal{F}_{HK}$ is the set of densities ρ that are (pure) ground state electron densities for some external potential v, a condition on ρ known as *v-representability*. Similarly, the domain of $\mathcal{F}_p$ ($\mathcal{F}_e$), are the sets of densities ρ such that there exists a pure (ensemble) state which yields ρ as the density, a condition called pure (ensemble) state N-representability.

For RDMFT, one considers a family of Hamiltonians of the form

$$H(h) = h + U,$$

where h is a single-particle operator, and U is any interaction. For atoms or molecules, h consists of both the kinetic energy term and the electron-nucleus interaction, and U is the electron-electron repulsion. However, we imagine that we are free to choose any single-particle

operator h, so in particular we may alter the kinetic term. This is physically relevant, for example, in ultracold atom experiments, where the hopping rates of atoms in an optical lattice can be tuned.

For each h, there corresponds a ground state $|\Psi\rangle$ and a 1RDM $\gamma = N\mathrm{Tr}_{N-1}(|\Psi\rangle\langle\Psi|)$. A generalization of the Hohenberg-Kohn theorem [17] implies that the map

$$h \mapsto |\Psi\rangle \mapsto \gamma$$

is invertible (up to shifting h by constants). We can then similarly define the Hohenberg-Kohn functional (also called the Gilbert functional)

$$\mathcal{F}_{HK}(\gamma) := \langle\Psi(\gamma)|U|\Psi(\gamma)\rangle\,,$$

which satisfies

$$E(h) = \min_{\gamma}\left[\mathrm{Tr}(h\gamma) + \mathcal{F}_{HK}(\gamma)\right],$$

where $E(h)$ is the ground state energy of $H(h)$. The constrained search functionals can also be defined [53]:

$$\mathcal{F}_p(\gamma) := \min_{|\Psi\rangle\mapsto\gamma} \langle\Psi|U|\Psi\rangle$$
$$\mathcal{F}_e(\gamma) := \min_{\Gamma\mapsto\gamma} \mathrm{Tr}(\Gamma U)$$

v-representability, pure state N-representability, and ensemble state N-representability are defined analogously.

Chapter 2

Generalized Functional Theory

2.1 Motivation

In DFT, the electron density $\rho(\mathbf{r})$ is the variable of the universal functional, and the external potential $v(\mathbf{r})$ can be seen as a parameter of the Hamiltonian. In RDMFT, the 1RDM γ plays the role of $\rho(\mathbf{r})$, and the single-particle operator h plays the role of $v(\mathbf{r})$. There is, in fact, nothing special about the electron density or the 1RDM. The underlying mathematical structure of both functional theories depends solely on the fact that there is a pair of conjugate variables, in the sense of Legendre-Fenchel transformations, in each case: (ρ, v) in DFT and (γ, h) in RDMFT. In both functional theories, we aim to obtain the ground state energies of a family of Hamiltonians, each consisting of a fixed part, which we shall also call the "interaction" in the abstract language that follows, and a variable part. For example, in DFT the Hamiltonian has the form $H(v) = \sum_i v(\mathbf{r}_i) + W$, where $W = T + U$, and in RDMFT one has $H(h) = \sum_i h_i + U$.

Each state Γ on the Hilbert space, pure or mixed, is assigned a "reduced quantity", which is ρ or γ respectively in DFT and RDMFT, with which the expectation value of the variable part of the Hamiltonian can be computed. For example, $\mathrm{Tr}[(\sum_i v(\mathbf{r}_i))\,\Gamma] = \int v(\mathbf{r})\rho(\mathbf{r})\mathrm{d}\mathbf{r}$ in DFT and $\mathrm{Tr}[(\sum_i h_i)\Gamma] = \mathrm{Tr}(h\gamma)$ in RDMFT. The reduced quantity, ρ or γ, is then the simplest object obtained from the full description Γ of the quantum mechanical system with which it is still possible to recover the expectation value of the variable part of the Hamiltonian. Indeed, the point of view that the full electronic wave function $\Psi(\mathbf{r}_1, \cdots, \mathbf{r}_N)$ might contain too many complicated and unnecessary details for determining the physical properties of the ground state was a significant driving force of the development of both DFT and RDMFT [8, 52, 54], going back to Thomas and Fermi [55, 56].

Mathematically, then, we see that there is a dual relation, in the algebraic sense, between the reduced quantity and the variable part of the Hamiltonian. By considering different scopes within which the latter is allowed to vary, we obtain distinct functional theories. This idea has been explored and exploited extensively in the literature, expanding the decades-old framework of Hohenberg and Kohn to an ever more diverse range of systems, sometimes even finding applications and connections to other branches of quantum many-body physics. Here, we give a list of examples of such developments, apologizing in advance for not being exhaustive due to the myriad of possible modifications and extensions of DFT that have been

proposed in the literature.

(1) Already in 1972, von Barth and Hedin devised a spin-polarized variant of DFT [57], called spin density functional theory (SDFT), taking $\rho_{\sigma\sigma'}(\mathbf{r}) := \left\langle \psi^\dagger_\sigma(\mathbf{r})\psi_{\sigma'}(\mathbf{r}) \right\rangle$ as the reduced quantity, coupled to an external potential $v_{\sigma\sigma'}(\mathbf{r})$ now sensitive to the spin degree of freedom. Equivalently, the electron density $\rho(\mathbf{r})$ is augmented with the magnetization, with the external potential being of magnetic nature. The Hohenberg-Kohn theorem famously fails in the sense that a one-to-one correspondence between $\rho_{\sigma\sigma'}(\mathbf{r})$ and $v_{\sigma\sigma'}(\mathbf{r})$ cannot be established [57, 58].

(2) In 1988, Oliveira, Gross, and Kohn proposed a functional theory for superconductors [59], where the reduced variable is the pair (ρ, Δ), where $\rho(\mathbf{r}) := \left\langle \sum_\sigma \psi^\dagger_\sigma(\mathbf{r})\psi_\sigma(\mathbf{r}) \right\rangle$ is the usual electron density and $\Delta(\mathbf{r}, \mathbf{r}') := \langle \psi_\uparrow(\mathbf{r})\psi_\downarrow(\mathbf{r}') \rangle$ is the nonlocal gap function. The resulting functional then depends on both ρ and Δ, and a procedure analogous to the Kohn-Sham scheme can be carried out to compute the ground state properties. Subsequent works led by Lüders, Marques, and Lathiotakis [60, 61] incorporated the full many-body nuclear density as an extra density component.

(3) Higuchi and Higuchi [62] considered a modification of the constrained search, introducing an arbitrary local reduced quantity $X(\mathbf{r})$ associated to each quantum state, thus yielding a constrained search functional $\mathcal{F}(\rho, X) := \min_{\Psi \mapsto (\rho, X)} \langle \Psi | T + U | \Psi \rangle$. However, in this formulation, $X(\mathbf{r})$ is treated as a purely auxiliary quantity and does not appear in the Hamiltonian. That is, the family of Hamiltonians of interest is still $H(v) = T + U + \sum_i v(\mathbf{r}_i)$.

(4) Líbero and Capelle [63] proposed a functional theory for spin lattice systems, in particular for the inhomogeneous Heisenberg model, with Hamiltonian $H = J \sum_{\langle i,j \rangle} \mathbf{S}^{(i)} \cdot \mathbf{S}^{(j)} + \sum_i \mathbf{B}^{(i)} \cdot \mathbf{S}^{(i)}$. Here, the reduced quantity is the list of all single-site spin expectation values $\langle \mathbf{S}^{(i)} \rangle = (\langle S^{(i)}_x \rangle, \langle S^{(i)}_y \rangle, \langle S^{(i)}_z \rangle)$, which are coupled to a site-dependent magnetic field $\{\mathbf{B}^{(i)}\}$. Later, in a similar vein, Alcaraz and Capelle [64] described a functional theory for quantum chains with Hamiltonian of the form $H = H_0 + \sum_i v_i N_i$, where each N_i is the "spin" operator associated to site i and takes discrete values. In both works, a variant of the local density approximation is employed to estimate the correlation functional, which is by definition the difference between the exact functional and the one obtained from mean-field theory.

(5) Quantum phase transitions (QPTs) refer to nonanalyticity of properties of a system at zero temperature caused by tuning control parameters of the Hamiltonian [65, 66]. Wu et al. [67] established a general functional-theoretic framework to investigate QPTs with Hamiltonian $H = H_0 + \sum_l \lambda_l A_l$, where H_0 and the A_l are arbitrary operators and the reduced quantity is the collection of expectation values $\{\langle A_l \rangle\}$. The authors were then able to apply results from DFT, such as the HK theorem, to examine the dependence of an entanglement measure on the control parameters λ_l. See also Refs. [68–70] for later refinements.

(6) More recently, Penz and van Leeuwen [71] studied a system consisting of spinless fermions placed on a finite graph. There, the Hamiltonian is $H = h + U$, where U is a fixed operator and h is a single-particle operator containing a vertex-dependent external potential

together with hopping terms specified by the edges of the graph. The reduced quantity is simply the discrete analog of the electron density, i.e., the electron number on each vertex, and the external potential is allowed to vary. The pure state N-representability problem is solved without complication, and the mechanism behind potential violation of the HK theorem in the same sense as in (1) is discussed in detail. Note that an intermediate case between this setting and ordinary DFT, where electrons move in a continuum, is discussed by Chayes et al. [72], in which infinite but discrete lattice systems are considered.

(7) Within RDMFT, Liebert et al. [73] demonstrated the construction of new functional theories by further reducing the 1RDM γ in case the single-particle Hamiltonian h is allowed to vary within a restricted subspace. In this context, the reduced quantity corresponds to the orthogonal projection of γ onto said subspace.

In each variant or modification of the original DFT listed above, often nontrivial portions of the work are dedicated to presenting an adapted proof of (some version of) the HK theorem. It is almost a tradition at this point in the field that whenever a novel variant of functional theory is proposed, the main mathematical results are to be re-proved so as to reassure the reader that replacing the electron density $\rho(\mathbf{r})$ by some other reduced quantity does not break the mathematical foundations that underlie a functional theory.

The goal of this chapter is thus to extract the abstract structure constitutive to (1)–(7), thereby establishing a very general mathematical framework in which notions like "universal functionals", "v-representability", and "N-representability" can be defined. The advantage of doing so is that each time we prove a mathematical statement, it will be simultaneously true for all functional theories as long as a certain set of minimal assumptions are satisfied.

Before going into formal mathematics, three remarks are in order: (a) No attempt will be made to broaden the mathematical theory to cover excited states or finite temperatures; we focus only on functional theories for the ground state problem. This is not because generalizations are impossible, but because we will anyway be concerned with only the ground state and its properties in the subsequent chapters. (b) For technical reasons, we always work with finite-dimensional vector spaces. This means that terms like "Hilbert space" or "vector space" without modifiers are understood to be finite-dimensional. The generalization to infinite dimensions should be at least intuitively clear. (c) As already hinted at in Examples (1)–(7) above, the HK theorem in its strongest form, shown in the original paper [8], does not hold in general. This is because the proof relies on the hypothesis that the ground state wave function does not vanish on a set of positive measure (for a mathematically rigorous discussion, see Ref. [52]), a condition that is prohibitively strong when the Hilbert space is finite-dimensional. However, as will become clear in Section 2.3, there is a weaker version of the HK theorem that is good enough for defining a universal functional. The failure of the strong form of the HK theorem is intimately tied to the existence of ground state degeneracies, and a satisfactory investigation of this phenomenon can be found in Refs. [74–76]. In the following, we will not have much to say about this topic.

2.2 Definitions

Recall that if $\mathcal{H}$ is a Hilbert space, the Lie algebra $\mathfrak{u}(\mathcal{H})$ of the unitary group consists of all skew-Hermitian endomorphisms (operators). Hence, we will write $i\mathfrak{u}(\mathcal{H})$ for the real vector

space of Hermitian operators.

Definition 2.1. *A **generalized functional theory** is a tuple $(V, \mathcal{H}, \iota, W)$, where V is a real vector space, $\mathcal{H}$ a Hilbert space, $\iota : V \to i\mathfrak{u}(\mathcal{H})$ a linear map and $W \in i\mathfrak{u}(\mathcal{H})$ is any Hermitian operator. The map ι is called the **potential map**, and W is called the **interaction** or the **fixed part** of the generalized functional theory. An element of V is called an **external potential**.*

We are interested in determining the ground state energy of Hamiltonians of the form $H(v) = \iota(v) + W$, where $v \in V$ is any external potential. Thus, we define the *(ground state) energy function* to be

$$E(v) := \min_{\substack{|\Psi\rangle \in \mathcal{H}: \\ \|\Psi\|=1}} \langle \Psi | \iota(v) + W | \Psi \rangle \,. \tag{2.1}$$

For any $v \in V$, let $\mathbf{G}_p(v)$ denote the set of all pure ground states $|\Psi\rangle\langle\Psi|$ of $\iota(v) + W$ and let $\mathbf{G}_e(v)$ denote the set of all ensemble ground states. Clearly, (1) $\mathbf{G}_e(v)$ is the convex hull of $\mathbf{G}_p(v)$, (2) $\mathbf{G}_p(v) = \mathbf{G}_e(v) \cap \mathcal{P}$, and (3) $E(v) = \mathrm{Tr}(\Gamma H(v)) = \mathrm{Tr}(\Gamma(\iota(v) + W))$ for any $\Gamma \in \mathbf{G}_e(v)$, in particular for any $\Gamma \in \mathbf{G}_p(v)$.

Proposition 2.1. *The energy function $E : V \to \mathbb{R}$ is concave.*

Proof. Take $v_1, v_2 \in V$ and $t \in [0, 1]$. Then

$$\begin{aligned}
&E(tv_1 + (1-t)v_2) = \min_{\Gamma \in \mathcal{P}} \mathrm{Tr}[\Gamma(t\iota(v_1) + (1-t)\iota(v_2) + W)] \\
&= \min_{\Gamma \in \mathcal{P}} \Big[t\,\mathrm{Tr}[\Gamma(\iota(v_1) + W)] + (1-t)\,\mathrm{Tr}[\Gamma(\iota(v_2) + W)] \Big] \\
&\geq t \min_{\Gamma \in \mathcal{P}} \Big[\mathrm{Tr}[\Gamma(\iota(v_1) + W)] \Big] + (1-t) \min_{\Gamma \in \mathcal{P}} \Big[\mathrm{Tr}[\Gamma(\iota(v_2) + W)] \Big] \\
&= tE(v_1) + (1-t)E(v_2).
\end{aligned}$$

□

Example 2.1. Consider an N-electron system on discrete lattice sites in any spatial dimension labeled by $i = 1, 2, \cdots, d$. The single-particle Hilbert space is

$$\mathcal{H}_1 = \mathbb{C}^d = \mathrm{span}_{\mathbb{C}}\{|1\rangle, |2\rangle, \cdots, |d\rangle\},$$

where $|i\rangle$ is the single-particle state where an electron is located at site i. The N-particle Hilbert space is $\mathcal{H} = \bigwedge^N \mathcal{H}_1$. Take $V = \mathbb{R}^d$ and $\iota : V \to i\mathfrak{u}(\mathcal{H}), v \mapsto \sum_{i=1}^d v_i a_i^\dagger a_i$. Let T be a suitable discretized kinetic energy operator, and U the discretized Coulomb interaction, and take $W = T + U$. Then the generalized functional theory defined by $(V, \mathcal{H}, \iota, W)$ is the ordinary DFT (on a finite lattice).

Example 2.2. Consider a system of N indistinguishable particles, either bosons or fermions. The N-particle Hilbert space is $\mathcal{H} = \bigwedge^N \mathcal{H}_1$ for fermions and $\mathcal{H}_N = \mathrm{Sym}^N \mathcal{H}_1$ for bosons. Take $V = i\mathfrak{u}(\mathcal{H}_1)$ and $\iota : V \to i\mathfrak{u}(\mathcal{H})$ to be the map that sends a single-particle operator to the corresponding operator on $\mathcal{H}$, i.e.,

$$\iota\left(u\,|i\rangle\langle j| + \bar{u}\,|j\rangle\langle i|\right) = u a_i^\dagger a_j + \bar{u} a_j^\dagger a_i,$$

where $(|i\rangle)_{i=1}^{\dim \mathcal{H}_1}$ is any orthonormal basis for $\mathcal{H}_1$ and $a_i^\dagger, a_i$ are the respective creation and annihilation operators. Take W to be an arbitrary interaction, e.g. $W = \sum_{ijkl} w_{ijkl} a_i^\dagger a_j^\dagger a_k a_l$

for a two-particle interaction. The generalized functional theory defined by $(V, \mathcal{H}, \iota, W)$ is RDMFT.

Example 2.3. Consider a chain of N two-level spins, whose Hilbert space is $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$. Take $V = (\mathbb{R}^3)^N$ and

$$\iota : (\mathbb{R}^3)^N \to i\mathfrak{u}(\mathcal{H})$$
$$(v^1, \cdots, v^N) \mapsto \sum_{i=1}^{N} (v_x^i X_i + v_y^i Y_i + v_z^i Z_i),$$

where X_i, Y_i, Z_i are the Pauli operators of the i-th spin. We can take the interaction to be, for example

$$W = -\sum_{i}^{N} (X_i X_{i+1} + Y_i Y_{i+1} + Z_i Z_{i+1}).$$

Then $(V, \mathcal{H}, \iota, W)$ recovers the spin-$\frac{1}{2}$ case of the spin functional theory discussed by Líbero and Capelle [63]. Generalization to arbitrary spin is straightforward.

Remark 2.1. A note on terminology: calling W the "interaction" comes from RDMFT, where W contains only the particle-particle interaction term and *not* the kinetic term. This piece of terminology, however, need not make any physical sense in other systems. For example, in DFT, we want W to be the sum of both the electron-electron interaction *and* the kinetic energy. Hence, the word "interaction" should be regarded merely as a reminder of how RDMFT is a special case of Definition 2.1. Whenever there is an opportunity for confusion, we shall call W the "fixed part" of the functional theory. Similarly, the term "external potential", which now refers to any $v \in V$ in our abstract setting, comes from the common nomenclature used in DFT. The corresponding term in RDMFT would be "single-particle operator", which has a very specific definition, and consequently is not chosen for the general terminology. Sometimes, by abuse of language, we call $\iota(v)$, instead of v, an external potential.

Remark 2.2. In many cases V will be a subspace of $i\mathfrak{u}(\mathcal{H})$ and $\iota : V \hookrightarrow i\mathfrak{u}(\mathcal{H})$ is the inclusion. However, we allow the possibility that ι is not injective.

In the following, we fix a generalized functional theory $(V, \mathcal{H}, \iota, W)$. Recall that an *ensemble state* (or *density operator*) on $\mathcal{H}$ is a positive semidefinite operator $\Gamma \in i\mathfrak{u}(\mathcal{H})$ such that $\mathrm{Tr}(\Gamma) = 1$. The set of all ensemble states forms a convex compact subset of $i\mathfrak{u}(\mathcal{H})$, whose extreme points are exactly the pure states, i.e., rank-one projectors. We will use $\mathcal{E}$ and $\mathcal{P}$ to denote the sets of all ensemble states and pure states on $\mathcal{H}$ respectively.

It turns out to be more convenient to think of ensemble states not as operators on $\mathcal{H}$, but rather as linear functionals on $i\mathfrak{u}(\mathcal{H})$. This means that we will think of an ensemble state Γ as belonging to $(i\mathfrak{u}(\mathcal{H}))^*$ via the canonical isomorphism $i\mathfrak{u}(\mathcal{H}) \to (i\mathfrak{u}(\mathcal{H}))^*, \Gamma \mapsto \mathrm{Tr}(\Gamma(\cdot))$.

Take any ensemble state $\Gamma \in (i\mathfrak{u}(\mathcal{H}))^*$ and any external potential $v \in V$. The expectation value of the operator $\iota(v) \in i\mathfrak{u}(\mathcal{H})$ with respect to Γ is $\langle \Gamma, \iota(v) \rangle$, where $\langle \cdot, \cdot \rangle : (i\mathfrak{u}(\mathcal{H}))^* \times i\mathfrak{u}(\mathcal{H}) \to \mathbb{R}$ is the natural pairing. Then, we have by duality

$$\langle \Gamma, \iota(v) \rangle = \langle \iota^*(\Gamma), v \rangle . \tag{2.2}$$

In this equation, the brackets $\langle \cdot, \cdot \rangle$ on the right hand side refer to the pairing $V^* \times V \to \mathbb{R}$. For a fixed ensemble state Γ, Eq. (2.2) holds for any external potential $v \in V$. In other

words, $\iota^*(\Gamma) \in V^*$ contains all the information needed to compute the expectation value of any $v \in V$. Hence, the dual vector $\iota^*(\Gamma) \in V^*$ corresponds to exactly the electron density $\rho(\mathbf{r})$ in DFT, which can be used to compute the expectation value of any external potential $v(\mathbf{r})$, or the 1RDM γ in RDMFT, which is sufficient to compute the expectation value of any single-particle operator h. This motivates the following:

Definition 2.2. *The dual map $\iota^* : (i\mathfrak{u}(\mathcal{H}))^* \to V^*$ is the **density map**. A **density** is a dual vector $\rho \in V^*$.*

Thus, for example, $\iota^*(\mathbf{G}_p(v))$ is the set of all densities in V^* that come from a pure ground state of $\iota(v) + W$ for each external potential $v \in V$.

We discuss now the notion of *representability* of densities. Roughly speaking, this has to do with whether a density $\rho \in V^*$ comes from a quantum state Γ. In RDMFT, for example, one speaks of the *N-representability problem*, which asks whether a given single-particle density operator γ is the 1RDM of some pure or ensemble N-particle state. Depending on whether we allow Γ to be an ensemble state, and whether Γ has to be a ground state of $\iota(v) + W$ for some external potential $v \in V$, there are in total four types of representability that can be defined. The importance of this notion will become clear when we define the various universal functionals in Section 2.4: we will see that the domain of each of the universal functionals is the set of those densities $\rho \in V^*$ that are representable in one of the following four senses.

Definition 2.3. *A density $\rho \in V^*$ is:*

- ***pure (resp. ensemble) state representable** if there exists a pure (resp. ensemble) state $\Gamma \in (i\mathfrak{u}(\mathcal{H}))^*$ such that $\iota^*(\Gamma) = \rho$.*
- ***pure (resp. ensemble) state v-representable** if there exists an external potential $v \in V$ and a pure (resp. ensemble) state $\Gamma \in (i\mathfrak{u}(\mathcal{H}))^*$ such that Γ is a ground state of $\iota(v) + W$ and $\iota^*(\Gamma) = \rho$.*

That is, the set of pure (resp. ensemble) state representable densities is $\iota^*(\mathcal{P})$ (resp. $\iota^*(\mathcal{E})$), and the set of pure (resp. ensemble) state v-representable densities is $\bigcup_{v\in V} \iota^*(\mathbf{G}_p(v))$ (resp. $\bigcup_{v\in V} \iota^*(\mathbf{G}_e(v))$). We will refer to the problem of determining and characterizing the sets $\iota^*(\mathcal{P})$ and $\iota^*(\mathcal{E})$ as the *representability problem*.

Applied to DFT or RDMFT, the notion of pure/ensemble state representability coincides exactly with that of pure/ensemble state N-representability. Since the present framework (Definition 2.1) generalizes beyond N-particle systems, we drop the prefix "N-" in the abstract terminology.

Example 2.4. Consider a system of N electrons. Let V consist of all two-particle operators, i.e., those of the form $a_i^\dagger a_j^\dagger a_k a_l$, and let $\iota : V \hookrightarrow i\mathfrak{u}(\mathcal{H}_N)$ be the inclusion. The density map ι^* is, up to normalization, the partial trace Tr_{N-2} over all but two particles, and $\iota^*(\Gamma)$ is proportional to the 2RDM. The problem of determining whether a two-particle operator ρ is in the set $\iota^*(\mathcal{E})$ is QMA complete [77].

The preceding example shows that determining the sets $\iota^*(\mathcal{E})$ and $\iota^*(\mathcal{P})$ can be extremely difficult in general. Indeed, if the Hamiltonian H of a system contains at most two-particle interactions, then the energy expectation value $\langle\Psi|H|\Psi\rangle$ depends linearly on the 2RDM. Consequently, an efficient characterization of the set of (ensemble state) representable 2RDMs would

lead to an efficient algorithm, which works *independently of the Hamiltonian*, for obtaining the exact ground state energy and ground state 2RDM using standard convex optimization techniques. This serves to warn us against being overly ambitious about the representability problem, and motivates us to single out particular classes of generalized functional theories for which the representability problem is comparatively more tractable.

With the notion of representability defined, we can now address a technical subtlety that we left out in Definition 2.1: when defining the potential map $\iota : V \to i\mathfrak{u}(\mathcal{H})$, we could have chosen to exclude the possibility that a nonzero element $v \in V$ is sent to a multiple of $\mathbb{1} \in i\mathfrak{u}(\mathcal{H})$. Evidently, this does not affect the physics, since adding a multiple of the identity to the Hamiltonian merely shifts the energy levels (in particular the ground state energy $E(v)$) by a constant. However, due to our definition of densities as elements of V^*, the existence of $v \neq 0$ such that $\iota(v) \propto \mathbb{1}$ will reduce the dimension of the set of representable densities in V^*. In RDMFT, for example, this reduction of dimension corresponds to the normalization of the trace of the 1RDM: $\mathrm{Tr}(\gamma) = N$.

Recall that an *affine combination* of finitely many vectors is a linear combination with coefficients summing to one. The *affine hull* of a subset X of a vector space, denoted as $\mathrm{aff}(X)$, is the set of all affine combinations of vectors in X. If A is an affine set, we write $\vec{A}$ to denote the vector space of translations, i.e., $\vec{A} := \{a - b \mid a, b \in A\}$. Also recall that the *annihilator* v° of a vector v is the vector space of all dual vectors that vanish on v.

Lemma 2.2. $\overrightarrow{\mathrm{aff}(\iota^*(\mathcal{P}))} = \overrightarrow{\mathrm{aff}(\iota^*(\mathcal{E}))} \cong \left(\frac{V}{\iota^{-1}(\mathrm{span}\{\mathbb{1}\})}\right)^*$ *canonically.*

Proof. The first equality is obvious.

$\overrightarrow{\mathrm{aff}(\iota^*(\mathcal{E}))} = \overrightarrow{\iota^*(\mathrm{aff}(\mathcal{E}))} = \iota^*\left(\overrightarrow{\mathrm{aff}(\mathcal{E})}\right)$ by linearity. But $\overrightarrow{\mathrm{aff}(\mathcal{E})} = \mathbb{1}^\circ \subset (i\mathfrak{u})^*$. Indeed, "$\subset$" is obvious due to normalization of density operators, and for each traceless Hermitian operator B, we see from the spectral decomposition $B = \sum_n \lambda_n |n\rangle\langle n|$, $\sum_n \lambda_n = 0$, that $B \in \overrightarrow{\mathrm{aff}(\mathcal{E})}$ (here we identify $(i\mathfrak{u}(\mathcal{H}))^* \cong i\mathfrak{u}(\mathcal{H})$ by the Hilbert-Schmidt inner product), showing "$\supset$". Thus, proving the lemma is equivalent to showing

$$\iota^*(\mathbb{1}^\circ) \cong \left(\frac{V}{\iota^{-1}(\mathrm{span}\{\mathbb{1}\})}\right)^*,$$

which is true because

$$\left(\frac{V}{\iota^{-1}(\mathrm{span}\{\mathbb{1}\})}\right)^* \cong (\iota^{-1}(\mathrm{span}\{\mathbb{1}\}))^\circ = \iota^*(\mathbb{1}^\circ)$$

canonically. □

Proposition 2.3. *The dimension of the set $\iota^*(\mathcal{E})$ of ensemble state representable densities is equal to $\dim V - \dim \iota^{-1}(\mathrm{span}\{\mathbb{1}\})$. In particular, $\iota^*(\mathcal{E})$ has full dimension if and only if the potential map ι is injective and $\mathbb{1} \in i\mathfrak{u}(\mathcal{H})$ is not contained in the image of ι. To be more precise,*

$$\mathrm{aff}(\iota^*(\mathcal{E})) = V^* \Leftrightarrow \iota^{-1}(\mathrm{span}\{\mathbb{1}\}) = 0.$$

Proof. This is a direct consequence of Lemma 2.2. □

Example 2.5. Consider the functional theory defined by $(i\mathfrak{u}(2), \mathbb{C}^2, \mathrm{Id}_{i\mathfrak{u}(2)}, W)$, where the interaction W is irrelevant. We have $\dim \mathrm{Id}^{-1}(\mathrm{span}\{\mathbb{1}\}) = \dim \mathrm{span}\{\mathbb{1}\} = 1$, so Proposition 2.3 states that $\mathrm{Id}^*(\mathcal{E})$ has dimension $\dim(i\mathfrak{u}(2)) - 1 = 3$, which is expected since $\mathcal{E}$ is just the Bloch ball.

It will be convenient to have a formula for the derivative of $\iota^*|_{\mathcal{P}}$. For any pure state $|\Psi\rangle\langle\Psi| \in \mathcal{P}$, we may identify the tangent space $T_{|\Psi\rangle\langle\Psi|}\mathcal{P}$ with $|\Psi\rangle^\perp = \{|\phi\rangle \in \mathcal{H} \mid \langle\phi|\Psi\rangle = 0\}$ (to get the actual identification, note that the derivative of the map $\mathcal{H} \to \mathcal{P}, |\Psi'\rangle \mapsto |\Psi'\rangle\langle\Psi'|$ at $|\Psi\rangle$ has kernel $\mathbb{C}\,|\Psi\rangle$, thus $T_{|\Psi\rangle\langle\Psi|}\mathcal{P} \cong \mathcal{H}/\mathbb{C}\,|\Psi\rangle \cong |\Psi\rangle^\perp$ canonically).

Lemma 2.4. *The derivative of the density map ι^* along the set of pure states $\mathcal{P}$ is characterized by*

$$\left\langle D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}(|\phi\rangle), v\right\rangle = 2\mathrm{Re}\,\langle\phi|\iota(v)|\Psi\rangle \tag{2.3}$$

for any $|\Psi\rangle \in \mathcal{H}$, $\|\Psi\|^2 = 1$, $|\phi\rangle \in T_{|\Psi\rangle\langle\Psi|}\mathcal{P}$, and $v \in V$.

Proof.

$$\begin{aligned}
&\left\langle D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}(|\phi\rangle), v\right\rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle \iota^*((\langle\Psi| + t\,\langle\phi|)(|\Psi\rangle + t\,|\phi\rangle)), v\rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\langle\Psi| + t\,\langle\phi|)\iota(v)(|\Psi\rangle + t\,|\phi\rangle) \\
&= \langle\phi|\iota(v)|\Psi\rangle + \langle\Psi|\iota(v)|\phi\rangle\,.
\end{aligned}$$

□

2.3 The Hohenberg-Kohn Theorem

The conceptual basis of DFT is the so-called Hohenberg-Kohn theorem [8], which establishes a one-to-one correspondence between external potentials and densities in the case of the inhomogeneous electron gas (that is, a system of N electrons moving in $\mathbb{R}^3$ interacting with each other via the Coulomb interaction and subject to an external potential $v(\mathbf{r})$), which in turn allows us to define the *Hohenberg-Kohn functional*. Actually proving this statement in full mathematical rigor is very subtle, and it does not apply to finite-dimensional systems (see discussions below). Nevertheless, for the definition of the functional, it turns out to be sufficient to prove a weaker version of the Hohenberg-Kohn theorem.

Theorem 2.5 (weak Hohenberg-Kohn; see, for example, Theorem 7 of Ref. [71])**.** *For any $v, v' \in V$, if $\rho \in \iota^*(\mathbf{G}_p(v)) \cap \iota^*(\mathbf{G}_p(v'))$, then $(\iota^*)^{-1}(\rho) \cap \mathbf{G}_p(v) = (\iota^*)^{-1}(\rho) \cap \mathbf{G}_p(v')$. In other words, if $\iota(v) + W$ and $\iota(v') + W$ both share a ground state density ρ, then any pure ground state Γ $(= |\Psi\rangle\langle\Psi|)$ of $\iota(v) + W$ whose density is ρ is also a ground state of $\iota(v') + W$ and vice versa. The same statement holds if we replace "p" with "e" (and "pure" with "ensemble").*

Proof. Assume $\rho \in \iota^*(\mathbf{G}_p(v)) \cap \iota^*(\mathbf{G}_p(v'))$. Take $\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathbf{G}_p(v)$ and $\Gamma' \in (\iota^*)^{-1}(\rho) \cap \mathbf{G}_p(v')$. In other words, we take a pure ground state Γ of $\iota(v) + W$ and a pure ground state Γ' of $\iota(v') + W$ such that $\iota^*(\Gamma) = \iota^*(\Gamma') = \rho$. Then

$$\begin{aligned}
&\langle\Gamma, \iota(v) + W\rangle \le \langle\Gamma', \iota(v) + W\rangle \\
&= \langle\Gamma', \iota(v)\rangle + \langle\Gamma', W\rangle = \langle\iota^*(\Gamma'), v\rangle + \langle\Gamma', W\rangle \\
&= \langle\rho, v\rangle + \langle\Gamma', W\rangle \\
&= \langle\rho, v - v'\rangle + \langle\rho, v'\rangle + \langle\Gamma', W\rangle \\
&= \langle\rho, v - v'\rangle + \langle\iota^*(\Gamma'), v'\rangle + \langle\Gamma', W\rangle \\
&= \langle\rho, v - v'\rangle + \langle\Gamma', \iota(v') + W\rangle\,,
\end{aligned} \tag{2.4}$$

implying $\langle \Gamma, \iota(v) + W\rangle + \langle \rho, v' - v\rangle \leq \langle \Gamma', \iota(v') + W\rangle$. Therefore,

$$\langle \Gamma', \iota(v') + W\rangle \leq \langle \Gamma, \iota(v') + W\rangle = \langle \Gamma, \iota(v) + W\rangle + \langle \rho, v' - v\rangle \leq \langle \Gamma', \iota(v') + W\rangle\,. \tag{2.5}$$

Both ends of this chain of inequalities agree, so the inequalities must have been equalities. In particular, $\langle \Gamma', \iota(v') + W\rangle = \langle \Gamma, \iota(v') + W\rangle$, showing that Γ is also a ground state of $\iota(v') + W$. Similarly, Γ' is also a ground state of $\iota(v) + W$.

It is easy to check that everything carries through if we replace "p" with"e" and "pure" with "ensemble". □

In contrast to this weakened version, the original Hohenberg-Kohn theorem claims that from the ground state density ρ, one can already recover the external potential v up to a constant. More precisely:

Theorem 2.6 (strong Hohenberg-Kohn, **not true in general!**)**.** *If two external potentials $v, v' \in V$ are such that $\iota(v) + W$ and $\iota(v') + W$ have respective pure ground states Γ, Γ' that share the same density, i.e., $\iota^*(\Gamma) = \iota^*(\Gamma')$, then $\iota(v) - \iota(v')$ is proportional to the identity.*

At this point, it is convenient to introduce the notion of *unique v-representability*:

Definition 2.4. *A density $\rho \in V^*$ is **uniquely (pure state) v-representable** if there exists an external potential $v \in V$ such that $\rho \in \iota^*(\mathbf{G}_p(v))$, and any other external potential $v' \in V$ for which $\rho \in \iota^*(\mathbf{G}_p(v'))$ satisfies $\iota(v') - \iota(v) \propto \mathbb{1}$.*

The strong Hohenberg-Kohn theorem then claims that all pure state v representable densities are uniquely v-representable.

This stronger version holds, with some mathematical subtleties that were settled rather recently [52, 78, 79], for N-electron systems in $\mathbb{R}^3$, which was the context in Hohenberg and Kohn's seminal paper. In finite systems, counterexamples are found and discussed in Refs. [71, 75, 76], where a functional theory for fermions on finite graphs are studied. Here, we will construct an extremely simple example that violates Theorem 2.6.

Example 2.6 (counterexample to the strong Hohenberg-Kohn theorem)**.** Take $V = \mathbb{R}$, $\mathcal{H} = \mathbb{C}^2 = \mathrm{span}\{|\uparrow\rangle, |\downarrow\rangle\}$, $\iota(v) = vZ$, where $Z = \mathrm{diag}(1, -1)$ is the Pauli Z operator, and $W = 0$. Physically, we might want to think of a spin-$\frac{1}{2}$ particle fixed at the origin and subject to an external tunable magnetic field in the z direction. We claim that in the generalized functional theory given by $(V, \mathcal{H}, \iota, W)$, the density $1 \in V^* = \mathbb{R}$ violates the strong HK theorem. Take distinct external potentials $v, v' < 0$. Then both $\iota(v) + W = vZ$ and $\iota(v') + W = v'Z$ share the same ground state $|\uparrow\rangle$, hence the same density, which can be readily computed:

$$\rho = \iota^*(|\uparrow\rangle\langle\uparrow|) = (\langle\uparrow| \cdot |\uparrow\rangle) \circ (v \mapsto vZ) = (v \mapsto v\,\langle\uparrow|Z|\uparrow\rangle) = (v \mapsto v) = 1.$$

Note that this is the minimal example where the strong Hohenberg-Kohn theorem can be violated: if $\mathcal{H}$ is one-dimensional, then $\iota(V) \subset \mathbb{1} \cdot \mathbb{R}$, so the conclusion of the theorem trivially holds.

Even though the strong Hohenberg-Kohn theorem does not hold in general, something useful can still be said about unique v-representability. Recall that if $f : M \to N$ is a smooth map of smooth manifolds, a point $q \in N$ is called a *regular value* if the derivative $D_p f : T_p M \to T_q N$ is surjective for all $p \in f^{-1}(q)$. Otherwise, $q \in N$ is said to be a *critical value* (if q lies outside of the image of f, it is a regular value vacuously).

Proposition 2.7. *Let $\rho \in V^*$ be a pure state v-representable density. If ρ is a regular value of the map $\iota^*|_{\mathcal{P}} : \mathcal{P} \to \mathrm{aff}(\iota^*(\mathcal{P}))$, then ρ is uniquely v-representable.*

Proof. Suppose $v, v' \in V$ are such that ρ is a (pure) ground state density for both $\iota(v) + W$ and $\iota(v') + W$. Then by the weak Hohenberg-Kohn theorem (Theorem 2.5), there is a state $|\Psi\rangle$ such that $\iota^*(|\Psi\rangle\langle\Psi|) = \rho$ and

$$(\iota(v) + W)\,|\Psi\rangle = E\,|\Psi\rangle$$
$$(\iota(v') + W)\,|\Psi\rangle = E'\,|\Psi\rangle$$

for some E, E'. Taking the difference, we get $\iota(v - v')\,|\Psi\rangle = (E - E')\,|\Psi\rangle$.

Suppose $\iota(v - v') \not\propto \mathbb{1}$, then there is some $\gamma \in \mathbb{1}^\circ$ such that $\langle \iota^*(\gamma), v - v'\rangle = \langle \gamma, \iota(v - v')\rangle \neq 0$. But $\iota^*(\mathbb{1}^\circ) = \overrightarrow{\mathrm{aff}(\iota^*(\mathcal{P}))}$ (see the proof of Lemma 2.2), so $\alpha := \iota^*(\gamma)$ is a vector in $\overrightarrow{\mathrm{aff}(\iota^*(\mathcal{P}))}$ such that $\langle \alpha, v - v'\rangle \neq 0$.

Since the map $D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}} : |\Psi\rangle^\perp \to \overrightarrow{\mathrm{aff}(\iota^*(\mathcal{P}))}$, $|\phi\rangle \mapsto (v \mapsto 2\mathrm{Re}\,\langle\phi|\iota(v)|\Psi\rangle)$ is surjective (where we used Lemma 2.4), we can find a $|\phi\rangle \in |\Psi\rangle^\perp$ such that $(v \mapsto 2\mathrm{Re}(\langle\phi|\iota(v)|\Psi\rangle) = \alpha$. But then

$$2\mathrm{Re}\,\langle\phi|\iota(v - v')|\Psi\rangle = \langle \alpha, v - v'\rangle\,.$$

The right hand side is nonzero, while the left hand side vanishes since $\langle\phi|\iota(v - v')|\Psi\rangle = (E - E')\,\langle\phi|\Psi\rangle = 0$. Hence, we have a contradiction and it must have been that $\iota(v - v') \propto \mathbb{1}$. □

Remark 2.3. In the statement of Proposition 2.7, we have to restrict the codomain of $\iota^*|_{\mathcal{P}}$ to $\mathrm{aff}(\iota^*(\mathcal{P}))$, which is fine because we know the image of $\iota^*|_{\mathcal{P}}$ is contained in the latter set anyway. We do this to prevent the condition that a density be a regular value from being too restrictive: indeed, if $\mathrm{aff}(\iota^*(\mathcal{P}))$ has positive codimension in V^*, then every density will be a critical value of the map $\iota^*|_{\mathcal{P}} : \mathcal{P} \to V^*$! Of course, in this case, the conclusion of Proposition 2.7 would be still be true, albeit vacuously, but then we would obtain a much weaker statement. In light of Proposition 2.3, we see that this sort of inconvenience is the price we pay for allowing the potential map ι to map nonzero elements $v \in V$ to multiples of $\mathbb{1}$.

Applying Sard's theorem, which states that the set of critical values has measure zero for any smooth map, we get:

Corollary 2.8. *In any generalized functional theory, the set of pure state v-representable densities which are not uniquely v-representable, i.e., densities which violate the strong Hohenberg-Kohn theorem, has measure zero in* $\mathrm{aff}(\iota^*(\mathcal{P}))$.

Thus, we conclude that the strong HK theorem is not violated *too drastically*.

The study of the properties of set of (uniquely) pure state v-representable densities has a history in the literature. For more detail, see Refs. [58, 71, 74, 75, 80–82].

2.4 Universal Functionals

In this section, we will define a total of three functionals within a generalized density functional theory $(V, \mathcal{H}, \iota, W)$. In practice, that is, in physically interesting and large systems, none of the

functionals can be exactly obtained due to either the non-constructive nature of the definition (for the Hohenberg-Kohn functional) or the sheer complexity of the optimization problems involved (for the constrained search functionals). However, as we will see, each of these three functionals plays a significant theoretical role in understanding functional theories in general, and they are related in precise ways discussed in the second half of this section.

Definition 2.5. *The **Hohenberg-Kohn functional** $\mathcal{F}_{HK}$ is the real-valued function defined on the set of pure state v-representable densities in the following way: given a v-representable density $\rho \in V^*$, take an external potential $v \in V$ so that $\rho \in \iota^*(\mathbf{G}_p(v))$. Then define $\mathcal{F}_{HK}(\rho) = \langle \Gamma, W \rangle$ for any $\Gamma \in \mathbf{G}_p(v) \cap (\iota^*)^{-1}(\rho)$.*

The definition of the value of the Hohenberg-Kohn functional at ρ depends on two choices: first, one has to choose an external potential $v \in V$ such that ρ is a ground state density of $\iota(v) + W$, then one specific pure ground state $\Gamma \in \mathbf{G}_p(v)$ such that $\iota^*(\Gamma) = \rho$ is chosen. It is precisely the weak Hohenberg-Kohn theorem (Theorem 2.5) that guarantees that $\mathcal{F}_{HK}(\rho)$ does not depend on these two choices.

Proposition 2.9. *The Hohenberg-Kohn functional is well-defined.*

Proof. We need to show that $\mathcal{F}_{HK}(\rho)$ does not depend on the choice of v and Γ. Suppose $v' \in V$ also satisfies $\rho \in \iota^*(\mathbf{G}_p(v'))$. Take any $\Gamma' \in \mathbf{G}_p(v') \cap (\iota^*)^{-1}(\rho)$. By Theorem 2.5, we have $\Gamma' \in \mathbf{G}_p(v) \cap (\iota^*)^{-1}(\rho)$. Therefore

$$\langle \Gamma', \iota(v) + W \rangle = \langle \Gamma, \iota(v) + W \rangle \Rightarrow \langle \Gamma', W \rangle = \langle \Gamma, W \rangle \,. \tag{2.6}$$

□

Definition 2.6. *The **pure functional** $\mathcal{F}_p$ and the **ensemble functional** $\mathcal{F}_e$ are defined as*

$$\begin{aligned} \mathcal{F}_p(\rho) : \iota^*(\mathcal{P}) &\to \mathbb{R} \\ \rho &\mapsto \min_{\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathcal{P}} \langle \Gamma, W \rangle \end{aligned} \tag{2.7}$$

$$\begin{aligned} \mathcal{F}_e(\rho) : \iota^*(\mathcal{E}) &\to \mathbb{R} \\ \rho &\mapsto \min_{\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathcal{E}} \langle \Gamma, W \rangle \,. \end{aligned} \tag{2.8}$$

That is, $\mathcal{F}_p(\rho)$ (resp. $\mathcal{F}_e(\rho)$) is defined to be the minimum of the expectation value $\langle \Gamma, W \rangle$ over all pure (resp. ensemble) states Γ such that $\iota^*(\Gamma) = \rho$.

Remark 2.4. The density map ι^* is linear, in particular continuous, so both $(\iota^*)^{-1}(\rho) \cap \mathcal{P}$ and $(\iota^*)^{-1}(\rho) \cap \mathcal{E}$ are compact. Since the function $\Gamma \mapsto \langle \Gamma, W \rangle$ is continuous, the infima in Eqs. (2.7) and (2.8) are attained. This justifies writing min instead of inf.

Remark 2.5. In the literature, these are also often called *constrained search* functionals because of the constrained minimizations appearing in their definitions. In the DFT literature, the pure functional is sometimes also called the *Levy-Lieb functional*, and the ensemble functional is sometimes called the *Lieb functional*. In the RDMFT literature, the Hohenberg-Kohn functional is called the *Gilbert functional*, due to Gilbert [17], and the ensemble functional is often called the *Valone functional*, due to Valone [53]. Throughout this thesis, we will stick to the more descriptive names given in Definition 2.6. We will also refer to both $\mathcal{F}_p$ and $\mathcal{F}_e$ as the constrained search functionals, and to the minimizations (2.7) and (2.8) as the *constrained search*.

Remark 2.6. A brief note on the word "functional": in DFT, a suitable function space which is an infinite-dimensional Banach space (see, for example, Section 2 of Ref. [47]) plays the role of V, so $\mathcal{F}_{HK}, \mathcal{F}_p, \mathcal{F}_e$ are functionals in the usual sense. In our setting, V is finite-dimensional, so it may be more natural to call them "functions". However, we choose to adopt the traditional terminology.

Example 2.7. We construct a simple functional theory and compute its functionals. Let $V = \mathbb{R}, \mathcal{H} = \mathbb{C}^2 = \operatorname{span}\{|\uparrow\rangle, |\downarrow\rangle\}, \iota(v) = vZ$, and $W = \lambda X$, where λ is a parameter. Note that the functional theory defined by $(V, \mathcal{H}, \iota, W)$ is identical to the one in Example 2.6 when $\lambda = 0$. It is easy to verify that $\iota^*(\mathcal{E}) = \iota^*(\mathcal{P}) = [-1, 1] \subset V^* = \mathbb{R}$.

We consider first the pure functional $\mathcal{F}_p(\rho) := \min_{|\Psi\rangle:\langle\Psi|Z|\Psi\rangle=\rho} \langle\Psi|\lambda X|\Psi\rangle$, where $\rho \in [-1, 1]$ is some fixed density. Every state $|\Psi\rangle$ which satisfies $\langle\Psi|Z|\Psi\rangle = \rho$ is of the form

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\Big(\sqrt{1+\rho}\,|\uparrow\rangle + e^{i\phi}\sqrt{1-\rho}\,|\downarrow\rangle\Big),$$

where $e^{i\phi}$ is any phase. Hence,

$$\mathcal{F}_p(\rho) = \min_{\phi\in[0,2\pi]} \lambda\sqrt{1+\rho}\sqrt{1-\rho}\cos\phi = -|\lambda|\sqrt{1+\rho}\sqrt{1-\rho}.$$

Next, we determine the ensemble functional $\mathcal{F}_e(\rho) := \min_{\Gamma:\langle\Gamma,Z\rangle=\rho} \lambda \operatorname{Tr}(\Gamma X)$. Recall that any ensemble state Γ on $\mathbb{C}^2$ can be written as $\Gamma = \frac{1}{2}(\mathbb{1} + a_x X + a_y Y + a_z Z)$, where $a = (a_x, a_y, a_z) \in \mathbb{R}^3$ satisfies $\|a\| \le 1$. Clearly, the condition $\langle\Gamma, Z\rangle = \rho$ is the same as $a_z = \rho$, so

$$\Gamma = \frac{1}{2}(\mathbb{1} + \rho Z + a_x X + a_y Y),$$

where $a_x^2 + a_y^2 \le 1 - \rho^2$. Since $\langle\Gamma, X\rangle = a_x$, we have

$$\mathcal{F}_e(\rho) = \min_{(a_x,a_y):a_x^2+a_y^2\le 1-\rho^2} \lambda a_x = -|\lambda|\sqrt{1-\rho^2}.$$

So we see that $\mathcal{F}_e(\rho) = \mathcal{F}_p(\rho)$. We can visualize the constrained search for both the pure functional and the ensemble functional as shown in Fig. 2.1. Since the Hilbert space $\mathcal{H}$ is two-dimensional, the set of all pure states $\mathcal{P}$ is the Bloch sphere, and the set of ensemble states $\mathcal{E}$ is the convex hull, i.e., the ball. For any ensemble (in particular pure) state Γ, the density map ι^* sends Γ to its z-coordinate on the Bloch ball. The preimage $(\iota^*)^{-1}(\rho)$ then intersects $\mathcal{P}$ and $\mathcal{E}$ in a circle and a disk respectively.

Of course, since $\mathcal{H} = \mathbb{C}^2$, we can also directly calculate the ground state energy $E(v)$ and the ground state $|\Psi(v)\rangle$ for each external potential $v \in \mathbb{R}$. The Hamiltonian is $H(v) = vZ + \lambda X$, so $E(v) = -\sqrt{\lambda^2 + v^2}$ and, for $\lambda \neq 0$,

$$|\Psi(v)\rangle\langle\Psi(v)| = \frac{1}{2}\left(\mathbb{1} - \frac{v}{\sqrt{\lambda^2+v^2}}Z - \frac{\lambda}{\sqrt{\lambda^2+v^2}}X\right).$$

It follows that the Hohenberg-Kohn functional satisfies

$$\begin{aligned}&\mathcal{F}_{HK}(\langle\Psi(v)|Z|\Psi(v)\rangle) = \lambda\,\langle\Psi(v)|X|\Psi(v)\rangle\\ \Rightarrow&\mathcal{F}_{HK}\left(-\frac{v}{\sqrt{\lambda^2+v^2}}\right) = -\frac{\lambda^2}{\sqrt{\lambda^2+v^2}}.\end{aligned}$$

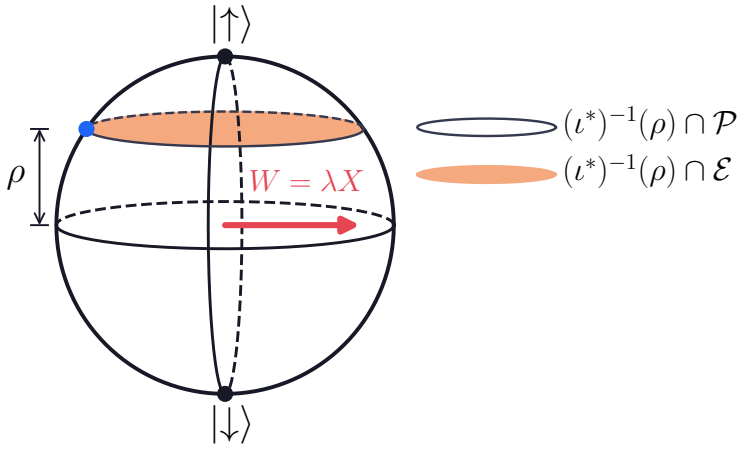


Figure 2.1 Illustration of the geometry of computing the pure and ensemble functionals. The sphere represents the set $\mathcal{P}$ of all pure states, and the ball is the set $\mathcal{E}$ of ensemble states. The value of the pure functional $\mathcal{F}_p$ at a density $\rho \in [-1, 1]$ is defined to be the minimum of $\langle \Gamma, W \rangle$, where Γ is a pure state mapping to ρ. The set of all admissible pure states Γ is $(\iota^*)^{-1}(\rho) \cap \mathcal{P}$, which is nothing but the horizontal circle at height ρ in the figure. Similarly, for the ensemble functional, the set of admissible ensemble states is the orange disk. We may think of the fixed interaction $W = \lambda X$ as a vector in $\mathbb{R}^3$, and in this case it points towards the positive x axis if $\lambda > 0$ (red arrow). The minimizer of $\langle \Gamma, W \rangle$ is then the blue dot on the left for both pure and ensemble states, which explains why $\mathcal{F}_p$ and $\mathcal{F}_e$ agree.

Inverting $\rho = -v/\sqrt{\lambda^2 + v^2}$ then gives

$$\mathcal{F}_{HK}(\rho) = -\frac{\lambda^2}{\sqrt{\lambda^2 + \frac{\lambda^2 \rho^2}{1-\rho^2}}} = -|\lambda|\sqrt{1-\rho^2} \qquad \rho \in (-1, 1).$$

Note that 1 and -1 are not pure state v-representable (unless $\lambda = 0$), so $\mathcal{F}_{HK}$ is not defined there.

If $\lambda = 0$, then $\mathcal{F}_{HK}(\rho) = 0$ by definition, so the formula above still holds. We conclude that the three functionals in the functional theory $(V, \mathcal{H}, \iota, W)$ are identical on $(-1, 1)$ for all λ.

2.5 Properties of the Functionals

So far, we have defined three functionals. These are the Hohenberg-Kohn functional $\mathcal{F}_{HK}$ and the pure and ensemble functionals $\mathcal{F}_p, \mathcal{F}_e$. The practical usefulness of these functionals lies in the fact that they all determine the ground state energy $E(v)$ for *all* external potentials $v \in V$.

Theorem 2.10. *For any external potential $v \in V$, we have*

$$E(v) = \min_{\rho \in \bigcup_{v \in V} \iota^*(\mathbf{G}_p(v))} [\langle \rho, v \rangle + \mathcal{F}_{HK}(\rho)] = \min_{\rho \in \iota^*(\mathcal{P})} [\langle \rho, v \rangle + \mathcal{F}_p(\rho)] = \min_{\rho \in \iota^*(\mathcal{E})} [\langle \rho, v \rangle + \mathcal{F}_e(\rho)]. \tag{2.9}$$

Remark 2.7. The sets involved in the three minimizations are exactly the sets of pure state v-representable, pure state representable and ensemble state representable densities respectively (see comments after Definition 2.3).

Proof. Fix an arbitrary external potential $v \in V$.

For any pure state v-representable density ρ', let Γ' be a pure ground state of $\iota(v') + W$ for some $v' \in V$ such that $\iota^*(\Gamma') = \rho'$. Then

$$\langle \rho', v \rangle + \mathcal{F}_{HK}(\rho') = \langle \rho', v \rangle + \langle \Gamma', W \rangle = \langle \Gamma', \iota(v) + W \rangle \geq E(v).$$

On the other hand, let Γ be a pure ground state of $\iota(v) + W$. Then $E(v) = \langle \iota^*(\Gamma), v \rangle + \langle \Gamma, W \rangle$. This shows the first equality.

The second equality follows from splitting the minimization in the Rayleigh-Ritz variational principle into two steps:

$$\begin{aligned} E(v) &= \min_{\Gamma \in \mathcal{P}} \langle \Gamma, \iota(v) + W \rangle = \min_{\rho \in \iota^*(\mathcal{P})} \min_{\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathcal{P}} \langle \Gamma, \iota(v) + W \rangle \\ &= \min_{\rho \in \iota^*(\mathcal{P})} \Big[\langle \rho, v \rangle + \min_{\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathcal{E}} \langle \Gamma, W \rangle \Big]. \end{aligned}$$

The third equality can be shown in exactly the same way. □

Since the three functionals yield the same ground state energy in the precise sense of Theorem 2.10, one might hope to establish simple relations among them, such as one being an extension of another. As a first step towards understanding what potentially could be said, note that the domains of these functionals sit inside each other:

$$\operatorname{dom} \mathcal{F}_{HK} \subset \operatorname{dom} \mathcal{F}_p \subset \operatorname{dom} \mathcal{F}_e,$$

which are the sets of pure state v-representable, pure state representable and ensemble state representable densities respectively. Hence, we may ask whether $\mathcal{F}_p$ and $\mathcal{F}_e$ are extensions of $\mathcal{F}_{HK}$, the answer to which is positive:

Proposition 2.11. *The functionals $\mathcal{F}_{HK}, \mathcal{F}_p, \mathcal{F}_e$ agree on their common domain, which is the set of pure state v-representable densities $\bigcup_{v \in V} \iota^*(\mathbf{G}_p(v))$.*

Proof. It is clear that $\bigcup_{v \in V} \iota^*(\mathbf{G}_p(v)) \subset \iota^*(\mathcal{P}) \subset \iota^*(\mathcal{E})$: the first inclusion follows from $\mathbf{G}_p(v) \subset \mathcal{P}$, and the second from $\mathcal{P} \subset \mathcal{E}$. Therefore, the common domain is the set of pure state v-representable densities.

Take a pure state v-representable density $\rho \in V^*$ and let $v \in V$ be an external potential with pure ground state Γ such that $\rho = \iota^*(\Gamma)$. Then $\mathcal{F}_{HK}(\rho) = \langle \Gamma, W \rangle$ by definition. We will show that $\mathcal{F}_p(\rho) = \langle \Gamma, W \rangle$.

Take any pure state Γ' such that $\iota^*(\Gamma') = \rho$, then

$$\begin{aligned} &\langle \Gamma', \iota(v) + W \rangle \geq \langle \Gamma, \iota(v) + W \rangle \\ &\Rightarrow \langle \rho, v \rangle + \langle \Gamma', W \rangle \geq \langle \rho, v \rangle + \langle \Gamma, W \rangle \\ &\Rightarrow \langle \Gamma', W \rangle \geq \langle \Gamma, W \rangle, \end{aligned}$$

implying $\mathcal{F}_p(\rho) \geq \langle \Gamma, W \rangle$. But clearly $\mathcal{F}_p(\rho) \leq \langle \Gamma, W \rangle$, so we conclude $\mathcal{F}_p(\rho) = \langle \Gamma, W \rangle$.

It is straightforward to show $\mathcal{F}_e(\rho) = \langle \Gamma, W \rangle$ by repeating the same argument. □

The preceding proposition does not say anything about the values of the functionals outside the set of pure state v-representable densities. In particular, it does *not* claim (nor does it rule out) that the pure and ensemble functionals agree on their common domain, which is $\iota^*(\mathcal{P})$. In fact, there are cases where the two do not agree [83].

Another important relation among $\mathcal{F}_{HK}, \mathcal{F}_p$, and $\mathcal{F}_e$ can be established using some machinery from convex analysis. Recall that if $g : \mathcal{V} \to \mathbb{R} \cup \{\pm\infty\}$ is any function on a vector space $\mathcal{V}$, the *convex conjugate* or *Legendre-Fenchel transform* of g is the function $g^* : \mathcal{V}^* \to \mathbb{R} \cup \{\pm\infty\}$ defined by $g^*(y) = \sup_{x \in U}(\langle x, y\rangle - g(x))$. If a function g is defined on a proper subset of $\mathcal{V}$, we extend it by setting $g(x) = +\infty$ for x outside the domain and define its convex conjugate as that of the extension.

A basic result is that g^* is convex for any g, and g^{**} is the lower convex envelope of g. Moreover, $g^{**} = g$ if and only if g is itself convex, proper and lower semicontinuous (or g is identically $+\infty$ or $-\infty$) by the Moreau-Fenchel theorem. We now show that the ensemble functional satisfies this condition:

Proposition 2.12. *The ensemble functional $\mathcal{F}_e : \iota^*(\mathcal{E}) \to \mathbb{R}$ is convex and continuous. Moreover, its extension to all of V^* by $+\infty$ is convex, proper, and lower semicontinuous.*

Proof. For the proof, let $\tilde{\mathcal{F}}_e : V^* \to \mathbb{R} \cup \{-\infty, +\infty\}$ denote the extension.

Take ensemble state representable densities $\rho_1, \rho_2 \in \iota^*(\mathcal{E})$, and $t \in [0, 1]$. Then the density $\rho := t\rho_1 + (1-t)\rho_2$ is also ensemble state representable (because $\mathcal{E}$ is convex and ι^* is linear), hence $\mathcal{F}_e(\rho)$ is defined. We have

$$\mathcal{F}_e(\rho) = \min_{\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathcal{E}} \langle \Gamma, W \rangle .$$

Since the set $\mathcal{E}$ of all ensemble states on $\mathcal{H}$ is convex, the set $\{\Gamma \in \mathcal{E} \mid \iota^*(\Gamma) = \rho\}$ contains $\{t\Gamma_1 + (1-t)\Gamma_2 \mid \Gamma_i \in \mathcal{E}, \iota^*(\Gamma_1) + \rho_1, \iota^*(\Gamma_2) = \rho_2\}$ as a subset. It follows that we can bound the minimum from above by restricting to the latter set, obtaining

$$\begin{aligned}
\mathcal{F}_e(\rho) &\leq \min_{\Gamma_1 \in (\iota^*)^{-1}(\rho_1) \cap \mathcal{E}} \min_{\Gamma_2 \in (\iota^*)^{-1}(\rho_2) \cap \mathcal{E}} \langle t\Gamma_1 + (1-t)\Gamma_2, W \rangle \\
&= t \min_{\Gamma_1 \in (\iota^*)^{-1}(\rho_1) \cap \mathcal{E}} \langle \Gamma_1, W \rangle + (1-t) \min_{\Gamma_2 \in (\iota^*)^{-1}(\rho_2) \cap \mathcal{E}} \langle \Gamma_2, W \rangle \\
&\leq t\mathcal{F}_e(\rho_1) + (1-t)\mathcal{F}_e(\rho_2).
\end{aligned}$$

So $\mathcal{F}_e$, and therefore also $\tilde{\mathcal{F}}_e$, is convex.

Let $\rho \in \iota^*(\mathcal{E})$ be any ensemble representable density. Suppose $(\rho_i)_i$ is a sequence in $\iota^*(\mathcal{E})$ converging to ρ. Let $\Gamma \in (\iota^*)^{-1}(\rho) \cap \mathcal{E}$ be a minimizer of the constrained search, so that $\mathcal{F}_e(\rho) = \langle \Gamma, W \rangle$. The set-valued map $\rho \mapsto (\iota^*)^{-1}(\rho) \cap \mathcal{E}$ is lower hemicontinuous (see, for example, Proposition 1.5.1 of Ref. [84]), so there exists a sequence $(\Gamma_i)_i$ in $\mathcal{E}$ such that $\iota^*(\Gamma_i) = \rho_i$ for each i and $\Gamma_i \to \Gamma$. Therefore

$$\limsup_{i \to \infty} \mathcal{F}_e(\rho_i) \leq \limsup_{i \to \infty} \langle \Gamma_i, W \rangle = \langle \Gamma, W \rangle = \mathcal{F}_e(\rho).$$

For each i, let $\Gamma_i' \in (\iota^*)^{-1}(\rho_i) \cap \mathcal{E}$ be a minimizer of the constrained search at ρ_i. We can extract a subsequence $(i_j)_j$ such that $\mathcal{F}_e(\rho_{i_j}) \xrightarrow{j \to \infty} \liminf_{i \to \infty} \mathcal{F}_e(\rho_i)$ and $\Gamma_{i_j}' \xrightarrow{j \to \infty} \Gamma'$ for some $\Gamma' \in \mathcal{E}$ by the compactness of $\mathcal{E}$. Then

$$\iota^*(\Gamma') = \lim_{j \to \infty} \iota^*(\Gamma_{i_j}') = \lim_{j \to \infty} \rho_{i_j} = \rho,$$

from which it follows that

$$\liminf_{i \to \infty} \mathcal{F}_e(\rho_i) = \lim_{j \to \infty} \mathcal{F}_e(\rho_{i_j}) = \lim_{j \to \infty} \langle \Gamma_{i_j}', W \rangle = \langle \Gamma', W \rangle \geq \mathcal{F}_e(\rho).$$

This shows that the ensemble functional $\mathcal{F}_e$ is continuous.

The extension $\tilde{\mathcal{F}}_e$ is proper because $\operatorname{dom}\mathcal{F}_e = \iota^*(\mathcal{E}) \neq \varnothing$, and $\operatorname{im}(\tilde{\mathcal{F}}_e) \subset (-\infty, \infty]$. It is lower semicontinuous because $\mathcal{F}_e$ is continuous and $\iota^*(\mathcal{E})$ is closed. □

Remark 2.8. Since $\mathcal{F}_e$ is convex, it is continuous on $\operatorname{relint}(\iota^*(\mathcal{E}))$, so technically we only had to show continuity on $\partial(\iota^*(\mathcal{E}))$.

Next, note that Theorem 2.10 is the same as saying that the ground state energy function E is, up to some minus signs, the convex conjugate of the three functionals. More precisely, one has

$$E(v) = -\mathcal{F}_{HK}^*(-v) = -\mathcal{F}_p^*(-v) = -\mathcal{F}_e^*(-v). \tag{2.10}$$

This implies immediately that the ground state energy function E is concave, giving an alternative proof of Proposition 2.1.

Multiplying Eq. (2.10) by -1 and taking the convex conjugate gives

$$(-E)^*(-\rho) = \mathcal{F}_{HK}^{**}(\rho) = \mathcal{F}_p^{**}(\rho) = \mathcal{F}_e^{**}(\rho) = \mathcal{F}_e(\rho), \tag{2.11}$$

where the last equality follows from the fact that the ensemble functional $\mathcal{F}_e$ is convex (more precisely, that its extension is convex, proper, and lower semicontinuous) due to Proposition 2.12. This proves the following:

Proposition 2.13. *The ensemble functional $\mathcal{F}_e$ is the lower convex envelope of both the Hohenberg-Kohn functional $\mathcal{F}_{HK}$ and the pure functional $\mathcal{F}_p$.*

2.6 The Purification Trick

Recall that any ensemble state Γ on $\mathcal{H}$ is the marginal of some pure state on $\mathcal{H} \otimes \mathcal{H}$. In other words, there always exists a pure state $\bar{\Gamma}$ on $\mathcal{H} \otimes \mathcal{H}$, called a *purification* of Γ, such that $\Gamma = \operatorname{Tr}_2(\bar{\Gamma})$. More precisely, if $\Gamma = \sum_i t_i |\Psi_i\rangle\langle\Psi_i|$ is any ensemble state, we may take $|\bar{\Psi}\rangle = \sum_i \sqrt{t_i}\,|\Psi_i\rangle \otimes |i\rangle$, where $(|i\rangle)_{i=1}^{\dim\mathcal{H}}$ is any orthonormal basis, and $\bar{\Gamma} := |\bar{\Psi}\rangle\langle\bar{\Psi}|$, for which it holds that $\Gamma = \operatorname{Tr}_2(\bar{\Gamma})$. Based on this idea, we now show that the ensemble functional of any generalized functional theory is, in fact, a pure functional of a suitably constructed functional theory.

This has both conceptual and practical appeal. Conceptually, this means that Definition 2.1 gives rise to a class of functional theories large enough such that the notion of "ensemble functional" is a redundant or automatic one as soon as that of "pure functional" is defined. Practically, this result is extremely useful for proving properties of the ensemble functional: any general statement which is true about the pure functional also applies to the ensemble functional, precisely because the latter is a special case of the former. As an example, when we derive a formula for the boundary force via the pure state constrained search (Eq. (2.7)) in Chapter 4, the same formula will apply to the ensemble functional after replacing the Hilbert space with a suitably larger one.

Even though we will primarily be interested in the ensemble functional, which is the lower convex envelope of the pure functional (Proposition 2.13), we present here a slightly more general construction, which will result in a family of not necessarily convex functionals, with the pure and ensemble functionals being two extreme cases.

Definition 2.7. *The k-**convex hull** of a subset X of a vector space is* $\mathrm{conv}_k(X) := \{\sum_{i=1}^k t_i s_i \mid s_i \in X, \sum_{i=1}^k t_i = 1, t_i \geq 0\}$. *The k-**convexification** of a function $f : X \to \mathbb{R}$ is the function* $\mathrm{conv}_k(f) : \mathrm{conv}_k(X) \to \mathbb{R}$, $s \mapsto \inf_{\{t_i\},\{s_i\}:\sum_i t_i s_i = s} \sum_i t_i f(s_i)$ *(where the conditions $\sum_i t_i = 1$ and $t_i \geq 0$ are left implicit).*

Remark 2.9. $\mathrm{conv}_k(\cdot)$ is not idempotent in general, in contrast to $\mathrm{conv}(\cdot)$. However, if $\mathrm{conv}(X)$ is at most $(k-1)$-dimensional, then $\mathrm{conv}(X) = \mathrm{conv}_k(X)$ by Carathéodory's theorem.

Let $(V, \mathcal{H}, \iota, W)$ be a generalized functional theory.

Definition 2.8. *For a positive integer k, the associated k-**convexification** or k-**convexified functional theory** is the generalized functional theory constructed from the data $(V, \mathcal{H} \otimes \mathbb{C}^k, \bar{\iota}, W \otimes \mathbb{1})$, where $\bar{\iota}(v) := \iota(v) \otimes \mathbb{1}$.*

The intuition behind Definition 2.8 is that we want to represent ensemble states of rank at most k on the "primal" Hilbert space $\mathcal{H}$ as pure states on the "primal + ancillas" Hilbert space $\mathcal{H} \otimes \mathbb{C}^k$. If $k = \dim \mathcal{H}$, we will simply refer to the functional theory constructed this way as the *convexified functional theory*.

Objects and notions like the pure functional, pure state representability, etc. in this newly constructed functional theory are then automatically defined. We will typically use a bar $(\bar{\ })$ to indicate that an object belongs to the k-convexified functional theory, so, for example, $\bar{\mathcal{P}}$ refers to the set of pure states on $\mathcal{H} \otimes \mathbb{C}^k$, and $\bar{\mathcal{F}}_p$ denotes the pure functional of the k-convexification.

Lemma 2.14. *The k-convex hull of $\iota^*(\mathcal{P})$ is equal to the set of pure state representable densities in the k-convexification. That is,*

$$\mathrm{conv}_k(\iota^*(\mathcal{P})) = \bar{\iota}^*(\bar{\mathcal{P}}). \tag{2.12}$$

Proof. Take a density $\rho \in V^*$ in the k-convex hull of $\iota^*(\mathcal{P})$. Then $\rho = \sum_{i=1}^k t_i \iota^*(\Gamma_i)$, where $\sum_i t_i = 1$, $t_i \geq 0$, and the $\Gamma_i = |\Psi_i\rangle\langle\Psi_i|$ are pure states on $\mathcal{H}$. Let $(|i\rangle)_{i=1}^k$ be the standard orthonormal basis for $\mathbb{C}^k$. Define $|\bar{\Psi}\rangle := \sum_k \sqrt{t_i}\, |\Psi_i\rangle \otimes |i\rangle$ and $\bar{\Gamma} := |\bar{\Psi}\rangle\langle\bar{\Psi}|$. Then for all $v \in V$

$$\begin{aligned}
\left\langle \bar{\iota}^*(\bar{\Gamma}), v \right\rangle &= \langle \bar{\Psi} | \iota(v) \otimes \mathbb{1} | \bar{\Psi} \rangle = \sum_{i,j=1}^k \sqrt{t_i t_j} (\langle \Psi_i | \otimes \langle i |)(\iota(v) \otimes \mathbb{1})(|\Psi_j\rangle \otimes |j\rangle) \\
&= \sum_{i,j=1}^k \sqrt{t_i t_j}\, \langle \Psi_i | \iota(v) | \Psi_j \rangle\, \delta_{ij} = \sum_{i=1}^k t_i \, \langle \Psi_i | \iota(v) | \Psi_i \rangle = \sum_{i=1}^k t_i \, \langle \iota^*(\Gamma_i), v \rangle \\
&= \left\langle \sum_{i=1}^k t_i \iota^*(\Gamma_i), v \right\rangle = \langle \rho, v \rangle \Rightarrow \bar{\iota}^*(\bar{\Gamma}) = \rho.
\end{aligned}$$

This shows "$\subset$".

To show inclusion in the other direction, take any $|\bar{\Psi}\rangle \in \mathcal{H} \otimes \mathbb{C}^k$. Consider the Schmidt decomposition $|\bar{\Psi}\rangle = \sum_{i=1}^k \lambda_i |\Psi_i\rangle \otimes |\tilde{i}\rangle$, where $(|\tilde{i}\rangle)_{i=1}^k$ is an orthonormal basis for $\mathbb{C}^k$. By the same calculation, we see that $\bar{\iota}^*(|\bar{\Psi}\rangle\langle\bar{\Psi}|) = \sum_{i=1}^k |\lambda_i|^2 \iota^*(|\Psi_i\rangle\langle\Psi_i|$. This shows "$\supset$". □

Corollary 2.15. *The set of ensemble state representable densities is equal to the set of pure state representable densities of the convexified functional theory.*

Lemma 2.14 is equivalent to saying that the domains of the k-convexified functional and the pure functional in the k-convexified functional theory coincide. We now show that the two functionals themselves are indeed the same. In loose terms, "k-convexifying commutes with constructing the pure functional".

Proposition 2.16. *For all k, the k-convexification of the pure functional of a functional theory is the pure functional of the k-convexification of the functional theory. That is,*

$$\operatorname{conv}_k(\mathcal{F}_p) = \bar{\mathcal{F}}_p. \tag{2.13}$$

Proof. Fix a density $\rho \in \operatorname{conv}_k(\iota^*(\mathcal{P})) = \bar{\iota}^*(\bar{\mathcal{P}})$. Unpacking the definitions, we get

$$\begin{aligned}
\operatorname{conv}_k(\mathcal{F}_p)(\rho) &= \inf\Big\{ \sum_{i=1}^{k} t_i \min\{\langle \Gamma, W\rangle \,|\iota^*\Gamma = \rho_i, \Gamma \in \mathcal{P}\} \mid \sum_i t_i\rho_i = \rho, \rho_i \in \iota^*(\mathcal{P}), t \in \Delta_{k-1}\Big\} \\
&= \inf\Big\{ \min\Big\{ \sum_{i=1}^{k} t_i \langle \Gamma_i, W\rangle \,|\iota^*\Gamma_i = \rho_i, \Gamma_i \in \mathcal{P}\Big\} \mid \sum_i t_i\rho_i = \rho, \rho_i \in \iota^*(\mathcal{P}), t \in \Delta_{k-1}\Big\} \\
&= \inf\Big\{ \sum_{i=1}^{k} t_i \langle \Psi_i|W|\Psi_i\rangle \mid \iota^*(|\Psi_i\rangle\langle\Psi_i|) = \rho_i, \sum_i t_i\rho_i = \rho, \rho_i \in \iota^*(\mathcal{P}), \langle\Psi_i|\Psi_i\rangle = 1, t \in \Delta_{k-1}\Big\} \\
&= \inf\Big\{ \sum_{i=1}^{k} t_i \langle \Psi_i|W|\Psi_i\rangle \mid \sum_i t_i\iota^*(|\Psi_i\rangle\langle\Psi_i|) = \rho, \langle\Psi_i|\Psi_i\rangle = 1, t \in \Delta_{k-1}\Big\},
\end{aligned}$$

where we have introduced the standard $(k-1)$-simplex $\Delta_k := \{t \in \mathbb{R}^k \mid \sum_{i=1}^k t_i = 1, t_i \geq 0\}$ to make the constraints look simpler. Also note that the set of admissible $((|\Psi_i\rangle)_{i=1}^k, t)$ in the last line is compact, so we may replace inf with min.

We turn our attention to the pure functional of the k-convexified functional theory. It is defined (see Definition 2.6) as a constrained minimization on $\mathcal{H} \otimes \mathbb{C}^k$:

$$\begin{aligned}
\bar{\mathcal{F}}_p(\rho) &= \min\Big\{ \langle \bar{\Gamma}, W \otimes \mathbb{1}\rangle \mid \bar{\iota}^*(\bar{\Gamma}) = \rho, \bar{\Gamma} \in \bar{\mathcal{P}}\Big\} \\
&= \min\Big\{ \langle \bar{\Psi}|W \otimes \mathbb{1}|\bar{\Psi}\rangle \mid \bar{\iota}^*(|\bar{\Psi}\rangle\langle\bar{\Psi}|) = \rho, \langle\bar{\Psi}|\bar{\Psi}\rangle = 1\Big\}.
\end{aligned}$$

Each pure state $|\bar{\Psi}\rangle \in \mathcal{H} \otimes \mathbb{C}^k$ can be written as $\sum_i \sqrt{t_i}\,|\Psi_i\rangle \otimes |i\rangle$, where $\langle\Psi_i|\Psi_i\rangle = 1$ and $\sum_i t_i = 1$ (note that the $|\Psi_i\rangle$ are not required to be orthogonal). It follows that

$$\bar{\mathcal{F}}_p(\rho) = \min\Big\{ \sum_i t_i \langle \Psi_i|W|\Psi_i\rangle \mid \sum_i t_i\iota^*(|\Psi_i\rangle\langle\Psi_i|) = \rho, \sum_i t_i = 1, t_i \geq 0, \langle\Psi_i|\Psi_i\rangle = 1\Big\}.$$

Therefore $\bar{\mathcal{F}}_p(\rho) = \operatorname{conv}_k(\mathcal{F}_p)$. □

Corollary 2.17. *The ensemble functional of a functional theory is the pure functional of the convexification of the functional theory.*

Chapter 3

Translation Invariant RDMFT for Bosons

In Chapter 2, we presented an abstract framework of functional theories. Apart from the motivating examples and a few rather artificial and simple constructions such as those presented in Examples 2.6 and 2.7, the focus was on extracting the essence shared by all functional theories (for finite-dimensional systems at zero temperature) and exploring the consequences of our abstract definitions. In particular, we did not address in depth how functional theories can arise "in nature".

In this chapter, we discuss specific instances of functional theories that occur naturally when one considers a system of interacting bosons with translation invariance. Recall that in RDMFT, one is concerned with a family of Hamiltonians $H(h) = h + W$, where h is a single-particle operator and W is a fixed interaction. The assumption that the system respects translational symmetry then has two implications: (1) The fact that $H(h)$ commutes with the group of translations means that each Hamiltonian is block-diagonal with respect to the orthogonal decomposition of the Hilbert space into subspaces labeled by irreducible representations of the group of translations. (2) Since not all single-particle operators are translation invariant, the space of allowed h is significantly reduced. In the language of Chapter 2, the space of external potentials is replaced by a subspace of much lower dimension. Consequently, the dimension of the density space, which is the dual, is also reduced.

Compared to Chapter 2, the present one will have more emphasis on physical ideas and less on mathematical rigor. In Chapter 4, we will define and discuss a class of functional theories known as *abelian functional theories*, of which the present bosonic momentum space functional theory is a special case. Hence, this chapter serves also as a precursor to the next one, providing concrete examples to guide our intuition when developing the abstract theory.

The contents of this chapter are based on a forthcoming paper in collaboration with Christian Schilling.

3.1 Motivation

Following the experimental realization of Bose-Einstein condensation (BEC) [85–87], the field of ultracold atoms has undergone rapid advancements [88, 89], providing a versatile platform for quantum simulations [90–92], among other applications. Atoms loaded into an optical lattice generated by laser beams can simulate many-body systems with accurately controllable parameters [93–95], facilitating the study of the bosonic superfluid-Mott insulator transition [96, 97]. Therefore, it is of interest to develop a functional theoretic description of bosonic lattice systems, which was proposed in Ref. [25]. Since the particle density alone contains insufficient information for the characterization of BEC in terms of Penrose and Onsager's off-diagonal long-range order criterion [98], RDMFT is naturally advantageous over (the bosonic variant of) DFT. The theory put forward in Ref. [25], however, neither assumes nor takes advantage of symmetries.

In the following, a functional theory (strictly speaking, a family of functional theories) for bosonic lattice systems with translation invariance is constructed by carrying out a two-step reduction of RDMFT based on symmetry, resulting in a simple characterization of the domains of the pure functional and the ensemble functional. This then allows us to derive an exact functional form, which is in turn intimately tied to the domain's geometry. Unlike previous works, we establish an exact formula for the boundary force, a diverging repulsive force from the boundary of the functional domain, a phenomenon previously observed in a broad range of systems [25, 37–39], generalizing in particular the notion of the BEC force [25, 38].

3.2 Incorporating Translational Symmetry in RDMFT

Consider a system of N spinless bosons on a one-dimensional periodic lattice with d sites. The group of translations is then the cyclic group $\mathbb{Z}/d\mathbb{Z} = \{e, g, g^2, \cdots, g^{d-1}\}$, where the generator g acts on $\mathcal{H}_1$ by shifting site i to site $i+1$. In momentum space, we have $g\,|k\rangle = e^{\frac{2\pi i}{d}\cdot k}\,|k\rangle$, where $|0\rangle, \cdots, |d-1\rangle$ denote the momentum space basis for $\mathcal{H}_1$. Let $|m_0, \cdots, m_{d-1}\rangle \propto \prod_k (b_k^\dagger)^{m_k}\,|\rangle$ denote the normalized N-particle state with m_k bosons having momentum k for each k, where the $b_k^\dagger$ and b_k are the creation and annihilation operators in momentum space respectively. Such a state will be referred to as a *permanent* with *occupation number vector* $m = (m_0, \cdots, m_{d-1}) \in \mathbb{R}^d$.

In terms of the permanents, the action of the translation group on $\mathcal{H}_N$ is given by

$$g\,|m_0, \cdots, m_{d-1}\rangle = e^{\frac{2\pi i}{d}\sum_{k=0}^{d-1} k m_k}\,|m_0, \cdots, m_{d-1}\rangle\,. \tag{3.1}$$

The Hilbert space $\mathcal{H}_N$ decomposes into the direct sum of isotypic components with respect to the $\mathbb{Z}/d\mathbb{Z}$-representation:

$$\mathcal{H}_N = \bigoplus_{P=0}^{d-1} \mathcal{H}_N^{(P)},$$

where the total momentum P labels the isomorphism classes of the irreducible representations of $\mathbb{Z}/d\mathbb{Z}$. For each P, the subspace $\mathcal{H}_N^{(P)}$ is spanned by all permanents with total momentum

P:

$$\mathcal{H}_N^{(P)} = \mathrm{span}_{\mathbb{C}} \left\{ |m_0, \cdots, m_{d-1}\rangle \;\middle|\; \sum_{k=0}^{d-1} k m_k \equiv P \pmod d \right\}. \tag{3.2}$$

As an example, suppose there are $N = 3$ bosons on $d = 3$ lattice sites. The N-particle Hilbert space decomposes into $d = 3$ subspaces, labeled by the total momentum $P = 0, 1, 2$. For $P = 0$, we have $\mathcal{H}_3^{(0)} = \mathrm{span}_{\mathbb{C}}\{|3,0,0\rangle, |0,3,0\rangle, |0,0,3\rangle, |1,1,1\rangle\}$.

Note that the condition $\sum_k k m_k \equiv P \pmod d$ is equivalent to $e^{\frac{2\pi i}{d}\sum_k k m_k} = e^{2\pi i P/d}$, so we are simply decomposing $\mathcal{H}_N$ into eigenspaces of the generator g of the translation group $\mathbb{Z}/d\mathbb{Z}$. Since the Hamiltonian is block-diagonal with respect to this decomposition, the ground state problem can be considered separately in each symmetry subspace. Hence, we fix a specific total momentum $P \in \{0, \cdots, d-1\}$ and restrict our attention to the subspace $\mathcal{H}_N^{(P)}$ in the following.

The second part of the construction of the symmetry-adapted functional theory starts with the fact that any one-particle operator h which commutes with $\mathbb{Z}/d\mathbb{Z}$ must be of the form $h = \sum_{k=0}^{d-1} v_k n_k$, where $n_k := b_k^\dagger b_k$ is the number operator in momentum space. Hence, the space of allowed h is parametrized by $v = (v_0, \cdots, v_{d-1}) \in \mathbb{R}^d$. To be precise, we have a map

$$\begin{aligned} \iota : \mathbb{R}^d &\to i\mathfrak{u}(\mathcal{H}_N^{(P)}) \\ v &\mapsto \sum_{k=0}^{d-1} v_k n_k. \end{aligned}$$

Note that in the RDMFT literature v is usually used to denote the actual external potential, which is diagonal in position space, and does not include the kinetic operator. Here, however, we choose to be as consistent as possible with the notation used in Chapter 2.

Definition 3.1. *For particle number N, number of lattice sites d, total momentum $P \in \{0, \cdots, d-1\}$, and interaction W, the* ***bosonic momentum space functional theory at*** *(d, N, P) is the generalized functional theory defined by the data $(\mathbb{R}^d, \mathcal{H}_N^{(P)}, \iota, \tilde{W})$, where $\tilde{W}$ is the restriction of W to $\mathcal{H}_N^{(P)}$.*

In this chapter, we will often simply write W instead of $\tilde{W}$ for the restriction. Also note that although the potential map ι depends on P, this dependence is suppressed.

One can check by direct calculation that the density map is given by (recall that $\langle \Gamma, n_i \rangle = \mathrm{Tr}(\Gamma n_i)$)

$$\begin{aligned} \iota^* : \mathcal{E}^{(P)} &\to \mathbb{R}^d \\ \Gamma &\mapsto \left(\langle \Gamma, n_0 \rangle, \cdots, \langle \Gamma, n_{d-1} \rangle \right), \end{aligned}$$

where $\mathcal{E}^{(P)}$ denotes the set of ensemble states on $\mathcal{H}_N^{(P)}$. Similarly, $\mathcal{P}^{(P)}$ will denote the set of pure states. Let $\mathcal{F}_{HK}^{(P)}, \mathcal{F}_p^{(P)}$, and $\mathcal{F}_e^{(P)}$ be the Hohenberg-Kohn functional, pure functional, and ensemble functional of the bosonic momentum space functional theory respectively. Unpacking the definitions, we have

$$\mathcal{F}_p^{(P)}(m) = \min_{\substack{|\Psi\rangle \in \mathcal{H}_N^{(P)}: \\ \langle \Psi | n_k | \Psi \rangle = m_k \forall k}} \langle \Psi | W | \Psi \rangle, \tag{3.3}$$

and similarly for $\mathcal{F}_e^{(P)}$. The domain $\operatorname{dom}\mathcal{F}_p^{(P)}$, resp. $\operatorname{dom}\mathcal{F}_e^{(P)}$, is then the set of all momentum occupation number vectors $m \in \mathbb{R}^d$ for which there exists a pure, resp. ensemble, N-particle state Γ on $\mathcal{H}_N^{(P)}$ such that $\langle \Gamma, n_k \rangle = m_k$ for all k (c.f. Definition 2.6). This representability problem will be addressed in Section 3.3.

Starting from Eq. (3.3), it is a straightforward exercise to verify that the ground state energy of $\sum_{k=0}^{d-1} v_k n_k + W$ on all of $\mathcal{H}$ is given by $\min_P E^{(P)}(v)$, where $E^{(P)}$ is the ground state energy function of the momentum space functional theory at (d, N, P). Although all results in this chapter are general, we will often choose a specific interaction W for definiteness. A convenient and also physically relevant choice is the two-particle interaction of the Bose-Hubbard model, which is given by

$$W = \frac{1}{d} \sum_{k_1,k_2,k_3,k_4=0}^{d-1} \delta_{k_1+k_2,k_3+k_4} b_{k_1}^\dagger b_{k_2}^\dagger b_{k_3} b_{k_4}, \tag{3.4}$$

where the Kronecker delta is evaluated modulo d.

3.3 Functional Domains

The goal of this section is to determine the domains of the functionals $\mathcal{F}_p^{(P)}, \mathcal{F}_e^{(P)}$ in each subspace $\mathcal{H}_N^{(P)}$, which amounts to characterizing the sets $\iota^*(\mathcal{P}^{(P)})$ and $\iota^*(\mathcal{E}^{(P)})$.

Obvious necessary representability conditions on $m \in \mathbb{R}^d$ are $\sum_k m_k = N$ and $m_k \geq 0$ for all $k \in \{0, \cdots, d-1\}$, which cut out a $(d-1)$-simplex in $\mathbb{R}^d$. This is expected: the identity $\mathbb{1} = N^{-1} \sum_{k=0}^{d-1} n_k$ is in the image of the potential map ι, so the codimension of $\iota^*(\mathcal{E}^{(P)})$ must be at least one by Proposition 2.3. These constraints are, however, not sufficient in general. As we will show in a moment, for most choices of (d, N, P), the domains $\operatorname{dom}\mathcal{F}_p^{(P)}, \operatorname{dom}\mathcal{F}_e^{(P)}$ are simplices with certain "forbidden regions" removed around the vertices, characterized by an additional set of inequality constraints. These forbidden regions strongly constrains the kinematics of the system, not only preventing the quantum state (strictly speaking, its momentum occupation number vector m) from exiting the domain, but actually exerting a *repulsive force* that drives the state away from the boundary. This latter claim, which is at this point not at all obvious, will be made precise and shown in Section 3.5, and provides a solid motivation for investigating the geometric structure of $\operatorname{dom}\mathcal{F}_p^{(P)}$ and $\operatorname{dom}\mathcal{F}_e^{(P)}$.

Theorem 3.1. *For each value of the total momentum $P \in \{0, \cdots, d-1\}$, the domains of $\mathcal{F}_p^{(P)}$ and $\mathcal{F}_e^{(P)}$ coincide, and are both given by*

$$\operatorname{conv} \underbrace{\left\{ m \in \mathbb{N}_0^d \;\middle|\; \sum_{k=0}^{d-1} m_k = N,\ \sum_{k=0}^{d-1} k m_k \equiv P \pmod d \right\}}_{=:\Omega^{(P)}}, \tag{3.5}$$

the convex hull of the occupation number vectors of all permanents with total momentum P.

Proof. For any N-particle state $|\Psi\rangle$ with momentum P, i.e., $|\Psi\rangle \in \mathcal{H}_N^{(P)}$, we have $|\Psi\rangle = \sum_{\alpha=1}^{\dim \mathcal{H}_N^{(P)}} c_\alpha |m^{(\alpha)}\rangle$, where each $m^{(\alpha)}$ satisfies $\sum_{k=0}^{d-1} k m_k^{(\alpha)} \equiv P \pmod d$. Hence, the momen-

tum occupation number vector m of $|\Psi\rangle$ is given by

$$m_k = \langle\Psi|n_k|\Psi\rangle = \sum_{\alpha,\beta=1}^{\dim \mathcal{H}_N^{(P)}} \bar{c}_\alpha c_\beta \langle m^{(\alpha)}|n_k|m^{(\beta)}\rangle = \sum_\alpha |c_\alpha|^2 m_k^{(\alpha)}, \tag{3.6}$$

where we have used $\langle m^{(\alpha)}|n_k|m^{(\beta)}\rangle = m_k^{(\alpha)}\delta_{\alpha\beta}$. Looking at Eq. (3.6), we immediately see that the occupation number vector m sweeps through the entirety of the convex hull of the $m^{(\alpha)}$ as the wave function coefficients c_α are varied. This shows $\iota^*(\mathcal{P}^{(P)}) = \mathrm{conv}(\Omega^{(P)})$.

$\iota^*(\mathcal{E}^{(P)})$ coincides with $\iota^*(\mathcal{P}^{(P)})$ because the former is the convex hull of the latter, which is already convex. □

Example 3.1. For $d = 2$ lattice sites, the N-particle Hilbert space is spanned by the permanents

$$|N,0\rangle, |N-1,1\rangle, \cdots, |0,N\rangle.$$

There are two possible values of the total momentum P, namely 0 and 1. Assuming N is even, then the $P = 0$ subspace is spanned by $|N,0\rangle, |N-2,2\rangle, \cdots, |0,N\rangle$ and the $P = 1$ subspace by $|N-1,1\rangle, |N-3,3\rangle, \cdots, |1,N-1\rangle$. According to Theorem 3.1, we have

$$\begin{aligned} \mathrm{dom}(\mathcal{F}_e^{(0)}) &= \mathrm{dom}(\mathcal{F}_p^{(0)}) = \{(n_0,n_1) \in \mathbb{R}^2 \mid n_0 + n_1 = N, 0 \le n_1 \le N\} \\ \mathrm{dom}(\mathcal{F}_e^{(1)}) &= \mathrm{dom}(\mathcal{F}_p^{(1)}) = \{(n_0,n_1) \in \mathbb{R}^2 \mid n_0 + n_1 = N, 1 \le n_1 \le N-1\}. \end{aligned} \tag{3.7}$$

The domains for $N = 4$ are illustrated in Fig. 3.1a and Fig. 3.1b.

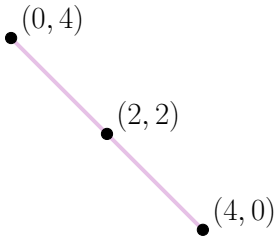
For general (d, N, P), the domain will be more complicated. However, we can make several observations:

1. Whenever d divides N, the states $|N,0,\cdots,0\rangle, |0,N,0,\cdots,0\rangle, \cdots$ all belong to the zero-momentum subspace. Therefore, $\mathrm{dom}\,\mathcal{F}_e^{(0)} = \mathrm{dom}\,\mathcal{F}_p^{(0)}$ is a simplex (see Figs. 3.1a, 3.1c, 3.1e and 3.1g).
2. In the limit $N \to \infty$ at fixed d and P, each vertex of the simplex $\{m \in \mathbb{R}^d \mid \sum_k m_k = N, m_k \ge 0\}$ becomes closer (relative to N) to the nearest occupation number vector $m^{(\alpha)}$ with total momentum P, so the domain approaches a simplex (see Fig. 3.1f).

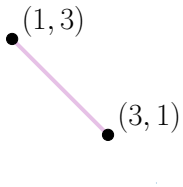
Figs. 3.1a, 3.1c and 3.1e to 3.1g also show a distinctive feature of the bosonic setting in contrast to fermions (see Ref. [37]): in general, there are occupation number vectors that are not extreme points. As we will see in Section 3.5, both the shape of the domain and the distribution of the occupation number vectors of the permanents therein will influence the behavior of the function near the boundary.

3.4 Form of the Universal Functional

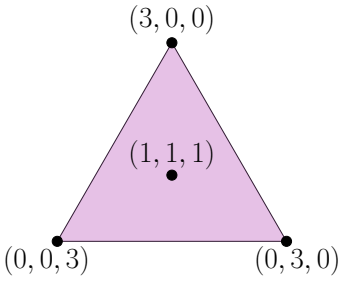
In usual RDMFT without incorporating symmetries, as outlined in Chapter 1, various approximate forms of the universal functional have been proposed in the literature, mostly for electronic systems [20, 29, 32, 34, 36, 99]. In isolated cases, exact analytical functionals of



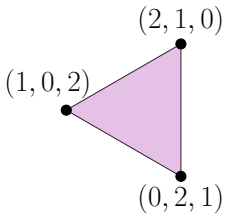
(a) $d = 2, N = 4, P = 0$



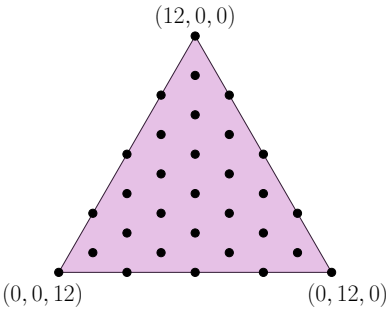
(b) $d = 2, N = 4, P = 1$



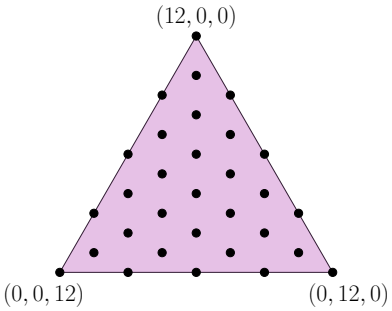
(c) $d = 3, N = 3, P = 0$



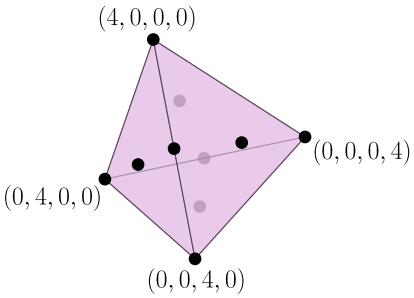
(d) $d = 3, N = 3, P = 1$



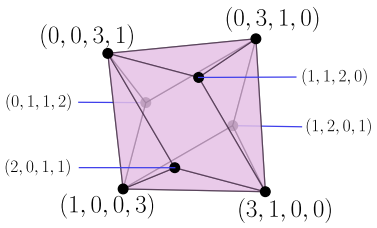
(e) $d = 3, N = 12, P = 0$



(f) $d = 3, N = 12, P = 1$



(g) $d = 4, N = 4, P = 0$



(h) $d = 4, N = 4, P = 1$

Figure 3.1 Structure of N-particle momentum space permanents and domains of the functional.

particular model systems have been derived and studied [24, 25]. However, the presence of symmetries is rarely exploited in the way it is in Sections 3.2 and 3.3. These simplifications due to translational symmetry were explored in Ref. [37], where the pure functional $\mathcal{F}_p$ was found to take a form that heavily depends on the geometry of the domain. The goal of this section is to adapt the results of Ref. [37] to bosons. In light of the purification trick (see Section 2.6), we will focus on the pure functional from now on, and the superscript on $\mathcal{F}_p^{(P)}$ will be dropped, with the understanding that $\mathcal{F}_p$ always refers to the pure state functional of the bosonic momentum space functional theory at some fixed (d, N, P).

As is apparent in Section 3.3, the domain of $\mathcal{F}_p$ is a subset of the hyperplane $\{m \in \mathbb{R}^n \mid \sum_k m_k = N\}$ bounded by a list of $J = J(d, N, P)$ affine constraints

$$D^{(j)}(m) = \langle m, S^{(j)} \rangle - \nu^{(j)} \geq 0, \tag{3.8}$$

where $S^{(j)} \in \mathbb{R}^d$, $\nu^{(j)} \in \mathbb{R}$, and $j = 1, \cdots, J$ labels the facets. Note that there is some ambiguity in the choice of $S^{(j)}$ and $\nu^{(j)}$. For instance, scaling them or replacing them with

$$\begin{aligned} S^{(j)} &\to S^{(j)} + (\mu, \cdots, \mu) \\ \nu^{(j)} &\to \nu^{(j)} + \mu N \end{aligned}$$

for any μ results in an equivalent constraint. This ambiguity, however, will not affect any physical result in this section.

Example 3.2. Take $(d, N, P) = (3, 3, 0)$ as an example. As one can see in Fig. 3.1c, the domain is $\{m \in \mathbb{R}^3 \mid m_k \geq 0, m_1 + m_2 + m_3 = 3\}$, whence

$$\begin{aligned} D^{(1)}(m) &= m_1 \\ D^{(2)}(m) &= m_2 \\ D^{(3)}(m) &= m_3. \end{aligned} \tag{3.9}$$

That is, we can take $S^{(j)}$ to be the j-th standard basis vector e_j and $\nu^{(j)} = 0$.

The idea is to express $\mathcal{F}_p(m)$ in terms of $D^{(j)}(m), j = 1, \cdots, J$, with $D^{(j)}(m)$ having the geometric interpretation of being the distance from m to the j-th facet (up to a multiplicative factor). We proceed in two steps: first, we treat a simplified scenario, the so-called "simplex setting" (see below for precise definition). Although a general combination of (d, N, P) will not satisfy the simplex assumption, this idealized setting will allow us to derive the functional in a straightforward way, illustrating the essential features without encountering technical complications. Next, in Section 3.4.2, armed with the intuition obtained in the simplex setting, we will drop this assumption and derive a functional form in full generality.

3.4.1 The Simplex

In this section, we will assume that the domain is a simplex with exactly $\dim \mathcal{H}_N^{(P)}$ vertices, meaning in particular that the momentum occupation number vector $m^{(\alpha)}$ of each permanent $|m^{(\alpha)}\rangle$ has to be a vertex. This is an extremely strong condition to impose on (d, N, P), and only specific combinations of low dimension and low particle number will result in such a scenario. For example, Figs. 3.1b and 3.1d satisfy this assumption, while the remaining do

not. Note that although the functional domains in Figs. 3.1a, 3.1c, 3.1e and 3.1g are simplices geometrically, they are excluded from the present section because of the extra $m^{(\alpha)}$ vectors which are not extreme points. Hereafter, *the simplex assumption* or *simplex setting* will always refer to this stronger condition than merely that the domain be a simplex.

Let $(m^{(\alpha)})_{\alpha=1}^{\dim \mathcal{H}_N^{(P)}}$ be the list of vertices. The Hilbert space $\mathcal{H}_N^{(P)}$ is then spanned by the permanents $|m^{(1)}\rangle, |m^{(2)}\rangle, \cdots$. Since there is a one-to-one correspondence between vertices and facets, the number J of facets is equal to $\dim \mathcal{H}_N^{(P)}$, and we will use Greek letters α, β to label the constraints instead of j, as we did in Eq. (3.8). To be more precise, we label the constraints so that $D^{(\alpha)} = 0$ corresponds to the facet opposite the vertex $m^{(\alpha)}$.

Suppose we are interested in the value of $\mathcal{F}_p$ at some $m \in \operatorname{dom} \mathcal{F}_p$. If $|\Psi\rangle = \sum_\alpha c_\alpha |m^{(\alpha)}\rangle$ is a state such that $\iota^*(|\Psi\rangle\langle\Psi|) = m$, i.e., $\langle\Psi|n_k|\Psi\rangle = m_k$ for all k, then

$$\sum_\alpha |c_\alpha|^2 m^{(\alpha)} = m$$

(see Eq. (3.6)). Applying $D^{(\beta)}$ to both sides, we get $\sum_\alpha |c_\alpha|^2 D^{(\beta)}(m^{(\alpha)}) = D^{(\beta)}(m)$. But $D^{(\beta)}(m^{(\alpha)}) = L_\alpha \delta_{\alpha\beta}$, where $L_\alpha := D^{(\alpha)}(m^{(\alpha)})$, because every vertex except $m^{(\beta)}$ lies on the hyperplane $D^{(\beta)} = 0$. It follows that

$$|c_\beta|^2 = \frac{D^{(\beta)}(m)}{L_\alpha}, \tag{3.10}$$

which has a geometric interpretation as a distance ratio. Therefore

$$\mathcal{F}_p(m) = \sum_{\alpha,\beta=1}^{\dim \mathcal{H}_N^{(P)}} W_{\alpha\beta} \bar{\xi}_\alpha \xi_\beta \sqrt{\frac{D^{(\alpha)}(m)}{L_\alpha}} \sqrt{\frac{D^{(\beta)}(m)}{L_\beta}}, \tag{3.11}$$

where $W_{\alpha\beta} := \langle m^{(\alpha)}|W|m^{(\beta)}\rangle$ and the ξ_α are phases chosen so that Eq. (3.11) is minimized.

Example 3.3. Consider $N = 3$ bosons on $d = 3$ lattice sites with total momentum $P = 1$ (Fig. 3.1d). The spanning permanents are

$$|m^{(1)}\rangle = |2,1,0\rangle, |m^{(2)}\rangle = |0,2,1\rangle, |m^{(3)}\rangle = |1,0,2\rangle.$$

The domain of the functional is given by the three inequality constraints

$$\begin{aligned} D^{(1)}(m) &= m_0 - m_2 + 1 \geq 0 \\ D^{(2)}(m) &= m_1 - m_0 + 1 \geq 0 \\ D^{(3)}(m) &= m_2 - m_1 + 1 \geq 0, \end{aligned} \tag{3.12}$$

together with the equality constraint $m_0 + m_1 + m_2 = 3$. We have $L_1 = L_2 = L_3 = 3$, so Eq. (3.11) becomes

$$\begin{aligned} &\mathcal{F}_p(m) = \\ &U_0 + \frac{2}{3} U_1 \operatorname{Re}\left[\bar{\xi}_1 \xi_2 \sqrt{D^{(1)}(m) D^{(2)}(m)} + \bar{\xi}_2 \xi_3 \sqrt{D^{(2)}(m) D^{(3)}(m)} + \bar{\xi}_3 \xi_1 \sqrt{D^{(3)}(m) D^{(1)}(m)}\right], \end{aligned} \tag{3.13}$$

where we have assumed $W_{11} = W_{22} = W_{33} = U_0$ and $W_{\alpha\neq\beta} = U_1$, which is the case for the Hubbard model. If U_1 is positive, minimization over the phases ξ_α is not trivial but still relatively simple. A straightforward calculation shows that $\mathcal{F}_p(m) = U_0 - U_1$ if $\sqrt{D^{(1)}(m)}, \sqrt{D^{(2)}(m)}, \sqrt{D^{(3)}(m)}$ satisfy the triangle inequality (that is, $\sqrt{D^{(1)}(m)} \leq \sqrt{D^{(2)}(m)} + \sqrt{D^{(3)}(m)}$ plus cyclic permutations), otherwise

$$\mathcal{F}_p(m) = U_0 + \frac{2}{3}U_1 \left(\sqrt{D^{(1)}(m)D^{(2)}(m)} - \sqrt{D^{(2)}(m)D^{(3)}(m)} - \sqrt{D^{(3)}(m)D^{(1)}(m)} \right)$$

if $D^{(3)}(m) \geq D^{(1)}(m), D^{(2)}(m)$ (and similarly for the other two possibilities).

3.4.2 Beyond Simplex

Outside of the simplex setting, that is, whenever the domain is not a simplex with $\dim \mathcal{H}_N^{(P)}$ vertices, it is no longer possible to recover $(|c_\alpha|^2)_{\alpha=1}^{\dim \mathcal{H}_N^{(P)}}$ uniquely from $(D^{(j)}(m))_{j=1}^J$ as we did in Eq. (3.10). Instead, the collection of numbers $D^{(j)}(m)$ depend linearly on the squared moduli of the wave function coefficients $|c_\alpha|^2$, with the associated linear map having a nontrivial kernel. Hence, the procedure to derive an equation analogous to Eq. (3.11) involves partially inverting said linear map, a method elaborated in the following.

Fix some $m \in \operatorname{dom} \mathcal{F}_p$, and suppose $|\Psi\rangle = \sum_\alpha c_\alpha |m^{(\alpha)}\rangle$ is a state in $\mathcal{H}_N^{(P)}$ with momentum occupation number vector m. Then, by Eq. (3.6),

$$\begin{aligned} D^{(j)}(m) &:= \langle m, S^{(j)} \rangle - \nu^{(j)} = \langle \Psi | \sum_{k=0}^{d-1} S_k^{(j)} n_k - \nu^{(j)} \mathbb{1} | \Psi \rangle \\ &= \sum_\alpha |c_\alpha|^2 \left(\langle S^{(j)}, m^{(\alpha)} \rangle - \nu^{(j)} \right) \\ &= \sum_\alpha |c_\alpha|^2 D^{(j)}(m^{(\alpha)}). \end{aligned} \tag{3.14}$$

Eq. (3.14) can be written more compactly as

$$D(m) = Ty, \tag{3.15}$$

where $D(m) := (D^{(1)}(m), \cdots, D^{(J)}(m))$, T is the $J \times \dim \mathcal{H}_N^{(P)}$ matrix defined by $T_{j\alpha} := D^{(j)}(m^{(\alpha)})$, and $y \in \mathbb{R}^{\dim \mathcal{H}_N^{(P)}}$ is the vector with components $y_\alpha := |c_\alpha|^2$. The situation is as follows: we know m, and hence $D(m)$ (remember the components of $D(m)$ are, up to scaling, the distances from m to the J facets), and would like to find all y that solve Eq. (3.15) (and have nonnegative entries).

Recall that if $A = U\Sigma V^\top$ is the singular value decomposition of a real matrix A, the *Moore-Penrose inverse* of A is defined by $A^+ = V\Sigma^+ U^\top$, where Σ^+ is obtained from taking the transpose of Σ and inverting its nonzero entries. The Moore-Penrose inverse satisfies $AA^+A = A$ and $A^+AA^+ = A^+$.

Let T^+ denote the Moore-Penrose inverse of T. Then Eq. (3.15) is solved by

$$y = T^+ D(m) + x, \tag{3.16}$$

where $x \in \mathbb{R}^{\dim \mathcal{H}_N^{(P)}}$ is any vector such that (1) $Tx = 0$ and (2) the resulting y has nonnegative components, with condition (2) depending on m. To see that Eq. (3.16) really solves Eq. (3.15), let T act on both sides of Eq. (3.16) to get $Ty = TT^+D(m)$. By pure state representability of m, there exists a state $|\Psi'\rangle = \sum_\alpha c'_\alpha |m^{(\alpha)}\rangle$ such that $\iota^*(|\Psi'\rangle\langle\Psi'|) = m$, implying $D(m) = Ty'$, where $y'_\alpha := |c'_\alpha|^2$. So we have $Ty = TT^+Ty' = Ty' = D(m)$.

Therefore

$$|\Psi\rangle = \sum_{\alpha=1}^{\dim \mathcal{H}_N^{(P)}} \xi_\alpha \sqrt{(T^+D(m))_\alpha + x_\alpha}\, |m^{(\alpha)}\rangle\,, \tag{3.17}$$

where the ξ_α are phases, which at this point are arbitrary. As a result,

$$\mathcal{F}_p(m) = \sum_{\alpha,\beta=1}^{\dim \mathcal{H}_N^{(P)}} W_{\alpha\beta}\bar{\xi}_\alpha\xi_\beta \sqrt{(T^+D(m))_\alpha + x_\alpha}\sqrt{(T^+D(m))_\beta + x_\beta}. \tag{3.18}$$

For each $m \in \operatorname{dom}\mathcal{F}_p$, the phases ξ_α and the vector x are to be chosen so that Eq. (3.18) is minimized. If, in addition to translation invariance, we require that the total Hamiltonian $H(v)$ also be invariant under combined action of lattice inversion and time-reversal, then we may take the phases ξ_α to be real-valued, which would lead to a different but equivalent functional $\tilde{\mathcal{F}}_p$ (in the sense of yielding the same ground state energy, see Ref. [73] for more detail). One approximation strategy for $\tilde{\mathcal{F}}_p$ would then be to divide the domain into disjoint regions, and fixing a sign pattern $(\xi_\alpha)_\alpha$ for all m in each region. We do not pursue this further here.

Eq. (3.18) reveals a remarkable relation between the functional $\mathcal{F}_p$ and its domain. We see that the "weight" of each permanent $|m^{(\alpha)}\rangle$ is controlled by $D(m)$, which quantifies the distances from m to each facet. Moreover, all quantities in the two square roots in Eq. (3.18) are purely geometric: they have nothing to do with the interaction W (of course, the minimization over x will, in general, depend on W). To see how Eq. (3.18) reduces to Eq. (3.11) in the simplex setting (remember that by "simplex setting" we really mean *two* conditions: that the domain is a simplex and that $\dim \mathcal{H}_N^{(P)}$ is equal to the number of vertices), observe that $T_{\beta\alpha} = L_\alpha\delta_{\beta\alpha}$ by the definition of L_α, and that the kernel of T is trivial. Therefore, the extra degrees of freedom parametrized by x disappear from Eq. (3.18), and $T^+_{\beta\alpha}$ is simply $L_\beta^{-1}\delta_{\beta\alpha}$. Consequently, we recover Eq. (3.11).

Example 3.4. Take $d = 2, N = 4, P = 0$. The domain of $\mathcal{F}_p$ is given by the convex hull of $m^{(1)} = (4,0)$, $m^{(2)} = (2,2)$, $m^{(3)} = (0,4)$ (see Fig. 3.1a). There are two facets, which in this case are simply the vertices $m^{(1)}$ and $m^{(3)}$ since the domain is one-dimensional, leading to two respective constraints $D^{(1)}(m) = m_0$ and $D^{(2)}(m) = m_1$. Thus

$$T = (D^{(j)}(m^{(\alpha)}))_{j\alpha} = \begin{pmatrix} 4 & 2 & 0 \\ 0 & 2 & 4 \end{pmatrix}, \tag{3.19}$$

the Moore-Penrose inverse of which is

$$T^+ = \frac{1}{24}\begin{pmatrix} 5 & -1 \\ 2 & 2 \\ -1 & 5 \end{pmatrix}. \tag{3.20}$$

The kernel of T is spanned by the vector $(-1,2,-1)$, so $x = q(-1,2,-1)$ with $q \in \mathbb{R}$. It follows that the radicands in Eq. (3.18) are

$$\begin{aligned} T^+D(m) + x &= \frac{1}{24}\begin{pmatrix} 5 & -1 \\ 2 & 2 \\ -1 & 5 \end{pmatrix}\begin{pmatrix} m_0 \\ m_1 \end{pmatrix} + q\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \\ &= \frac{1}{24}\begin{pmatrix} 5m_0 - m_1 \\ 2m_0 + 2m_1 \\ -m_0 + 5m_1 \end{pmatrix} + q\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} + \frac{1}{8}(m_0 - m_1) - q \\ \frac{1}{3} + 2q \\ \frac{1}{3} - \frac{1}{8}(m_0 - m_1) - q \end{pmatrix}, \end{aligned} \tag{3.21}$$

where we have used $m_0 + m_1 = 4$. If we assume $W_{13} = 0$, $W_{11} = W_{33}$, and $W_{12} = W_{23} =: w$, which is true for the Hubbard model (Eq. (3.4)), and set $W_{11} = W_{33} = 0$ without loss of generality (we can always do this by shifting the energy by a constant), then a choice of minimizing phases is $(q_1, q_2, q_3) = (1, -1, 1)$ and Eq. (3.18) gives

$$\mathcal{F}_p(m) = W_{22}\left(\frac{1}{3} + 2q\right) - 2|w|\sqrt{\frac{1}{3} + 2q}\left[\sqrt{\frac{1}{3} + \frac{1}{8}(m_0 - m_1) - q} + \sqrt{\frac{1}{3} - \frac{1}{8}(m_0 - m_1) - q}\right], \tag{3.22}$$

minimized over q.

We would like to point out that Eq. (3.18) is an *exact* expression of the pure functional $\mathcal{F}_p$. Of course, expressing $\mathcal{F}_p$ in this form does not trivialize in any way the task of computing $\mathcal{F}_p$: for a general W, there is no simple way of determining x and the phases ξ_α (which will depend on m), and for interesting systems where $\dim \mathcal{H}_N^{(P)} \gg d$, the nullity of the matrix T, which is the dimension of the space of candidates for x, will be extremely large. Nonetheless, Eq. (3.18) establishes an important link between $\mathcal{F}_p$ and the geometry of its domain, and hints at a potential approximation scheme for the functional through a suitable estimate of $x(m)$.

3.5 Generalized BEC Force

In this section, we carry out a quantitative investigation of the behavior of the functional close to the boundary of its domain.

Observe that in Eq. (3.11) the functional takes the approximate form $\mathcal{F}_p(m) \sim F_0 - \mathcal{G}\sqrt{D^{(\omega)}(m)}$ whenever $D^{(\omega)}(m)$ approaches zero, where F_0 and $\mathcal{G}$ are constants. Consequently, the derivative $\mathrm{d}_m\mathcal{F}_p$ diverges as $1/\sqrt{D^{(\omega)}(m)}$ as $D^{(\omega)}(m) \to 0$. The same phenomenon is demonstrated in Fig. 3.2, where we plot the quantity $\|\mathrm{d}_m\mathcal{F}_p\|$ for the Hubbard model with $d = 3$ (so that the domain is 2-dimensional) and several combinations of (N, P). One may infer by inspecting the plots in Fig. 3.2 that $\mathrm{d}\mathcal{F}_p$ always diverges near the boundary in all cases.

This turns out to be a commonly observed feature in RDMFT: whenever the variable playing the role of a "density" approaches the boundary of the domain of the functional, the derivative diverges as one over the square root of the distance to the boundary [25, 37–39]. In fermionic systems [37], this has been dubbed the "exchange force", while for bosons analogous behavior of the functional near the BEC vertex $(N, 0, \cdots)$ is called the "BEC force" [38, 39].

As we will demonstrate in this section, this force exists not only in the vicinity of the BEC vertex, but more generally near any facet of $\mathrm{dom}(\mathcal{F}_p)$, which is why we will refer to it

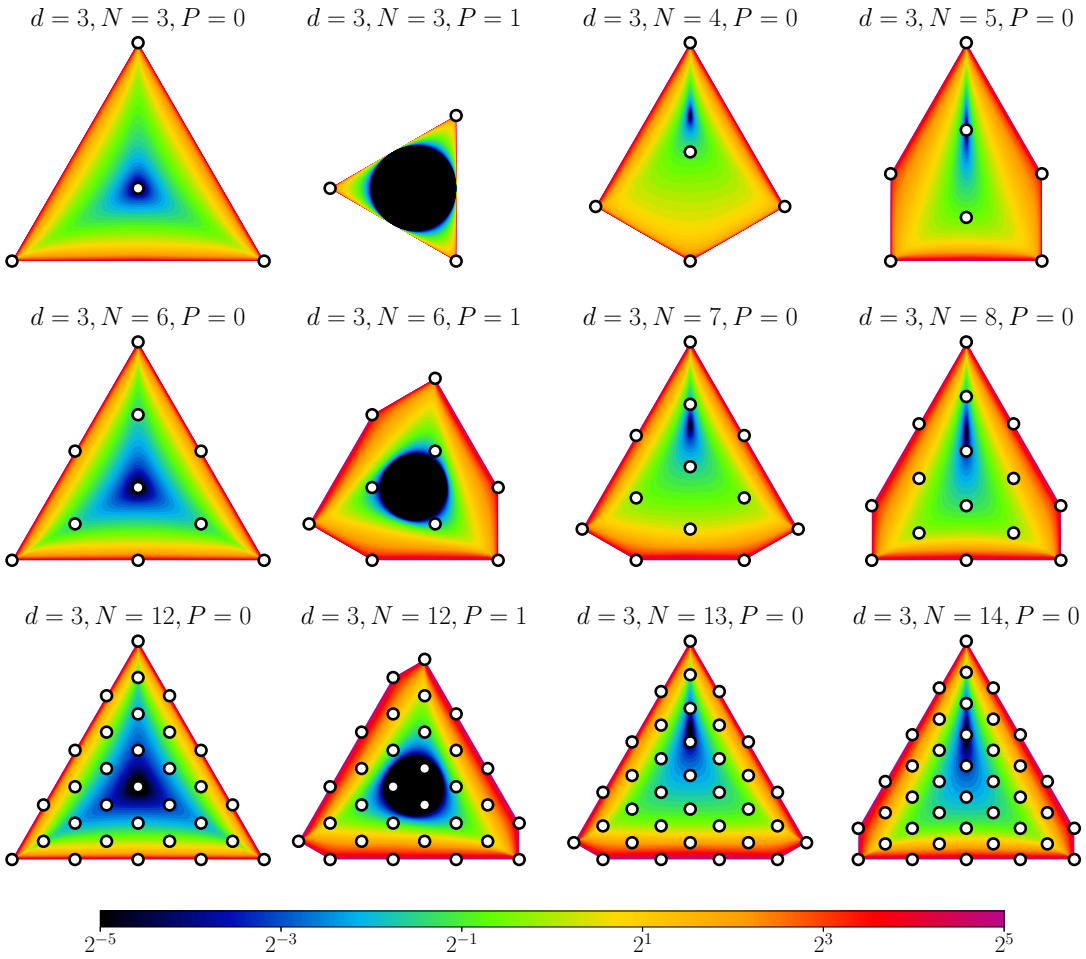


Figure 3.2 Plots of the magnitude of the derivative of the functional for various combinations of (N, P) with $d = 3$. Whenever N is not a multiple of 3, all momenta are equivalent, so only the case $P = 0$ is presented. If N is a multiple of 3, only $P = 1$ and $P = 2$ are equivalent, thus we present plots for $P = 0$ and $P = 1$. For $N = 3, 6, 12$, the functional is affine on the black disc in the center of its domain, which is a degeneracy region (see Refs. [75, 76] for more detailed discussions).

as the *generalized BEC force*, or simply *boundary force* when we move to the more general setting of abelian functional theories in Chapter 4. It is exactly the presence of such a force that prohibits the so-called *pinning*, which refers to situations where the 1RDM (or other generalized notions of density) of the ground state lies precisely on the boundary of the set of pure N-representable 1RDMs [45, 46, 100], i.e., the domain of $\mathcal{F}_p$, and is related to the question of whether natural occupation numbers of the ground state can vanish [101].

The existence of this repulsive force in various fermionic and bosonic systems has been quantitatively studied in Ref. [39] by constructing a family of suitably approximated wave functions, giving an upper bound of the pure functional $\mathcal{F}_p$, which then implies a divergence of the derivative near the boundary. The constructed wave functions also serve as a means to obtain a functional approximation. Here, in contrast, we will provide an *exact* formula for the "repulsion strength", a quantity which will be defined below, of the diverging generalized BEC force. Such a formula will not only provide a criterion for the presence or absence of pinning, but will also contribute an exact relation the functional of any translation invariant bosonic system has to satisfy, independently of the interaction W. Moreover, it will be evident from our derivation that repulsive forces from the boundary are a common feature in general functional theories, and we will keep coming back to this phenomenon throughout this thesis as we explore functional theory classes of different levels of generality.

Before proceeding, a brief note on convention: recall that in determining the facet-defining inequalities, there is an ambiguity in $S^{(j)} \in \mathbb{R}^d$ and $\nu^{(j)} \in \mathbb{R}$, as we remarked shortly after Eq. (3.8) in Section 3.4. Throughout the rest of this section, we will fix the normalization of the constraint coefficients $S^{(j)}, \nu^{(j)}$ by requiring that $S^{(j)}$ be tangent to $\operatorname{dom} \mathcal{F}_p$ and that $\|S^{(j)}\|$ be unity. If ι is injective, which is true for most (d, N, P), then $\operatorname{dom} \mathcal{F}_p$ has codimension one in $\mathbb{R}^d$, so the former condition is equivalent to imposing $\sum_\alpha S^{(j)}_\alpha = 0$. For example, the first constraint in Eq. (3.9) for $(d, N, P) = (3, 3, 0)$ would become $D^{(1)}(m) = (2m_1 - m_2 - m_3 + 3)/\sqrt{6}$, for which $S^{(1)} = \frac{1}{\sqrt{6}}(2, -1, -1)$ and $\nu^{(1)} = -\frac{3}{\sqrt{6}}$. This is merely a convenient choice that will simplify the formulas that follow.

3.5.1 The Simplex

As in Section 3.4.1, a clearer understanding can be gained by first analyzing the problem under the assumption that $\operatorname{dom}(\mathcal{F}_p)$ is a simplex with $\dim \mathcal{H}_N^{(P)}$ vertices. Although a heuristic argument for a diverging repulsive force in the simplex setting is given at the beginning of this section, here we will derive a formula for the repulsion strength (defined in Eq. (3.23)). Recall (see Section 3.4.1) that we have occupation number vectors $m^{(\alpha)}$, $\alpha = 1, \cdots, \dim \mathcal{H}_N^{(P)}$ corresponding to the permanents, with the respective constraint associated to the facet opposite $m^{(\alpha)}$ denoted as $D^{(\alpha)}(m)$. Pick any $\omega \in \{1, \cdots, \dim \mathcal{H}_N^{(P)}\}$, and let us focus on the values of the functional $\mathcal{F}_p$ in a small region near a point m_* on the facet labeled by ω, i.e., the one opposite the vertex $m^{(\omega)}$. That is to say, $D^{(\omega)}(m_*) = 0$ by assumption and we shall investigate $\mathcal{F}_p(m_* + \epsilon\eta)$ for some inward-pointing vector η. We will also assume that m_* does not simultaneously lie on another facet. In other words, $D^{(\alpha)}(m_*) > 0$ whenever $\alpha \neq \omega$.

In principle, we could carry out the calculations for an arbitrary direction η so long as it is inward-pointing. However, the outcome should not differ much: if the diverging boundary force exists, which we are yet to establish, then, intuitively, moving in any direction other than the one perpendicular to the facet should result in only minute changes in the value of

$\mathcal{F}_p$. For this reason, we will take here $\eta = S^{(\omega)}$, where $S^{(\omega)}$ consists of the coefficients of the linear part of the constraint $D^{(\omega)}$ as defined in Eq. (3.8). Then, almost by definition, the vector $S^{(\omega)}$ is the inward unit vector perpendicular to the facet ω.

We will make use of Eq. (3.11) to examine $\mathcal{F}_p(m_* + \epsilon S^{(\omega)})$. More specifically, the derivative

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(m_* + \epsilon S^{(\omega)}) \tag{3.23}$$

will be computed, and this quantity will be referred to as the *repulsion strength* of the boundary force at m_*. If one computed $\mathrm{d}\mathcal{F}_p/\mathrm{d}\epsilon$ near but not at m_* instead, one would get

$$\frac{1}{2\sqrt{\epsilon}}\left[\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(m_* + \epsilon S^{(\omega)})\right]$$

(plus other terms that are insignificant in the limit $\epsilon \to 0$). Therefore, the repulsion strength is, up to a factor of 2, the prefactor in front of $1/\sqrt{\epsilon}$ of the diverging derivative.

It is straightforward to check

$$D^{(\alpha)}(m_* + \epsilon S^{(\omega)}) = D^{(\alpha)}(m_*) + \epsilon \left\langle S^{(\alpha)}, S^{(\omega)} \right\rangle,$$

which, for $\alpha = \omega$, yields

$$D^{(\omega)}(m_* + \epsilon S^{(\omega)}) = \epsilon. \tag{3.24}$$

Of course, Eq. (3.24) is unsurprising: ϵ parameterizes the distance to m_*, whereas $D^{(\omega)}$ is the distance to the facet. Nonetheless, Eq. (3.24) confirms that $S^{(\omega)}$ is indeed inward-pointing as we had required. Now, if we plug Eq. (3.24) into Eq. (3.11), separate the terms involving ω, and perform the minimization over the phase ξ_ω, we get

$$\begin{aligned}\mathcal{F}_p(m_* + \epsilon S^{(\omega)}) = \min_{(\xi_\alpha)_{\alpha\neq\omega}} \Bigg[&\sum_{\alpha\neq\omega,\beta\neq\omega} \bar{\xi}_\alpha \xi_\beta W_{\alpha\beta} \sqrt{\frac{D^{(\alpha)}(m_* + \epsilon S^{(\omega)})}{L_\alpha}} \sqrt{\frac{D^{(\beta)}(m_* + \epsilon S^{(\omega)})}{L_\beta}} \\ &- 2\sqrt{\frac{\epsilon}{L_\omega}} \left| \sum_{\alpha\neq\omega} \xi_\alpha W_{\omega\alpha} \sqrt{\frac{D^{(\alpha)}(m_* + \epsilon S^{(\omega)})}{L_\alpha}} \right| + \frac{\epsilon}{L_\omega} W_{\omega\omega} \Bigg].\end{aligned} \tag{3.25}$$

The goal is to take the derivative with respect to $\sqrt{\epsilon}$ at $\epsilon = 0$ (Eq. (3.23)). This can be accomplished via Danskin's theorem [102], which asserts that it is valid to interchange the order of minimization and taking directional derivatives (which is what we need here; in particular, it does not imply that the pure functional $\mathcal{F}_p$ is differentiable), while restricting the range of the new minimization to the set of minimizers of the original one. That is, if we define

$$X_0 := \operatorname*{arg\,min}_{(\xi_\alpha)_{\alpha\neq\omega}} \sum_{\alpha\neq\omega,\beta\neq\omega} \bar{\xi}_\alpha \xi_\beta W_{\alpha\beta} \sqrt{\frac{D^{(\alpha)}(m_*)}{L_\alpha}} \sqrt{\frac{D^{(\beta)}(m_*)}{L_\beta}},$$

then

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(m_* + \epsilon S^{(\omega)}) = -\frac{2}{\sqrt{L_\omega}} \max_{(\xi_\alpha)_{\alpha\neq\omega}\in X_0} \left| \sum_{\alpha\neq\omega} \xi_\alpha W_{\omega\alpha} \sqrt{\frac{D^{(\alpha)}(m_*)}{L_\alpha}} \right|. \tag{3.26}$$

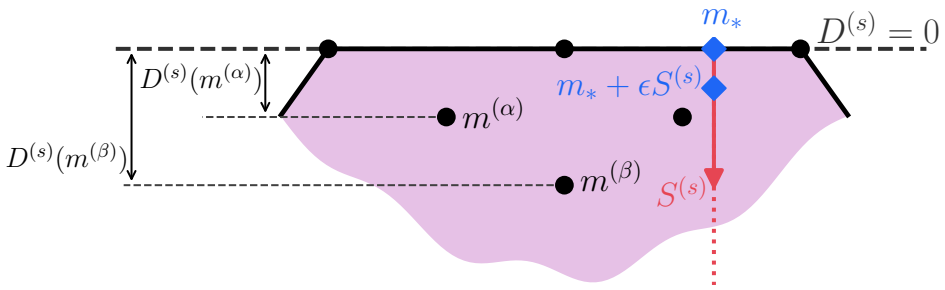


Figure 3.3 Illustration of the geometric objects involved in the analysis of the generalized BEC force. The upper facet is the facet labeled by s, i.e., the points on it satisfy $D^{(s)} = 0$. In the diagram, $m^{(\alpha)}$ and $m^{(\beta)}$ are the occupation number vectors of two permanents (arbitrarily chosen for the purpose of illustration) not lying on the facet. Their distances $D^{(s)}(m^{(\alpha)})$ and $D^{(s)}(m^{(\beta)})$ are indicated on the left. The repulsion strength is defined in terms of the values of the functional $\mathcal{F}_p$ along the straight path starting at a point m_* on the facet (blue diamond marker on top) and extending inwards perpendicularly to the facet (red dotted line), i.e., in the direction of the normal vector $S^{(s)}$ (red arrow).

Let $|\Phi\rangle$ denote the state $\sum_{\alpha\neq\omega} \xi_\alpha \sqrt{D^{(\alpha)}(m_*)/L_\alpha}\, |m^{(\alpha)}\rangle$ with the phases ξ_α chosen so that the maximum in Eq. (3.26) is achieved. Then we can write Eq. (3.26) more compactly as

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}(m_* + \epsilon S^{(\omega)}) = -2\frac{\left|\langle m^{(\omega)}|W|\Phi\rangle\right|}{\sqrt{L_\omega}}. \tag{3.27}$$

First, note that the minus sign indicates repulsion from the facet. The geometric factor $\sqrt{L_\omega}$ in the denominator implies that the repulsion strength is weaker as the distance from the vertex to the facet increases, which gives us the following physical interpretation: the farther away the vertex $m^{(\omega)}$ is from the facet $\langle m, S^{(\omega)}\rangle - \nu^{(\omega)} = 0$, the less relevant the state $|m^{(\omega)}\rangle$ is for describing the physics for densities near the facet, since having a nonzero overlap with $|m^{(\omega)}\rangle$ comes at the cost of deviating from the facet at a rate proportional to the distance $L_\omega = D^{(\omega)}(m^{(\omega)})$. (This is also why we imposed $\left\|S^{(\omega)}\right\| = 1$: so that $D^{(\omega)}(m)$ is not only proportional to the distance from m to the facet, but actually *is* the distance.)

3.5.2 Beyond Simplex

We now drop the simplex assumption and derive a formula for the repulsion strength for any (d, N, P). As is often the case, abandoning the simplex assumption generates substantial complications in establishing parallel results. First and foremost, there is no longer a one-to-one correspondence between the facets and the occupation number vectors $\{m^{(\alpha)}\}$ of the permanents, and instead we have J constraints $D^{(j)}, j = 1, \cdots, J$ and $\dim \mathcal{H}_N^{(P)}$ occupation number vectors $m^{(\alpha)}$, $\alpha = 1, \cdots, \dim \mathcal{H}_N^{(P)}$. For the notation, this means that we will use Latin letters $j, s, \cdots$ to label facets, rather than Greek letters $\alpha, \omega, \cdots$, which we reserve for labeling permanents. We proceed as before, choosing a specific $s \in \{1, \cdots, J\}$ and fixing a density m_* on the facet labeled by s, so that $D^{(s)}(m_*) = 0$. We will also assume $D^{(j)}(m_*) > 0$ whenever $j \neq s$, so that m_* does not simultaneously lie on another facet.

We will now present a physically motivated, and admittedly not completely rigorous derivation of the generalized BEC force. One subtlety (among many) that we sweep under the rug

here is that the following argument is valid only when m_* is not a critical value of the density map of the *restricted functional theory on the facet*, a notion which will be defined precisely in Chapter 4. Since the set of critical values has measure zero in the facet, the mistake we are about to make is not too serious. In Chapter 4, when we discuss abelian functional theories, which generalize bosonic momentum space functional theories, we will give a more general and strictly rigorous proof based on the constrained search.

We will adopt the notation that I_j denotes the set of all indices α for which the occupation number vector $m^{(\alpha)}$ satisfies $D^{(j)}(m^{(\alpha)}) = 0$. In other words, I_j is the set of indices of permanents lying on the facet labeled by j.

Let $|\Phi(\epsilon)\rangle$, $\epsilon \in [0, \epsilon')$ be a curve such that

$$\iota^*(|\Phi(\epsilon)\rangle\langle\Phi(\epsilon)|) = m_* + \epsilon S^{(s)}$$

for all ϵ and assume that $|\Phi(\epsilon)\rangle$ is a minimizer of the constrained search at $m_* + \epsilon S^{(s)}$ for each ϵ. That is,

$$\mathcal{F}_p(m_* + \epsilon S^{(s)}) = \langle\Phi(\epsilon)|W|\Phi(\epsilon)\rangle\,. \tag{3.28}$$

Write $|\Phi(\epsilon)\rangle = \sum_\alpha c_\alpha(\epsilon)\,|m^{(\alpha)}\rangle$. We will assume that the coefficients $c_\alpha(\epsilon)$ take the following form:

$$\begin{cases} c_\alpha(\epsilon) = p_\alpha + \epsilon q_\alpha + O(\epsilon^2) & \alpha \in I_s \\ c_\alpha(\epsilon) = \sqrt{\epsilon} r_\alpha + O(\epsilon) & \alpha \notin I_s. \end{cases} \tag{3.29}$$

Eq. (3.29) needs some justification. Since $\langle\Phi(\epsilon)|n_k|\Phi(\epsilon)\rangle$ depends linearly on each $|c_\alpha(\epsilon)|^2$, and only the "off-facet" permanents, i.e., $|m^{(\alpha)}\rangle$ with $\alpha \notin I_s$, can contribute to the deviation $\epsilon S^{(s)}$ from the facet, we must have $c_\alpha(\epsilon) \propto \sqrt{\epsilon}$ for $\alpha \notin I_s$, up to higher-order terms in ϵ. On the other hand, it is not immediately clear that the "on-facet" coefficients, that is, $c_\alpha(\epsilon)$ with $\alpha \in I_s$, should not contain terms proportional to $\sqrt{\epsilon}$. Indeed, if $c_\alpha(0) = 0$, it is conceivable that $c_\alpha(\epsilon) \sim \sqrt{\epsilon}$ even if $\alpha \in I_s$. However, the vanishing of $c_\alpha(0)$ actually implies that the state $|m^{(\alpha)}\rangle$ is decoupled from other permanents on the facet (this claim is made more precise by the "No-Mixing Lemma" (Lemma 4.9), which we prove in Chapter 4), so c_α does not contribute to the generalized BEC force.

Plugging Eq. (3.29) into Eq. (3.28) yields

$$\mathcal{F}_p(m_* + \epsilon S^{(s)}) = \sum_{\alpha,\beta \in I_s} \bar{p}_\alpha p_\beta W_{\alpha\beta} + 2\sqrt{\epsilon}\,\mathrm{Re} \sum_{\alpha \in I_s, \beta \notin I_s} \bar{p}_\alpha r_\beta W_{\alpha\beta} + O(\epsilon). \tag{3.30}$$

The coefficients p_α have to be chosen so as to minimize the first term, and the resulting minimum is clearly just $\mathcal{F}_p(m_*)$. Similarly, the r_β are fixed by minimizing the second term. Observe that there is always an obvious choice of phases of the r_β that will minimize the expression. Hence, we have

$$\frac{\mathcal{F}_p(m_* + \epsilon S^{(s)}) - \mathcal{F}_p(m_*)}{\sqrt{\epsilon}} = -2 \sum_{\beta \notin I_s} |r_\beta| \cdot \left| \sum_{\alpha \in I_s} \bar{p}_\alpha W_{\alpha\beta} \right| = -2 \sum_{\beta \notin I_s} |r_\beta| \cdot \left| \langle m^{(\beta)}|W|\Phi\rangle \right|, \tag{3.31}$$

where we have defined $|\Phi\rangle := |\Phi(0)\rangle = \sum_{\alpha \in I_s} p_\alpha\,|m^{(\alpha)}\rangle$. The $|r_\beta|$ cannot be chosen arbitrarily: the condition $\iota^*(|\Phi(\epsilon)\rangle\langle\Phi(\epsilon)|) = m_* + \epsilon S^{(s)}$ implies

$$\sum_{\alpha \in I_s} m^{(\alpha)} \left[|p_\alpha|^2 + 2\epsilon\,\mathrm{Re}(\bar{p}_\alpha q_\alpha) \right] + \epsilon \sum_{\alpha \notin I_s} m^{(\alpha)}\,|r_\alpha|^2 = m_* + \epsilon S^{(s)}$$

$$\Rightarrow 2\,\mathrm{Re}\sum_{\alpha\in I_s} m^{(\alpha)}\bar{p}_\alpha q_\alpha + \sum_{\alpha\notin I_s} m^{(\alpha)}\,|r_\alpha|^2 = S^{(s)}.$$

Take the inner product with $S^{(s)}$ and use $\langle m^{(\alpha)}, S^{(s)}\rangle = \nu^{(s)}$ for $\alpha \in I_s$ to get

$$2\nu^{(s)}\,\mathrm{Re}\sum_{\alpha\in I_s} \bar{p}_\alpha q_\alpha + \sum_{\alpha\notin I_s} |r_\alpha|^2\,\langle m^{(\alpha)}, S^{(s)}\rangle = \langle S^{(s)}, S^{(s)}\rangle = 1.$$

But $\langle\Phi(\epsilon)|\Phi(\epsilon)\rangle = 1$ implies $2\,\mathrm{Re}\sum_{\alpha\in I_s}\bar{p}_\alpha q_\alpha = -\sum_{\alpha\notin I_s}|r_\alpha|^2$, so we have

$$\sum_{\alpha\notin I_s} |r_\alpha|^2\, D^{(s)}(m^{(\alpha)}) = 1. \tag{3.32}$$

That is, the last line of Eq. (3.31) should be minimized subject to the constraint (3.32). This constrained minimization can be easily carried out and the result is

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0}\mathcal{F}_p(m_* + \epsilon S^{(s)}) = -2\left[\sum_{\alpha\notin I_s}\frac{\big|\langle m^{(\alpha)}|W|\Phi\rangle\big|^2}{D^{(s)}(m^{(\alpha)})}\right]^{\frac{1}{2}}, \tag{3.33}$$

Eq. (3.33), which presents an exact analytical relation satisfied by the functional $\mathcal{F}_p$ near any facet, is the main result of this section. It directly implies that the derivative $\mathrm{d}_m\mathcal{F}_p$ diverges as $1/\sqrt{\epsilon}$ when the density m is at distance $\epsilon \to 0$ to a facet, provided that the sum in Eq. (3.33) does not vanish. We would like to highlight a number of features exhibited by this formula:

(1) The sign is negative, indicating a repulsive force from the boundary as in Section 3.5.1.

(2) The contribution of each permanent $|m^{(\alpha)}\rangle$ is inversely weighted by the geometric factor $D^{(s)}(m^{(\alpha)})$ (see Fig. 3.3 for illustration), which is the distance from $m^{(\alpha)}$ to the facet. In other words, distant permanents have smaller effects on the strength of the repulsive force.

(3) The radicand is a sum of nonnegative terms. This means that the only way the repulsion strength $\mathrm{d}\mathcal{F}_p(m_* + \epsilon S^{(j)})/\mathrm{d}\sqrt{\epsilon}|_{\epsilon=0}$ can vanish is when $\langle m^{(\alpha)}|W|\Phi\rangle = 0$ for all α such that $D^{(s)}(m^{(\alpha)}) > 0$.

(4) To see how Eq. (3.33) reduces to Eq. (3.27) in the simplex setting, note that, in the latter situation, there is only one permanent $|m^{(\alpha)}\rangle$ that is not on the facet, so the sum contains only one term and we recover Eq. (3.27).

Example 3.5. Let us illustrate Eq. (3.33) by considering the Bose-Hubbard model (Eq. (3.4)) with $(d, N, P) = (3, N, 0)$, where N is a multiple of 3 (see Fig. 3.2 for $N = 6$ and $N = 12$). The domain is a 2-simplex (triangle), but this does not belong to what we call the "simplex setting" since there are occupation number vectors $m^{(\alpha)}$ which are not extreme points. We will choose s to be (the label of) the facet corresponding to the bottom edge (see Figs. 3.1 and 3.2), for which the inequality constraint is given by

$$D^{(s)}(m) = \frac{1}{\sqrt{6}}(2m_0 - m_1 - m_2) + \frac{N}{\sqrt{6}} \geq 0, \tag{3.34}$$

so $S^{(s)} = \frac{1}{\sqrt{6}}(2,-1,-1)$ and $\nu^{(s)} = -N/\sqrt{6}$. Take $m_* = (0, Nt, N(1-t))$ for some $t \in (0,1)$, then $D^{(s)}(m_*) = 0$, so m_* does lie on the facet labeled by s. To apply Eq. (3.33), first note that the permanents on the facet are

$$|0,N,0\rangle, |0,N-3,3\rangle, \cdots, |0,0,N\rangle .$$

Since W does not couple the states on the facet (true for any pair interaction) and $|0,N,0\rangle$ and $|0,0,N\rangle$ have the lowest energy among all permanents on the facet (special to the Hubbard model), the minimizer states of the constrained search at m_* are

$$|\Phi^*(\theta)\rangle = \sqrt{t}\,|0,N,0\rangle + \sqrt{1-t}e^{i\theta}\,|0,0,N\rangle . \tag{3.35}$$

The states $|0,N,0\rangle$ and $|0,0,N\rangle$ only couple to $|1,N-2,1\rangle$ and $|1,1,N-2\rangle$ respectively, with matrix elements

$$\langle 1,N-2,1|W|0,N,0\rangle = \langle 1,1,N-2|W|0,0,N\rangle = \frac{2}{3}\sqrt{N(N-1)}.$$

Moreover, their distances to the facet are equal and given by

$$D^{(s)}(1,N-2,1) = D^{(s)}(1,1,N-2) = \sqrt{6}/2.$$

Assuming $N > 3$ so that $|1,N-2,1\rangle$ and $|1,1,N-2\rangle$ do not coincide, Eq. (3.33) becomes

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0}\mathcal{F}_p(m_*+\epsilon S^{(j)}) &= -2\left[\frac{\left(\frac{2}{3}\cdot\sqrt{N(N-1)}\right)^2(\sqrt{t}^2+\sqrt{1-t}^2)}{\frac{\sqrt{6}}{2}}\right]^{\frac{1}{2}} \\ &= -\frac{4\cdot 2^{\frac{1}{4}}\cdot 3^{\frac{3}{4}}}{9}\sqrt{N(N-1)}\end{aligned} \tag{3.36}$$

The numerical value of the prefactor is about 1.205. Interestingly, the repulsion strength does not depend on the position t on the facet in this case.

In order to verify that Eq. (3.36) gives the correct generalized BEC force, we approximate the functional by

$$\mathcal{F}_p(m_*+\epsilon S^{(s)}) \approx \mathcal{F}_p(m_*) + \sqrt{\epsilon}\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0}\mathcal{F}_p(m_*+\epsilon S^{(s)})$$

and compare the approximation to numerical values of $\mathcal{F}_p$. We will choose $m_* = (0, \frac{N}{2}, \frac{N}{2})$, for which $\mathcal{F}_p(m_*) = N(N-1)/3$. It follows that

$$\frac{\mathcal{F}_p\left(m_*+\epsilon S^{(s)}\right)}{N^2} \approx \frac{N-1}{3N} - 1.205\sqrt{\epsilon}\frac{1}{N}\sqrt{\frac{N-1}{N}}. \tag{3.37}$$

In Fig. 3.4, this is compared to the exact functional for $N = 6, 18, 30$. Clearly, Eq. (3.37) yields the correct generalized BEC force at $m_* = (0, \frac{N}{2}, \frac{N}{2})$.

Finally, we would like to remark that Eq. (3.33) also accounts for the larger repulsion strength for $P = 1$ compared to $P = 0$ at fixed $(d = 3, N)$ when N is a multiple of 3, as can be observed in Fig. 3.2. When $P = 1$, it is still true that a minimizing state on the facet is always a superposition of the two extremal permanents (see Fig. 3.5). However, these couple to five off-facet permanents in total, as opposed to two in the case of $P = 0$. As a result, the sum in Eq. (3.33) contains five terms when $P = 1$ instead of two, resulting in a larger repulsion strength. This difference is illustrated in Fig. 3.5 for $N = 12$.

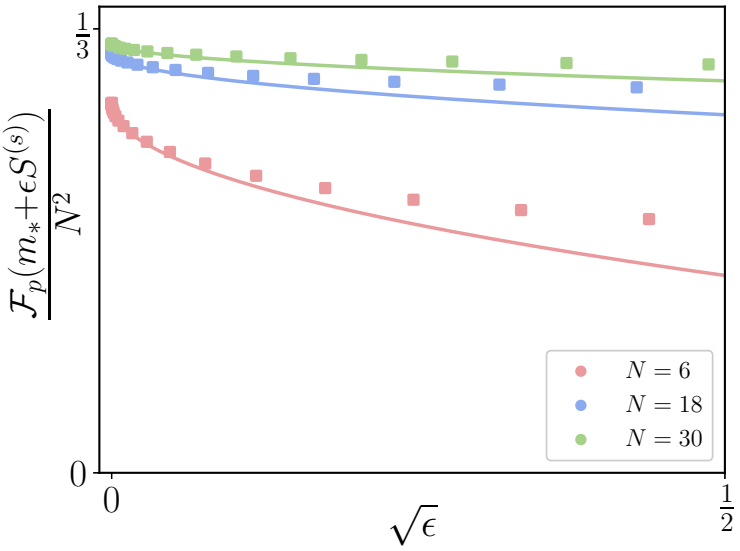


Figure 3.4 Comparison of the approximate functional given by Eq. (3.37) (solid lines) for the Hubbard model with exact numerical values (square dots).

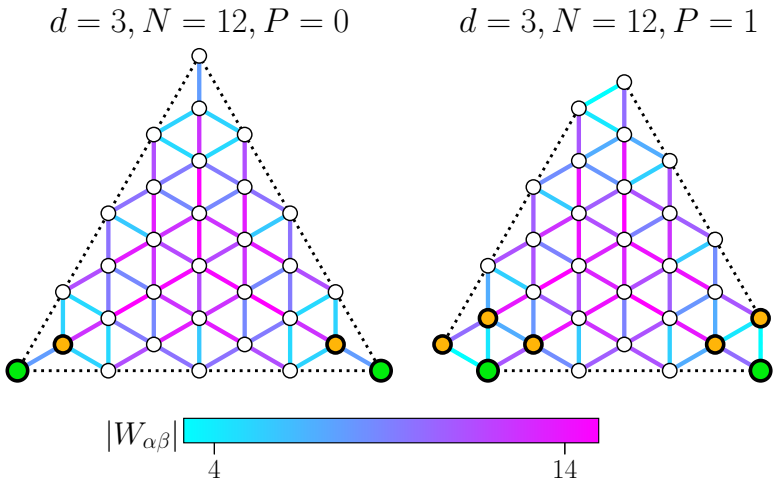


Figure 3.5 Graphical depiction of the underlying cause of the stronger generalized BEC force for $P = 1$ (see the two bottom-left plots in Fig. 3.2). In each diagram, two occupation number vectors $m^{(\alpha)}, m^{(\beta)}$ are connected by an edge if $W_{\alpha\beta} \neq 0$, with the value of $|W_{\alpha\beta}|$ encoded by the color. On a given facet, the extremal permanents (green) couple to two off-facet states (yellow) for $P = 0$ and to five off-facet states for $P = 1$.

Chapter 4

Abelian Functional Theory

The bosonic momentum space functional theory presented in Chapter 3 is a special case of an *abelian* functional theory. There, we consider the family of one-particle operators h which are diagonal in the momentum basis, which all commute and consequently allow us to determine the domain of the functional with relative ease. In this chapter, we single out a class of generalized functional theories which have this commutativity property and investigate its consequences. An important result of this chapter is the *boundary force formula* (Theorem 4.8), which we state and prove in Section 4.3.

4.1 Definitions

Definition 4.1. *A generalized functional theory $(V, \mathcal{H}, \iota, W)$ is **abelian** if the image of ι is commutative. In other words, $[\iota(v), \iota(v')] = 0$ for any $v, v' \in V$.*

Note that this definition is independent of the interaction W. That is, we do *not* require W to commute with the external potentials.

Example 4.1. The discrete DFT in Example 2.1 is abelian because the number operators $a_i^\dagger a_i$ $(i = 1, \cdots, d)$ commute. RDMFT (Example 2.2), on the other hand, is not abelian (unless $\dim \mathcal{H}_1 = 1$) because the set of all one-particle operators $i\mathfrak{u}(\mathcal{H}_1)$, or more precisely its image in $i\mathfrak{u}(\mathcal{H})$, where $\mathcal{H}$ is the N-particle Hilbert space, is not commutative.

Example 4.2. Consider a system of N two-level spins, whose Hilbert space is $\mathcal{H} = (\mathbb{C}^2)^{\otimes N}$. As the space of external potentials, we take $V = \mathbb{R}^N$ together with potential map $\iota(v_1, \cdots, v_N) = \sum_{i=1}^N v_i Z_i$, where Z_i is the Pauli Z operator of the i-th spin. In contrast to Example 2.3, the functional theory defined by $(V, \mathcal{H}, \iota, W)$ is abelian for any interaction W because the Z_i commute.

An (if not *the most*) important feature of abelian functional theories is that all external potentials can be simultaneously diagonalized, since they all commute by assumption.

Definition 4.2. *A density $\alpha \in V^*$ is a **weight** if there exists a nonzero vector $|\Psi\rangle \in \mathcal{H}$ such that*

$$\forall v \in V : \iota(v) |\Psi\rangle = \langle \alpha, v \rangle |\Psi\rangle .$$

Given any $\alpha \in V^*$, the corresponding ***weight space*** *is*

$$\mathcal{H}_\alpha := \{|\Psi\rangle \in \mathcal{H} \mid \forall v \in V : \iota(v)\,|\Psi\rangle = \langle \alpha, v\rangle\,|\Psi\rangle\}.$$

If α is a weight, its ***multiplicity*** *is defined as the dimension of $\mathcal{H}_\alpha$. A nonzero vector in $\mathcal{H}_\alpha$ is called a* ***weight vector****.*

Hence, a density α is a weight if and only if $\mathcal{H}_\alpha \neq 0$. It is an elementary fact that the Hilbert space decomposes into an orthogonal direct sum of the weight spaces:

$$\mathcal{H} = \bigoplus_{\alpha \in \Omega} \mathcal{H}_\alpha, \tag{4.1}$$

where Ω denotes the set of weights.

The importance of the weights is that they determine the domains of the pure and ensemble functionals according to the following theorem, which is the generalization of Theorem 3.1:

Theorem 4.1. *The set of pure state representable densities coincides with the set of ensemble state representable densities, and both are given by the convex hull of the set of weights. In other words,*

$$\iota^*(\mathcal{P}) = \iota^*(\mathcal{E}) = \mathrm{conv}(\Omega). \tag{4.2}$$

Proof. We will show

$$\iota^*(\mathcal{P}) \overset{(1)}{\subset} \iota^*(\mathcal{E}) \overset{(2)}{\subset} \mathrm{conv}(\Omega) \overset{(3)}{\subset} \iota^*(\mathcal{P}).$$

Inclusion (1) follows directly from $\mathcal{P} \subset \mathcal{E}$ (all pure states are ensemble states).

Let $(|E_i\rangle)_{i\in I}$ be an orthonormal basis compatible with the weight space decomposition (4.1), where I is some index set. That is, for each $i \in I$, there exists a weight $\omega_i \in \Omega$ such that the basis vector $|E_i\rangle$ is in the weight space $\mathcal{H}_{\omega_i}$. Take any ensemble state representable density $\rho \in \iota^*(\mathcal{E})$. Then, by definition,

$$\rho = \iota^* \left(\sum_{i,j\in I} c_{ij}\, |E_i\rangle\langle E_j| \right),$$

where the matrix $(c_{ij})_{ij}$ is positive semidefinite with unit trace. Hence, for any $v \in V$,

$$\begin{aligned}
\langle \rho, v\rangle &= \sum_{i,j\in I} c_{ij}\, \langle E_j|\iota(v)|E_i\rangle = \sum_{i,j\in I} c_{ij}\, \langle \omega_i, v\rangle\, \langle E_j|E_i\rangle \\
&= \sum_{i,j\in I} c_{ij}\, \langle \omega_i, v\rangle\, \delta_{ij} = \sum_{i\in I} c_{ii}\, \langle \omega_i, v\rangle \\
&= \left\langle \sum_{i\in I} c_{ii}\omega_i, v \right\rangle,
\end{aligned}$$

implying $\rho = \sum_{i\in I} c_{ii}\omega_i$. Since $c_{ii} \geq 0$ and $\sum_i c_{ii} = 1$, we conclude that ρ is a convex combination of the weights. This shows inclusion (2).

Take any $\rho = \sum_{\omega\in\Omega} t_\omega \omega$ with $\sum_{\omega\in\Omega} t_\omega = 1$ and $t_\omega \geq 0$. For each weight $\omega \in \Omega$, pick any weight vector $|\Phi_\omega\rangle \in \mathcal{H}_\omega$ with $\||\Phi_\omega\|| = 1$. Define $|\Psi\rangle = \sum_{\omega\in\Omega} \sqrt{t_\omega}\,|\Phi_\omega\rangle$. Then, for any $v \in V$,

$$\begin{aligned}\langle \iota^*(|\Psi\rangle\langle\Psi|), v\rangle &= \langle\Psi|\iota(v)|\Psi\rangle = \sum_{\omega,\omega'\in\Omega} \sqrt{t_{\omega'}t_\omega}\,\langle\Psi_{\omega'}|\iota(v)|\Psi_\omega\rangle \\ &= \sum_{\omega,\omega'\in\Omega} \sqrt{t_{\omega'}t_\omega}\,\langle\omega, v\rangle\,\delta_{\omega'\omega} = \sum_{\omega\in\Omega} t_\omega\,\langle\omega, v\rangle \\ &= \langle\rho, v\rangle\,,\end{aligned}$$

so $\iota^*(|\Psi\rangle\langle\Psi|) = \rho$. This shows inclusion (3). □

Remark 4.1. At least when the connected Lie group corresponding to $\iota(V) \subset i\mathfrak{u}(\mathcal{H})$ is compact, i.e., a torus, Theorem 4.1 is a corollary of *Atiyah's theorem* [103] (also due to Guillemin and Sternberg [104]), which will be introduced in Chapter 5. What Atiyah, Guillemin, and Sternberg proved is a much more general statement about Hamiltonian torus actions on symplectic manifolds. In our case, the symplectic manifold is simply the projective Hilbert space (which can be identified with $\mathcal{P}$), for which a simple proof like the one presented above works. We will never need the full power of Atiyah's theorem in this thesis.

In light of Theorem 4.1, we will simply call a density $\rho \in V^*$ *representable* if it is pure state representable or ensemble state representable, since $\iota^*(\mathcal{P}) = \iota^*(\mathcal{E})$. Equivalently, a density $\rho \in V^*$ is representable if and only if it belongs to $\mathrm{conv}(\Omega)$. The set $\mathrm{conv}(\Omega)$, being the convex hull of finitely many points, is a convex polytope. Clearly, the dimension of $\mathrm{conv}(\Omega)$ is equal to the dimension of $\mathrm{aff}(\iota^*(\mathcal{P}))$, which is in turn equal to $\dim V - \dim \iota^{-1}(\mathrm{span}\{\mathbb{1}\})$ by Proposition 2.3. In particular, if the potential map ι is injective and its image does not contain $\mathbb{1}$, then the convex polytope $\mathrm{conv}(\Omega)$ will have full dimension in V^*. In general, $\mathrm{conv}(\Omega)$ lies in a hyperplane of codimension $\dim(\iota^{-1}(\mathrm{span}\{\mathbb{1}\}))$. We summarize this discussion in the following proposition.

Proposition 4.2. *The set of representable densities* $\iota^*(\mathcal{P}) = \iota^*(\mathcal{E}) = \mathrm{conv}(\Omega)$ *is a convex polytope of dimension* $\dim V - \dim \iota^{-1}(\mathrm{span}\{\mathbb{1}\})$.

Example 4.3. Consider again the abelian functional theory of N two-level spins in Example 4.2. Since $\iota(v) = \sum_{i=1}^N v_i Z_i$, the weight vectors are exactly the simultaneous eigenvectors of the Pauli Z operators, which are product states of $|\uparrow\rangle$ and $|\downarrow\rangle$. The corresponding weights are then $\Omega = \{-1, 1\}^N \subset \mathbb{R}^N = (\mathbb{R}^N)^*$. For each weight $\alpha = (\alpha_1, \cdots, \alpha_N)$, $\alpha_i = \pm 1$, the weight space $\mathcal{H}_\alpha$ is one-dimensional, so there is, up to scalars, a unique weight vector for α. The set of representable densities is then $\iota^*(\mathcal{P}) = \iota^*(\mathcal{E}) = \mathrm{conv}(\Omega) = \{(\rho_1, \rho_2, \cdots, \rho_N) \mid \rho_i \in [-1, 1]\} = [-1, 1]^N$, which is an N-dimensional hypercube (see Fig. 4.1 for $N = 3$ spins).

Example 4.4. For the bosonic momentum space functional theory introduced in Chapter 3 (see Definition 3.1), the weight vectors are the permanents $|m^{(\alpha)}\rangle$ with momentum P, and the weights are the occupation number vectors $m^{(\alpha)}$ of the permanents.

Now that we have established that $\iota^*(\mathcal{P}) = \iota^*(\mathcal{E}) = \mathrm{conv}(\Omega)$ is a convex polytope, we can talk about its facets, which are by definition the codimension-1 faces of $\mathrm{conv}(\Omega)$. In particular, we will define the notion of the *restricted functional theory* corresponding to each facet of $\mathrm{conv}(\Omega)$. This will turn out to be useful when we discuss boundary forces in subsequent sections, where properties of the restricted functional theory will play an important role in the derivation of the boundary force formula.

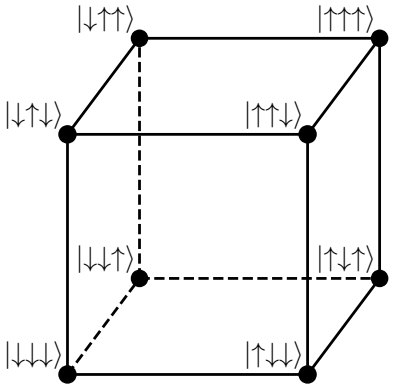


Figure 4.1 The weight structure for the abelian spin chain functional theory described in Example 4.3 with $N = 3$ two-level spins. For three spins, there are 8 possible product states which are simultaneous eigenstates of the Pauli Z operators, and they correspond to eight weights, given by $(\pm 1, \pm 1, \pm 1)$ in $(\mathbb{R}^3)^* = \mathbb{R}^3$, forming vertices of a cube of side length 2. According to Theorem 4.1, we have $\operatorname{dom} \mathcal{F}_p = \iota^*(\mathcal{P}) = \operatorname{dom} \mathcal{F}_e = \iota^*(\mathcal{E}) = [-1, 1]^3$.

Definition 4.3. *Let $(V, \mathcal{H}, \iota, W)$ be an abelian functional theory and let F be a facet of the convex polytope* $\operatorname{conv}(\Omega)$. *Let $\Omega_F := \Omega \cap F \subset \Omega$ be the subset of weights lying on F. Consider the subspace $\mathcal{H}_F = \bigoplus_{\omega \in \Omega_F} \mathcal{H}_\omega \subset \mathcal{H}$ and the corresponding orthogonal projection $\Pi_F : \mathcal{H} \to \mathcal{H}_F$. The* ***restricted functional theory on the facet*** *F or simply the* ***facet functional theory on*** *F is the generalized functional theory given by $(V, \mathcal{H}_F, \iota_F, \Pi_F W \Pi_F^\dagger)$, where $\iota_F(v) = \Pi_F \iota(v) \Pi_F^\dagger$.*

Note that we do not replace the vector space V that parametrizes the space of external potentials. This is mostly a matter of convention: intuitively, elements in V which are perpendicular to F will be mapped to something proportional to $\mathbb{1}_F$ under ι_F, where $\mathbb{1}_F$ denotes the identity on $\mathcal{H}_F$, which is fine, since we have always allowed the possibility for the image of the potential map to contain the identity. In fact, this is consistent with Proposition 4.2 as the dimension of the facet F is precisely one lower than that of $\operatorname{conv}(\Omega)$.

Proposition 4.3 (Basic properties of restricted functional theories)**.** *Let $F \subset \operatorname{conv}(\Omega)$ be any facet. Then:*

(1) The restricted functional theory on F is abelian.

(2) $\Omega_F := \Omega \cap F$ is the set of weights of the restricted functional theory

(3) $\dim(\iota_F^{-1}\operatorname{span}\{\mathbb{1}_F\}) = \dim(\iota^{-1}\operatorname{span}\{\mathbb{1}\}) + 1$.

(4) For all $\Gamma \in \mathcal{P}$, $\iota^*(\Gamma) \in F$ if and only if $\Gamma \in \mathcal{P}_F$, where $\mathcal{P}_F$ denotes the set of pure states on $\mathcal{H}_F$.

(5) For all $\Gamma \in \mathcal{E}$, $\iota^*(\Gamma) \in F$ if and only if $\Gamma \in \mathcal{E}_F$, where $\mathcal{E}_F$ denotes the set of ensemble states on $\mathcal{H}_F$.

(6) $\mathcal{F}_p^{(F)} = \mathcal{F}_p|_F$, where $\mathcal{F}_p^{(F)}$ denotes the pure functional of the restricted functional theory.

(7) $\mathcal{F}_e^{(F)} = \mathcal{F}_e|_F$, where $\mathcal{F}_e^{(F)}$ denotes the ensemble functional of the restricted functional theory.

Proof. (1) and (2) are obvious. (3) follows from Proposition 4.2 and the fact that $\dim \mathrm{conv}(\Omega_F) = \dim F = \dim \mathrm{conv}(\Omega) - 1$.

Take a pure state $|\Psi\rangle\langle\Psi| \in \mathcal{P}$ such that $\iota^*(|\Psi\rangle\langle\Psi|) \in F$. Let $|\Psi\rangle = \sum_{\omega\in\Omega} |\Psi_\omega\rangle$ be the weight space decomposition of $|\Psi\rangle$, with $|\Psi_\omega\rangle \in \mathcal{H}_\omega$. Then

$$\langle \iota^*(|\Psi\rangle\langle\Psi|), v\rangle = \sum_{\omega\in\Omega} \langle\Psi_\omega|\iota(v)|\Psi_\omega\rangle = \sum_{\omega\in\Omega} \langle\omega, v\rangle \, \|\Psi_\omega\|^2$$

for all $v \in V$. Hence, $\iota^*(|\Psi\rangle\langle\Psi|) = \sum_{\omega\in\Omega} \|\Psi_\omega\|^2 \omega$. But this convex combination has to lie on the facet F, so $|\Psi_\omega\rangle$ must vanish if $\omega \notin F$, which implies $|\Psi\rangle \in \mathcal{H}_F$. This shows (4).

Let $\Gamma = \sum_i t_i |\Psi_i\rangle\langle\Psi_i| \in \mathcal{E}$ be any ensemble state such that $\iota^*(\Gamma) \in F$. Then by linearity of ι^*, we have $\sum_i t_i \iota^*(|\Psi_i\rangle\langle\Psi_i|) \in F$, so $\iota^*(|\Psi_i\rangle\langle\Psi_i|) \in F$ for each i. (4) then implies $|\Psi_i\rangle \in \mathcal{H}_F$, which shows (5).

By definition, for $\rho \in F$,

$$\mathcal{F}_p^{(F)}(\rho) = \min\Big\{ \langle\Psi|W|\Psi\rangle \mid |\Psi\rangle \in \mathcal{H}_F, \|\Psi\| = 1, \iota^*(|\Psi\rangle\langle\Psi|) = \rho \Big\}$$
$$\mathcal{F}_p|_F(\rho) = \min\Big\{ \langle\Psi|W|\Psi\rangle \mid |\Psi\rangle \in \mathcal{H}, \|\Psi\| = 1, \iota^*(|\Psi\rangle\langle\Psi|) = \rho \Big\}.$$

By (4), we can replace the condition $|\Psi\rangle \in \mathcal{H}$ in the second line with $|\Psi\rangle \in \mathcal{H}_F$. This shows (6). Similarly, (7) follows from writing down the definitions of $\mathcal{F}_e^{(F)}$ and $\mathcal{F}_e|_F$ and applying (5). □

Example 4.5. Consider again the abelian spin functional theory constructed in Example 4.3. For each $j \in \{1, 2, \cdots, N\}$, a pair of facets of $\mathrm{conv}(\Omega) = [-1, 1]^N$ are given by $F_j^\pm := \{\rho \in [-1, 1]^N \mid \rho_j = \pm 1\}$, and every facet of $\mathrm{conv}(\Omega)$ is of this form. For each (j, q), where $q = \pm 1$ is a sign, the facet Hilbert space is $\mathcal{H}_{F_j^q} = \{|\Psi\rangle \in \mathcal{H} \mid Z_j |\Psi\rangle = q |\Psi\rangle\}$, the q-eigenspace of Z_j. It is easy to check that $e_j \in V$, the vector whose j-th entry is one and all other entries are zero, satisfies $\iota_{F_j^q}(e_j) = q \mathbb{1}_{F_j^q}$.

Finally, we introduce the notion of *strong orthogonality* of two vectors in $\mathcal{H}$, which will be relevant for the proof of the boundary force in Section 4.3. For any vector $|\Psi\rangle \in \mathcal{H}$, we will denote its component in a weight space $\mathcal{H}_\omega$ by $|\Psi_\omega\rangle$.

Definition 4.4. *Two vectors $|\Phi\rangle, |\Phi'\rangle$ are **strongly orthogonal** if their components are orthogonal in each weight space. That is, $\langle\Phi_\omega|\Phi'_\omega\rangle = 0$ for all $\omega \in \Omega$.*

Lemma 4.4. *For normalized and strongly orthogonal states $|\Phi\rangle$ and $|\Phi'\rangle$, define $|\Psi\rangle = x|\Phi\rangle + y|\Phi'\rangle$, where $|x|^2 + |y|^2 = 1$. Then*

$$\iota^*(|\Psi\rangle\langle\Psi|) = |x|^2 \iota^*(|\Phi\rangle\langle\Phi|) + |y|^2 \iota^*(|\Phi'\rangle\langle\Phi'|). \tag{4.3}$$

Proof.

$$\begin{aligned}
\langle \iota^*(|\Psi\rangle\langle\Psi|), v\rangle &= \sum_{\alpha,\beta\in\Omega} (\bar{x}\langle\Phi_\alpha| + \bar{y}\langle\Phi'_\alpha|)\iota(v)(x|\Phi_\beta\rangle + y|\Phi'_\beta\rangle) \\
&= |x|^2 \langle\Phi|\iota(v)|\Phi\rangle + |y|^2 \langle\Phi'|\iota(v)|\Phi'\rangle + 2\,\mathrm{Re}\Big[\bar{x}y \sum_{\alpha,\beta\in\Omega} \langle\beta, v\rangle \underbrace{\langle\Phi_\alpha|\Phi'_\beta\rangle}_{=0}\Big] \\
&= \big\langle |x|^2 \iota^*(|\Phi\rangle\langle\Phi|) + |y|^2 \iota^*(|\Phi'\rangle\langle\Phi'|), v\big\rangle
\end{aligned}$$

□

4.2 Unique v-representability

In the abelian setting, slightly more can be said about unique v-representability than what we proved in Section 2.3 for any general functional theory. Recall that a density $\rho \in V^*$ is said to be pure state v-representable if there exists an external potential $v \in V$ such that $\iota(v) + W$ has ρ as the density of a pure ground state (Definition 2.3). It is uniquely v-representable if any two such external potentials v, v' satisfy $\iota(v) - \iota(v') \propto \mathbb{1}$ (Definition 2.4). In Section 2.3, after proving the weak Hohenberg-Kohn theorem, we showed that its original version, which is stronger, is equivalent to the statement that all pure state v-representable densities are uniquely so. We then gave a simple counterexample (Example 2.6), which led to the conclusion that the strong HK theorem is false in general. However, we were able to provide a partial remedy to this failure. Namely, Theorem 2.7 states that the strong HK theorem can only fail at the critical values of the map $\iota^*|_{\mathcal{P}} : \mathcal{P} \to \mathrm{aff}(\iota^*(\mathcal{P}))$, which implies, through Sard's theorem, that the failure of the strong HK theorem only occurs on a set of measure zero (relative to $\mathrm{aff}(\iota^*(\mathcal{P}))$). When a functional theory is abelian, there is a satisfactory characterization of the set of critical values, which we will give now.

For any vector $|\Psi\rangle \in \mathcal{H}$, we denote by $\Omega_\Psi \subset \Omega$ the *support* of $|\Psi\rangle$, which is by definition the set of weights ω such that $|\Psi\rangle$ has nonzero component in the weight space $\mathcal{H}_\omega$. In other words, if $|\Psi\rangle = \sum_{\omega\in\Omega} |\Psi_\omega\rangle$ is the weight space decomposition, then $\Omega_\Psi = \{\omega \in \Omega \mid |\Psi_\omega\rangle \neq 0\}$. Clearly, $\iota^*(|\Psi\rangle\langle\Psi|) \in \mathrm{conv}(\Omega_\Psi)$.

Lemma 4.5. *For any pure state* $|\Psi\rangle\langle\Psi| \in \mathcal{P}$,

$$\mathrm{im}(D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}) = \overrightarrow{\mathrm{aff}(\Omega_\Psi)}.$$

Proof. Let $|\Psi\rangle = \sum_{\omega\in\Omega_\Psi} |\Psi_\omega\rangle$ be the weight space decomposition. Let $|\phi\rangle \in T_{|\Psi\rangle\langle\Psi|}\mathcal{P} \cong |\Psi\rangle^\perp$ be any tangent vector. Then by Lemma 2.4, we have

$$\big\langle D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}(|\phi\rangle), v\big\rangle = \sum_{\omega\in\Omega_\Psi} \langle\omega, v\rangle\, 2\,\mathrm{Re}\,\langle\phi|\Psi_\omega\rangle$$

for all $v \in V$. It follows that

$$D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}(|\phi\rangle) = 2\sum_{\omega\in\Omega_\Psi} \mathrm{Re}\,\langle\phi|\Psi_\omega\rangle\,\omega.$$

This lies in $\overrightarrow{\mathrm{aff}(\Omega_\Psi)}$ because $\sum_{\omega\in\Omega_\Psi}\langle\phi|\Psi_\omega\rangle = \langle\phi|\Psi\rangle = 0$. Since $|\phi\rangle$ was an arbitrary tangent vector, this shows "$\subset$".

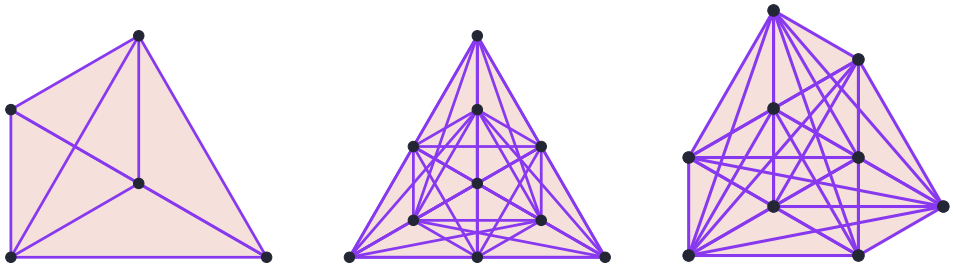
Conversely, any $\alpha \in \overrightarrow{\mathrm{aff}(\Omega_\Psi)}$ can be written as a linear combination of vectors of the form $\omega - \omega'$, where $\omega, \omega' \in \Omega_\Psi$. For any such pair, take

$$|\phi\rangle = \frac{|\Psi_\omega\rangle}{\langle\Psi_\omega|\Psi_\omega\rangle} - \frac{|\Psi_{\omega'}\rangle}{\langle\Psi_{\omega'}|\Psi_{\omega'}\rangle}.$$

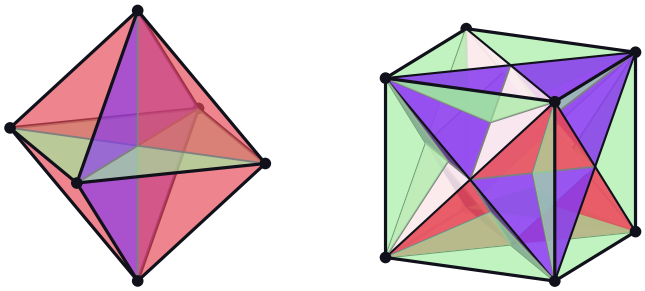
It is straightforward to verify that $\langle\phi|\Psi\rangle = 0$, so $|\phi\rangle$ is a tangent vector to $\mathcal{P}$ at $|\Psi\rangle\langle\Psi|$. We have

$$D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}(|\phi\rangle) = 2(\omega - \omega').$$

Since the derivative is linear, this shows "$\supset$". □



(a) 4 bosons on 3 sites, $P = 1$ **(b)** 6 bosons on 3 sites, $P = 0$ **(c)** 6 bosons on 3 sites, $P = 1$



(d) 2 fermions on 4 sites **(e)** 3 qubits

Figure 4.2 Critical values of several abelian functional theories. (a)–(c): translation invariant bosonic functional theories discussed in Chapter 3, with $d = 3$ sites. The critical values are the points lying in the convex hull of two weights, indicated in purple. The weight structure of (d) is that of two (spinless) fermions placed on four lattice sites with local external potentials. The six vertices correspond to the six Slater determinants $|1,1,0,0\rangle, |1,0,1,0\rangle, \cdots, |0,0,1,1\rangle$. In (e), we consider the abelian functional theory of three two-level spins (Example 4.2). The set of weights is $\{-1,1\}^3$, which forms the vertices of a cube (see also Fig. 4.1). Since $\dim \Omega = 3$ for both (d) and (e), the critical values are the points lying in planes spanned by the weights. In both cases, the boundary $\partial\mathrm{conv}(\Omega)$ (points on which are always critical values) is omitted for ease of visualization. The set of critical values of (d) (minus the boundary) is the union of three squares, shown in purple, red, and green. For (e), it is the union of two hollow tetrahedra (red and purple) and six rectangles (green).

Proposition 4.6. *Let $(V, \mathcal{H}, \iota, W)$ be an abelian functional theory. Let $\rho \in \mathrm{conv}(\Omega)$ be a representable density. Then ρ is a critical value of the map $\iota^*|_{\mathcal{P}} : \mathcal{P} \to \mathrm{aff}(\Omega)$ if and only if ρ lies in the convex hull of $\dim \mathrm{conv}(\Omega) = \dim V - \dim \iota^{-1}(\mathrm{span}\{\mathbb{1}\})$ weights. (See also Lemma 3.3 of Ref. [105].)*

Proof. If ρ is a critical value, then there is a pure state $|\Psi\rangle\langle\Psi|$ such that $D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}$ is not surjective onto $\overrightarrow{\mathrm{aff}(\Omega)}$. By Lemma 4.5, we have $\overrightarrow{\mathrm{aff}(\Omega_\Psi)} \subsetneq \overrightarrow{\mathrm{aff}(\Omega)}$. It follows that $\dim \mathrm{conv}(\Omega_\Psi) < \dim \mathrm{conv}(\Omega)$. Since $\rho \in \mathrm{conv}(\Omega_\Psi)$, it is a convex combination of $\dim \mathrm{conv}(\Omega_\Psi) + 1 \leq \dim \mathrm{conv}(\Omega)$ weights by Carathéodory's theorem.

Conversely, suppose ρ lies in the convex hull of $\dim \mathrm{conv}(\Omega)$ weights. In other words, there is a subset $\Omega' \subset \Omega$ and coefficients $t_\omega \geq 0$ such that $\sum_{\omega\in\Omega'} t_\omega \omega = \rho$, $\sum_{\omega\in\Omega'} t_\omega = 1$, and $|\Omega'| \leq \dim \mathrm{conv}(\Omega)$. Consider the state $|\Psi\rangle := \sum_{\omega\in\Omega'} |\Psi_\omega\rangle$, where the $|\Psi_\omega\rangle$ satisfy $|\Psi_\omega\rangle \in \mathcal{H}_\omega$ and $\||\Psi_\omega\rangle\|^2 = t_\omega$ for all $\omega \in \Omega'$. Then again by Lemma 4.5 we have $\mathrm{im}(D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}) = \overrightarrow{\mathrm{aff}(\Omega_\Psi)} = \overrightarrow{\mathrm{aff}(\Omega')}$. But $\dim \overrightarrow{\mathrm{aff}(\Omega')} \leq |\Omega'| - 1 \leq \dim \mathrm{conv}(\Omega) - 1$, so $D_{|\Psi\rangle\langle\Psi|}\iota^*|_{\mathcal{P}}$ cannot be surjective onto $\overrightarrow{\mathrm{aff}(\Omega)}$. This shows that ρ is a critical value. □

Remark 4.2. By Carathéodory's theorem, any representable density ρ can be written as a convex combination of $\dim \operatorname{conv}(\Omega)+1$ weights. However, almost none can be written as that of $\dim \operatorname{conv}(\Omega)$ weights, since those which can be expressed this way all lie in a finite union of hyperplanes of dimension $\dim \operatorname{conv}(\Omega) - 1$, which intersects $\operatorname{conv}(\Omega)$ in a set of measure zero. Alternatively, we can arrive at the same conclusion by applying Sard's theorem and Proposition 4.6.

Remark 4.3. A special case of Proposition 4.6 is when ρ lies on a facet F of $\operatorname{conv}(\Omega)$. In this case, ρ can be written as a convex combination of $\dim F + 1 = \dim \operatorname{conv}(\Omega)$ weights, hence ρ must be a critical value. On the other hand, we know this quite intuitively already from the fact that ρ lies on a facet of $\operatorname{conv}(\Omega)$: the derivative of $\iota^*|_{\mathcal{P}}$ at any pure state in the preimage of ρ cannot be surjective onto $\overrightarrow{\operatorname{aff}(\Omega)}$, because otherwise we would be able to walk outside of $\operatorname{conv}(\Omega)$ by perturbing the pure state in the right direction!

Combining Proposition 4.6 and Proposition 2.7, we get:

Corollary 4.7. *Let $\rho \in \operatorname{conv}(\Omega)$ be a representable density. If ρ is pure state v-representable and does not lie in the convex hull of $\dim \Omega$ weights, then ρ is uniquely v-representable.*

In Fig. 4.2, we show the sets of critical values in various abelian functional theories with $\dim \operatorname{conv}(\Omega) = 2$ or $\dim \operatorname{conv}(\Omega) = 3$. By Proposition 2.7, these lower dimensional polytopes are where the strong Hohenberg-Kohn theorem (Theorem 2.6) could potentially fail.

4.3 Boundary Forces

Finally, we turn our attention to boundary forces, with the main objective being to generalize and formalize the discussion in Section 3.5. The following theorem is the main result of this section:

Theorem 4.8. *Let $(V, \mathcal{H}, \iota, W)$ be an abelian functional theory. Let $F \subset \operatorname{conv}(\Omega)$ be a facet and $\rho_* \in F$ a density which is a regular value of the facet density map. In other words, ρ_* does not belong to the convex hull of $\dim F$ weights (Proposition 4.6). Take an inward-pointing vector $\eta \in \overrightarrow{\operatorname{aff}(\Omega)}$, and take $S \in V$ perpendicular to F such that $\langle \eta, S\rangle = 1$. Then*

$$\lim_{\epsilon \to 0^+} \frac{\mathcal{F}_p(\rho_* + \epsilon\eta) - \mathcal{F}_p(\rho_*)}{\sqrt{\epsilon}} = -2 \left[\sum_{\omega \in \Omega \setminus \Omega_F} \frac{\|\Pi_\omega W |\Phi\rangle\|^2}{\langle \omega, S\rangle - \langle \gamma, S\rangle} \right]^{\frac{1}{2}} \tag{4.4}$$

for any $\gamma \in \operatorname{aff}(F)$, where $|\Phi\rangle$ is a minimizer of the constrained search (2.7) at ρ_*. If there are multiple minimizers, $|\Phi\rangle$ is taken to minimize Eq. (4.4).

Remark 4.4. With the notation of Chapter 3, the facet F corresponds to an inequality constraint $D^F(\rho) = \langle \rho, S\rangle - \nu^F \geq 0$. It is easy to verify that $\nu^F = \langle \gamma, S\rangle$, so the denominator in Eq. (4.4) is nothing but $D^F(\omega)$.

Equation (4.4) implies $\mathcal{F}_p(\rho_* + \epsilon\eta) \approx \mathcal{F}_p(\rho_*) - \sqrt{\epsilon}\mathcal{G}$ for small ϵ, where $-\mathcal{G}$ is the right hand side of Eq. (4.4), and is the *repulsion strength* in Section 3.5.

In the rest of this section, we will present two derivations of this formula. The first one applies perturbation theory to ground states of Hamiltonians $H(v)$, where v is taken to be large in the direction of S, to deduce properties of ground states mapping to regions near the facet F. This approach has the advantage of being more or less intuitive and transparent. However, it has a crucial flaw: in order to use perturbation theory at all, we must assume, among other things, pure state v-representability. What this means is that we are actually making a statement about the Hohenberg-Kohn functional $\mathcal{F}_{HK}$, and not about the pure functional $\mathcal{F}_p$ as in the statement of Theorem 4.8. Thanks to Proposition 2.11, we know that $\mathcal{F}_{HK}$ and $\mathcal{F}_p$ at least agree on their common domain, so this approach is still valid for $\mathcal{F}_p$ as long as we stay within the set of pure state v-representable densities, which is in general a proper subset of $\mathrm{conv}(\Omega)$.

The second approach, which is much more involved, proves Theorem 4.8 directly from the definition of the pure functional $\mathcal{F}_p$ (Definition 2.6). In other words, the proof will be completely agnostic to v-representability issues. Despite the considerable number of technicalities one faces, the intuition is very simple: whenever a density ρ is close to the fixed density ρ_* on the facet F, say at "distance" ϵ (of course, we do not have an inner product on V^*, so we will have to be more precise), any state $|\Psi\rangle$ such that $\iota^*(|\Psi\rangle\langle\Psi|) = \rho$ must assume such a form where its overlap with $\mathcal{H}_\omega$ also approaches zero at a rate proportional to $\sqrt{\epsilon}$, where ω is any weight not on the facet. In that regime, the zeroth order contribution to the functional must be from the facet Hilbert space $\mathcal{H}_F$ (see Definition 4.3), to which a correction of order $\sqrt{\epsilon}$ from facet–off-facet interactions is to be added. This latter term is then the source of the boundary force.

4.3.1 Heuristic Derivation by Perturbation Theory

Here, we provide a quick derivation of Eq. (4.4) based on perturbation theory. As a precursor to the much more rigorous proof in Section 4.3.2, the present section will not dwell on precise language. The aim is to gain some intuition for why the boundary force should exist, and why the derivative of the functional should diverge as $1/\sqrt{\epsilon}$ as we get close to a facet.

Throughout this section, we will assume harmlessly that the functional domain $\iota^*(\mathcal{P}) = \iota^*(\mathcal{E})$ has full dimension in V^*. That is, we will assume $\dim \mathrm{conv}(\Omega) = \dim V^*$, or equivalently (Proposition 4.2) that $\iota : V \to i\mathfrak{u}(\mathcal{H})$ is injective and $\mathbb{1} \notin \iota(V)$. We will make the following further (extremely strong) assumptions:

(1) The densities in $\mathrm{relint}(\mathrm{conv}(\Omega))$ (which is just the usual interior) are pure state v-representable.

(2) $\mathcal{F}_p$ is differentiable on $\mathrm{relint}(\mathrm{conv}(\Omega))$.

(3) The ground state energy function $E(v) := \min_{\Gamma\in\mathcal{E}} \langle \Gamma, \iota(v) + W\rangle$ is differentiable.

(4) The densities in $\mathrm{relint}(F)$ are all uniquely pure state v-representable in the restricted functional theory on F.

By Proposition 2.11, (1) implies that the three functionals agree on the relative interior:

$$\mathcal{F}_p|_{\mathrm{relint\,conv}(\Omega)} \equiv \mathcal{F}_e|_{\mathrm{relint\,conv}(\Omega)} \equiv \mathcal{F}_{HK}|_{\mathrm{relint\,conv}(\Omega)}$$

Hence, in this section, we will ignore the distinction between the three functionals, and simply write $\mathcal{F}$ for all three. Assumptions (2) and (3) together allow us to use the following: if $\rho = \mathrm{d}_v E$, then $-v = \mathrm{d}_\rho \mathcal{F}$. (If differentiability is not assumed, we would have to replace the derivatives with the super/sub-differential instead.)

Set $U := \ker(\eta) \subset V$ and $\tilde{U} := S^\circ = \{\beta \in V^* | \langle \beta, S\rangle = 0\} \subset V^*$. We then get decompositions

$$V = \mathbb{R}S \oplus U$$
$$V^* = \mathbb{R}\eta \oplus \tilde{U}.$$

Note that $\tilde{U}$ is just $\overrightarrow{\mathrm{aff}(F)}$. Let $\pi_S : V \to \mathbb{R}S$ and $\pi_U : V \to U$ be the projections with respect to the given decomposition of V (i.e., $\pi_S|_{\mathbb{R}S} = \mathrm{Id}_{\mathbb{R}S}$ and $\ker \pi_S = U$, similarly for π_U). It is easy to show that $\pi_S(v) = \langle \eta, v\rangle\, S$. We are guided by the following intuition: for $v \in V$, we can write $v = v_S + v_U$ with respect to $V = \mathbb{R}S \oplus U$. If v_S is sent to infinity in the correct direction, the corresponding density will tend to the facet $F \subset \mathrm{conv}(\Omega)$.

For $v \in V$, we have $v = \pi_S(v) + \pi_U(v)$. The Hamiltonian can then be written as

$$H(v) = \iota(v) + W = \iota(\pi_S(v)) + \iota(\pi_U(v)) + W. \tag{4.5}$$

The subspace of ground states of such a Hamiltonian in the limit $\langle \eta, \pi_S(v)\rangle = \langle \eta, v\rangle \to +\infty$ is exactly the facet Hilbert space $\mathcal{H}_F \subset \mathcal{H}$, with energy

$$E_0 = \langle \Phi | \iota(\pi_S(v)) | \Phi\rangle = \langle \alpha, \pi_S(v)\rangle = \langle \gamma, \pi_S(v)\rangle\,, \tag{4.6}$$

where $|\Phi\rangle$ is any normalized state in $\mathcal{H}_F$ and $\alpha \in \Omega_F$ is any weight on the facet. The last equality holds because $\gamma - \alpha \in \overrightarrow{\mathrm{aff}(F)}$, and S is, by assumption, perpendicular to the facet F. We now apply perturbation theory to the Hamiltonian (4.5), treating $\iota(\pi_U(v)) + W$ as a small perturbation. Since the ground states are degenerate, the first order correction is the lowest eigenvalue of $\Pi_F[\iota(\pi_U(v)) + W]\Pi_F^\dagger$ (i.e., the perturbing Hamiltonian restricted/compressed to the ground state subspace $\mathcal{H}_F$), which is nothing but the ground state energy of the external potential $\pi_U(v) \in V$ in the restricted functional theory on F, corresponding to a unique ground state $|\Psi(v)\rangle \in \mathcal{H}_F$ by assumption (4). That is, we have the first order energy correction

$$E_1 = E_F(\pi_U(v)), \tag{4.7}$$

where $E_F : V \to \mathbb{R}$ is the ground state energy function of the facet functional theory. It is now straightforward to obtain the second order correction to the ground state energy: the excited states are the weight states not lying on F, with each weight space $\mathcal{H}_\omega$ having energy $\langle \omega, \pi_S(v)\rangle$. Combining the second order correction with Eq. (4.6) and Eq. (4.7), we get the ground state energy up to second order:

$$E(v) \approx \langle \gamma, \pi_S(v)\rangle + E_F(\pi_U(v)) + \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\overbrace{\|\Pi_\omega W\,|\Phi(v)\rangle\|^2}^{=:r_\omega(v)}}{\langle \gamma, \pi_S(v)\rangle - \langle \omega, \pi_S(v)\rangle}.$$

Note that $\langle \gamma, \pi_S(v)\rangle + E_F(\pi_U(v))$ is simply $E_F(v)$, due to the fact that $\iota_F(\pi_S(v)) = \langle \gamma, \pi_S(v)\rangle\, \mathbb{1}_{\mathcal{H}_F}$. It follows that

$$E(v) \approx E_F(v) + \sum_{\omega \in \Omega \setminus \Omega_F} \frac{r_\omega(v)}{\langle \gamma, \pi_S(v)\rangle - \langle \omega, \pi_S(v)\rangle}. \tag{4.8}$$

Taking the derivative at $v \in V$, we get

$$\mathrm{d}_v E \approx \underbrace{\mathrm{d}_v E_F}_{\in \iota_F^*(\mathcal{E}_F)=F} \underbrace{- \sum_{\omega \in \Omega \setminus \Omega_F} \frac{r_\omega(v)}{\langle \gamma - \omega, \pi_S(v) \rangle^2} (\gamma \circ \pi_S - \omega \circ \pi_S)}_{\in \mathbb{R}\eta} + \underbrace{\sum_{\omega \in \Omega \setminus \Omega_F} \frac{\mathrm{d}_v r_\omega}{\langle \gamma - \omega, \pi_S(v) \rangle}}_{\in \tilde{U}}, \tag{4.9}$$

which is valid only for $\langle \eta, v \rangle \to +\infty$. The second term is proportional to η because it vanishes on $\ker \eta = U$. Similarly, the third term belongs to $\tilde{U}$ because

$$\langle \mathrm{d}_v r_\omega, S \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} r_\omega(v + tS) = 0,$$

where the last equality follows from $|\Phi(v)\rangle = |\Phi(v + tS)\rangle$ since $\iota_F(e)$ is proportional to the identity in $\mathcal{H}_F$. Equation (4.9) has an intuitive interpretation. As $\langle \eta, v \rangle \to +\infty$, we are driving the ground state density (left hand side) close to the facet F, and this equation tells us how the ground state density ρ approaches F: the first term on the right hand side, which is the derivative of the energy function of the facet functional theory, can be interpreted as a density on F, and the last two terms represent how ρ deviates from the density on the facet, with the deviation decomposed into a "longitudinal" component (in $\mathbb{R}\eta$) and a "transverse" one (in $\tilde{U} = \overrightarrow{\mathrm{aff}(F)}$). Since $\langle \gamma - \omega, \pi_S(v) \rangle = \langle \gamma - \omega, S \rangle \cdot \langle \eta, v \rangle$, we immediately see that the longitudinal component is inversely proportional to $\langle \eta, v \rangle^2$. In other words, in order for ρ to be ϵ-close to the facet F, we must have $\langle \eta, v \rangle \propto 1/\sqrt{\epsilon}$. But the external potential v is just the negative of the derivative of the functional $\mathcal{F}$ at ρ, so $\mathrm{d}_\rho \mathcal{F}$ must behave like $1/\sqrt{\epsilon}$. We will now try to make this argument more precise.

For b small enough, define a path $c : (0, b) \to V$ by $\mathrm{d}_{c(\epsilon)} E = \rho_* + \epsilon\eta$. We know this characterizes c uniquely because $v \mapsto \mathrm{d}_v E$ and $\rho \mapsto -\mathrm{d}_\rho \mathcal{F}$ are inverses of each other. Applying Eq. (4.9) to $\mathrm{d}_{c(\epsilon)} E$ and rearranging, we get

$$\underbrace{\mathrm{d}_{c(\epsilon)} E_F - \rho_* + \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\mathrm{d}_{c(\epsilon)} r_\omega}{\langle \gamma - \omega, \pi_S(c(\epsilon)) \rangle}}_{\in \tilde{U}}$$
$$\underbrace{- \sum_{\omega \in \Omega \setminus \Omega_F} \frac{r_\omega(c(\epsilon))}{\langle \gamma - \omega, \pi_S(c(\epsilon)) \rangle^2} ((\gamma - \omega) \circ \pi_S) - \epsilon\eta}_{\in \mathbb{R}\eta} \approx 0 \in V^*.$$

Since $(\mathbb{R}\eta) \cap \tilde{U} = 0$, the two underlined groups of terms must vanish separately. For the second group of terms, we evaluate everything on S, keeping in mind that $\langle \eta, S \rangle = 1$ and $\pi_S(S) = S$:

$$\begin{cases} 0 \approx \mathrm{d}_{c(\epsilon)} E_F - \rho_* + \displaystyle\sum_{\omega \in \Omega \setminus \Omega_F} \frac{\mathrm{d}_{c(\epsilon)} r_\omega}{\langle \gamma - \omega, \pi_S(c(\epsilon)) \rangle} \\ \epsilon \approx - \displaystyle\sum_{\omega \in \Omega \setminus \Omega_F} \frac{r_\omega(c(\epsilon))}{\langle \gamma - \omega, \pi_S(c(\epsilon)) \rangle^2} \langle \gamma - \omega, S \rangle \end{cases} \tag{4.10}$$

Write $c(\epsilon) = \tau(\epsilon) S + c_U(\epsilon)$ where $\tau(\epsilon) \in \mathbb{R}$ and $c_U(\epsilon) \in U$. Then the second equation of Eq. (4.10) becomes

$$\epsilon \approx \frac{1}{\tau(\epsilon)^2} \sum_{\omega \in \Omega \setminus \Omega_F} \frac{r_\omega(c(\epsilon))}{\langle \omega - \gamma, S \rangle} \Rightarrow \tau(\epsilon) \approx \frac{1}{\sqrt{\epsilon}} \left(\sum_{\omega \in \Omega \setminus \Omega_F} \frac{r_\omega(c(\epsilon))}{\langle \omega - \gamma, S \rangle} \right)^{\frac{1}{2}}.$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* + \epsilon\eta) = \langle \eta, \mathrm{d}_{\rho_*+\epsilon\eta}\mathcal{F}\rangle = \langle \eta, \mathrm{d}_{\mathrm{d}_{c(\epsilon)}E}\mathcal{F}\rangle = -\langle \eta, c(\epsilon)\rangle = -\tau(\epsilon)$$
$$\approx -\frac{1}{\sqrt{\epsilon}}\left(\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega W\,|\Phi(c(\epsilon))\rangle\|^2}{\langle\omega-\gamma, S\rangle}\right)^{\frac{1}{2}}.$$

We may replace $|\Phi(c(\epsilon))\rangle$ with its limit as $\epsilon \to 0$, potentially making an error of higher order in ϵ. The limit is nothing but $|\Phi\rangle$, the unique state corresponding to the density $\rho_* \in F$ in the facet functional theory. Therefore, for ϵ small,

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\mathcal{F}(\rho_* + \epsilon\eta) = 2\sqrt{\epsilon}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F} \approx -2\left(\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega W\,|\Phi\rangle\|^2}{\langle\omega-\gamma, S\rangle}\right)^{\frac{1}{2}}, \tag{4.11}$$

which is the claimed boundary force formula (Eq. (4.4)).

4.3.2 Proof by Constrained Search

Here, we present a rigorous proof of Theorem 4.8 by constrained search. Since the proof is long, it will bit split into sections for the sake of comprehensibility.

Step I: The "No Mixing Lemma"

The first step is to have good control over minimizers $|\Phi\rangle$ of the constrained search on the facet. Informally speaking, we will show that "a pure state which is a minimizer of the constrained search does not mix with a state which is strongly orthogonal to it under the interaction W". Since this statement is entirely general, in the sense that it applies to any density lying in the relative interior of the functional domain (convex hull of all weights) of any abelian functional theory, we will not consider the facet functional theory here. In Step III, we will apply this result to the facet functional theory.

Lemma 4.9 (No Mixing Lemma)**.** *Let $(V, \mathcal{H}, \iota, W)$ be any abelian functional theory with weights Ω. Suppose ρ_* lies in the relative interior of* $\mathrm{conv}(\Omega)$*. Let $|\Phi\rangle$ be a minimizer of the pure state constrained search at ρ_*. In other words, $\iota^*(|\Phi\rangle\langle\Phi|) = \rho_*$ and $\langle\Psi|W|\Psi\rangle \geq \langle\Phi|W|\Phi\rangle$ for all states $|\Psi\rangle$ such that $\iota^*(|\Psi\rangle\langle\Psi|) = \rho_*$. Let $|\delta\rangle$ be a weight vector. If $\langle\delta|\Phi\rangle = 0$, then $\langle\delta|W|\Phi\rangle = 0$.*

Proof. We argue by contradiction. The strategy is as follows: suppose $\langle\delta|W|\Phi\rangle \neq 0$. We will find a state $|\Psi\rangle$ such that $\iota^*(|\Psi\rangle\langle\Psi|) = \rho_*$ and $\langle\Psi|W|\Psi\rangle < \langle\Phi|W|\Phi\rangle$, which contradicts the assumption that $|\Phi\rangle$ is a minimizer of the constrained search. Finding such a state will be achieved by constructing a one-parameter family of states $|\Psi(t)\rangle$ depending smoothly on t with $|\Psi(0)\rangle = |\Phi\rangle$, and showing that $\iota^*(|\Psi(t)\rangle\langle\Psi(t)|) = \rho_*$ for all t and that $\frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}\langle\Psi(t)|W|\Psi(t)\rangle \neq 0$.

Let $\omega \in \Omega$ denote the weight of $|\delta\rangle$. In other words, $\iota(v)\,|\delta\rangle = \langle\omega, v\rangle\,|\delta\rangle$ for all external potentials $v \in V$. We will distinguish between two cases depending on whether ω coincides with ρ_* or not. This is technically not necessary as the argument in the latter case also applies to the former, but we nevertheless choose to present the proof this way for more clarity.

Without loss of generality, we assume $\|\delta\| = 1$.

Case (a): $\omega = \rho_*$
For any angle ϕ, define $|\Psi(t)\rangle := te^{i\phi}|\delta\rangle + \sqrt{1-t^2}\,|\Phi\rangle$. Then $\iota^*(|\Psi(t)\rangle\langle\Psi(t)|) = t^2\omega + (1-t^2)\omega = \omega = \rho_*$. The derivative of the expectation value $\langle\Psi(t)|W|\Psi(t)\rangle$ at $t = 0$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle\Psi(t)|W|\Psi(t)\rangle = 2\,\mathrm{Re}\left[e^{-i\phi}\,\langle\delta|W|\Phi\rangle\right]. \tag{4.12}$$

Since $\langle\delta|W|\Phi\rangle \neq 0$, there is a suitable choice of angle ϕ such that the expression above does not vanish.

Case (b): $\omega \neq \rho_*$
If $\omega \neq \rho_*$, consider the ray starting from ω extending in the direction of ρ_*. This ray intersects the boundary of $\mathrm{conv}(\Omega)$ in some point μ. Since the boundary of $\mathrm{conv}(\Omega)$ is exactly the union of all its facets, μ belongs to at least one facet, which we will call H (see Figure 4.3).

Let $(|E_i\rangle)_{i=1}^{\dim\mathcal{H}}$ be an orthonormal basis for $\mathcal{H}$ which is compatible with the weight space decomposition. That is, each $|E_i\rangle$ is a weight vector with weight $\omega_i \in \Omega$. Without loss of generality, way may take $|E_1\rangle = |\delta\rangle$. Let $I_\Phi := \{i \mid \langle E_i|\Phi\rangle \neq 0\}$ and $I_H := \{i \mid \omega_i \in H\}$.

Since ρ_* is assumed to be in the relative interior of $\mathrm{conv}(\Omega)$, we have $\rho_* \neq \mu$. It follows that ω, ρ_*, μ are distinct and collinear, and hence ρ_* is a strict convex combination of the other two. In other words, $\rho_* = u_1\omega + u_2\mu$ with $u_1 + u_2 = 1$ and $u_1, u_2 \in (0,1)$. Since μ lies on the facet H, we can express μ as a convex combination $\mu = \sum_{i\in I_H} q_i\omega_i$.

Write $|\Phi\rangle = \sum_{i\in I_\Phi} c_i e^{i\theta_i}|E_i\rangle$, where the c_i are positive. For arbitrary angles $(\phi_i)_{i\in I_H\setminus I_\Phi}$ and ϕ, consider the smooth curve

$$\begin{aligned}|\Psi(t)\rangle := &\sum_{i\in I_\Phi\setminus I_H} \sqrt{c_i^2(1-t^2)}e^{i\theta_i}|E_i\rangle + \sum_{i\in I_\Phi\cap I_H} \sqrt{c_i^2(1-t^2)+u_2t^2q_i}e^{i\theta_i}|E_i\rangle \\ &+ t\sqrt{u_2}\sum_{i\in I_H\setminus I_\Phi}\sqrt{q_i}e^{i\phi_i}|E_i\rangle + t\sqrt{u_1}e^{i\phi}|\delta\rangle,\end{aligned} \tag{4.13}$$

which is defined on a sufficiently small open neighborhood of 0. We claim that the four terms in Eq. (4.13) are mutually strongly orthogonal. This is clear for the first three terms: they are spanned by disjoint sets of weight vectors. To see that $|E_1\rangle = |\delta\rangle$ is strongly orthogonal to the rest, note that $1 \notin I_\Phi$ because $\langle E_1|\Phi\rangle = 0$ by assumption. Moreover, $1 \notin I_H$ because ω clearly cannot lie on the facet H: indeed, if ω belonged to H, then $\rho_* = u_1\omega + u_2\mu$ would also lie on H, contradicting the assumption that $\rho_* \in \mathrm{relint}(\mathrm{conv}(\Omega))$. Thus $1 \notin I_\Phi \cup I_H$ and we obtain the desired conclusion.

The mutual orthogonality implies

$$\begin{aligned}\|\Psi(t)\|^2 &= \sum_{i\in I_\Phi\setminus I_H} c_i^2(1-t^2) + \sum_{i\in I_\Phi\cap I_H}[c_i^2(1-t^2)+u_2t^2q_i)] + t^2u_2\sum_{i\in I_H\setminus I_\Phi} q_i + t^2u_1 \\ &= (1-t^2)\sum_{i\in I_\Phi} c_i^2 + t^2u_2\sum_{i\in I_H} q_i + t^2u_1 \\ &= 1 - t^2 + t^2u_2 + t^2u_1 = 1.\end{aligned}$$

So $|\Psi(t)\rangle$ is normalized for all t. The mutual strong orthogonality implies that the density of

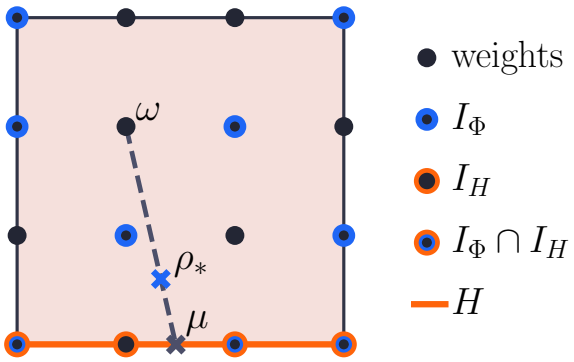


Figure 4.3 Illustration of the construction of the curve $|\Psi(t)\rangle$. In this diagram, we have assumed that each weight space has multiplicity one for simplicity.

$|\Psi(t)\rangle\langle\Psi(t)|$ is the sum of the densities of its constituents (Lemma 4.4):

$$\begin{aligned}
&\iota^*(|\Psi(t)\rangle\langle\Psi(t)|) \\
&= \sum_{i\in I_\Phi\setminus I_H} c_i^2(1-t^2)\omega_i + \sum_{i\in I_\Phi\cap I_H} [c_i^2(1-t^2)+u_2t^2q_i]\omega_i + t^2u_2\sum_{i\in I_H\setminus I_\Phi} q_i\omega_i + t^2u_1\omega \\
&= (1-t^2)\sum_{i\in I_\Phi} c_i^2\omega_i + t^2u_2\sum_{i\in I_H} q_i\omega_i + t^2u_1\omega \\
&= (1-t^2)\rho_* + t^2u_2\mu + t^2u_1\omega = (1-t^2)\rho_* + t^2\rho_* \\
&= \rho_*
\end{aligned}$$

So we see that $|\Psi(t)\rangle\langle\Psi(t)|$ and $|\Phi\rangle$ share the same density. Finally, it is easy to verify from Eq. (4.13) that $|\Psi(0)\rangle = |\Phi\rangle$.

Now we compute the derivative of the expectation value $\langle\Psi(t)|W|\Psi(t)\rangle$ at $t = 0$. The only contributions we get are from differentiating the last two terms of Eq. (4.13), since the other two are smooth in t^2. Therefore,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle\Psi(t)|W|\Psi(t)\rangle \\
&= 2\,\mathrm{Re}\left[\sqrt{u_2}\sum_{i\in I_H\setminus I_\Phi}\sqrt{q_i}e^{-i\phi_i}\langle E_i|W|\Phi\rangle\right] + 2\,\mathrm{Re}\left[\sqrt{u_1}e^{-i\phi}\langle\delta|W|\Phi\rangle\right].
\end{aligned} \tag{4.14}$$

Remember that the angles $(\phi_i)_{i\in I_H\setminus I_\Phi}$ and ϕ can be chosen arbitrarily. We then choose the ϕ_i so that the first term is nonnegative. Because $\langle\delta|W|\Phi\rangle \neq 0$ and $u_1 > 0$, we can choose ϕ so that the second term is positive. But this implies that there is a t_0 such that $\langle\Psi(t_0)|W|\Psi(t_0)\rangle < \langle\Phi|W|\Phi\rangle$. □

Lemma 4.9 can be made slightly stronger:

Corollary 4.10. *Let $\rho_*, |\Phi\rangle$ be as above. Then $\langle\Gamma|W|\Phi\rangle = 0$ for any vector $|\Gamma\rangle$ that is strongly orthogonal to $|\Phi\rangle$.*

Proof. Write $|\Gamma\rangle = \sum_{\omega\in\Omega_\Gamma}|\Gamma_\omega\rangle$. We have $\langle\Gamma_\omega|\Phi\rangle = 0$ by strong orthogonality. Applying Lemma 4.9 to each $|\Gamma_\omega\rangle$ then yields $\langle\Gamma_\omega|W|\Phi\rangle = 0$, which implies $\langle\Gamma|W|\Phi\rangle = 0$. □

To see how Lemma 4.9 can fail if we do not assume $\rho_* \in \mathrm{relint}(\mathrm{conv}(\Omega))$, let ρ_* be any density on a facet $F \subset \mathrm{conv}(\Omega)$. Let $|\Phi\rangle$ be a minimizer of the constrained search at ρ_*. Then $|\Phi\rangle$ is orthogonal to $|E_i\rangle$ for any i such that $\omega_i \notin F$. But there always exists an interaction W such that $\langle E_i|W|\Phi\rangle \neq 0$.

Step II: Facet – Off-Facet Decomposition

Next, we will decompose the pure functional $\mathcal{F}_p$ into three pieces. In essence, the strategy is to split each state $|\Phi\rangle$ whose density $\iota^*(|\Phi\rangle\langle\Phi|)$ is close to a facet $F \subset \mathrm{conv}(\Omega)$ into $|\Phi_F\rangle + |\Phi'\rangle$, where $|\Phi_F\rangle$ has support only on the facet, and $|\Phi'\rangle$ has support only off-facet. The expectation value $\langle\Phi|W|\Phi\rangle$ then splits into three terms, called "facet-facet", "facet–off-facet", and "off-facet–off-facet" contributions, which is the content of Theorem 4.23. The punchline is then that for the boundary force, only the facet–off-facet term will matter, which is shown in Step III.

For the actual boundary force computation, which we carry out in Step IV, it is necessary to have a manageable characterization of the fiber $(\iota^*)^{-1}(\rho)$ for a density ρ close to the facet F. To study the fiber, we first restrict our attention to the squared moduli of the coefficients $(y_i)_i$ (with respect to some chosen basis) of a state $|\Psi\rangle\langle\Psi| \in (\iota^*)^{-1}(\rho)$, which leads us naturally to the definition of the *classical density map*, which is an affine map. We will show (Proposition 4.17) that whenever a density ρ is close to some density ρ_* on the facet, the fiber of the classical density map over ρ is approximately the product of the fiber over ρ_* and another convex polytope, with the latter parametrizing, roughly speaking, the different directions in which a state can deviate from one mapping to the facet. We then argue (Lemma 4.20, Corollary 4.22) that the fiber can be replaced by an actual, honest product of two convex polytopes due to topological reasons, which allows us to obtain the so-called facet–off-facet decomposition of the pure functional (Theorem 4.23).

Definition 4.5. *Let $(V, \mathcal{H}, \iota, W)$ be an abelian functional theory together with an orthonormal basis $(|E_i\rangle)_{i\in I}$, $|I| = \dim \mathcal{H}$, for $\mathcal{H}$ which is compatible with the weight space decomposition $\mathcal{H} = \bigoplus_{\omega\in\Omega} \mathcal{H}_\omega$. For each $i \in I$, let $\omega_i \in \Omega$ be the weight of $|E_i\rangle$. The **space of classical states** is the $(\dim \mathcal{H} - 1)$-simplex $\Delta_I := \{y \in \mathbb{R}^I \mid \sum_{i\in I} y_i = 1, y_i \geq 0\}$. The **classical density map** is*

$$\begin{aligned} \mathcal{A} : \Delta_I &\to \mathrm{conv}(\Omega) \\ y &\mapsto \sum_{i\in I} y_i \omega_i. \end{aligned} \tag{4.15}$$

$\mathcal{A}$ can be extended to an affine map $\bar{\mathcal{A}} : \mathrm{aff}(\Delta_I) = \{y \in \mathbb{R}^I \mid \sum_i y_i = 1\} \to \mathrm{aff}(\Omega)$. We will denote the translation space of the space of classical states by $\mathfrak{D} := \overrightarrow{\mathrm{aff}(\Delta_I)} = \{t \in \mathbb{R}^I \mid \sum_i t_i = 0\}$. Finally, we write $\mathfrak{N}$ for the kernel of $D\mathcal{A} : \mathfrak{D} \to \overrightarrow{\mathrm{aff}(\Omega)}$.

Proposition 4.11. *If ρ lies in the relative interior of $\mathrm{conv}(\Omega)$, then $\mathcal{A}^{-1}(\rho)$ is a convex polytope of dimension $\dim \mathfrak{N}$.*

Proof. $\mathcal{A}^{-1}(\rho) = \bar{\mathcal{A}}^{-1}(\rho) \cap \Delta_I$ is a convex polytope because it is the intersection of an affine space with a simplex. Since $\bar{\mathcal{A}}^{-1}(\rho) \subset \mathrm{aff}(\Delta_I)$, such an intersection will have the same dimension as $\dim \bar{\mathcal{A}}^{-1}(\rho) = \dim \mathfrak{N}$ as long as $\bar{\mathcal{A}}^{-1}(\rho) \cap \mathrm{relint}(\Delta_I) \neq \varnothing$. Showing the latter condition is equivalent to finding $t \in \Delta_I$, $t_i > 0$, such that $\mathcal{A}(t) = \rho$.

Take any $i \in I$. If $\omega_i = \rho$, set $t_j^{(i)} = \delta_{ij}$. Otherwise, consider the ray from ω_i through ρ. This ray intersects the convex polytope $\mathrm{conv}(\Omega)$ in a point μ, which lies on at least one facet F. Expressing ρ as a strict (because ρ is in the relative interior) convex combination of ω_i and μ, and writing μ as a convex combination of $\omega_j, j \in I_F$, we obtain a convex combination $\rho = \sum_{j\in I} t_j^{(i)} \omega_j$, where $t_i^{(i)} > 0$. (Note that this argument is very similar to the one used in the proof of Lemma 4.9)

Finally, we have

$$\rho = \frac{1}{|I|} \sum_{i\in I} \sum_{j\in I} t_j^{(i)} \omega_j = \sum_{j\in I} \underbrace{\left(\sum_{i\in I} \frac{t_j^{(i)}}{|I|} \right)}_{=:t_j} \omega_j = \sum_{j\in I} t_j \omega_j.$$

For each $j \in I$, the coefficient t_j is a sum of nonnegative terms, with at least one being positive, namely $t_j^{(j)} > 0$. It follows that $t_j > 0$ and $t \in \mathrm{relint}(\Delta_I)$. □

The motivation for studying the classical density map $\mathcal{A}$ is that understanding its fibers allows us to understand the fibers of the (pure state) density map $\iota^*|_{\mathcal{P}}$ up to phases. More precisely:

Proposition 4.12. *For any $\rho \in \mathrm{conv}(\Omega)$,*

$$(\iota^*)^{-1}(\rho) \cap \mathcal{P} = \left\{ |\Psi\rangle\langle\Psi| \,\middle|\, |\Psi\rangle = \sum_{i\in I} \sqrt{y_i}\xi_i |E_i\rangle \,,\ y \in \mathcal{A}^{-1}(\rho), \xi_i \in S^1 \right\}, \tag{4.16}$$

where $S^1 = \{\xi \in \mathbb{C} \mid |\xi| = 1\}$ is the set of phases.

Proof. This is easily shown by direct calculation. □

Hence, we immediately obtain a formula of the pure functional expressed as a minimization over both the phases and the fiber of $\mathcal{A}$:

Corollary 4.13. *For any $\rho \in \mathrm{conv}(\Omega)$, the value of the pure functional at ρ is*

$$\mathcal{F}_p(\rho) = \min_{\xi\in(S^1)^I} \min_{y\in\mathcal{A}^{-1}(\rho)} \sum_{i,j\in I} \sqrt{y_i y_j}\bar{\xi}_i \xi_j W_{ij}, \tag{4.17}$$

where $W_{ij} := \langle E_i|W|E_j\rangle$.

The rest of this section will be dedicated to understanding the fiber $\mathcal{A}^{-1}(\rho)$ when ρ is close to a facet $F \subset \mathrm{conv}(\Omega)$.

If $F \subset \mathrm{conv}(\Omega)$ is a facet, there is a natural choice of orthonormal basis for $\mathcal{H}_F$, namely $(|E_i\rangle)_{i\in I_F}$, where $I_F := \{i \in I \mid \omega_i \in F\}$. Moreover, this basis is easily seen to be compatible with the weight space decomposition $\mathcal{H}_F = \bigoplus_{\omega\in\Omega_F} \mathcal{H}_\omega$. Hence, there corresponds a classical density map for the facet functional theory, which we will denote by $\mathcal{A}_F : \Delta_{I_F} \to \mathrm{conv}(\Omega_F)$. We will think of Δ_{I_F} as a subset of Δ_I by mapping $y \in \Delta_{I_F}$ to $\tilde{y} \in \Delta_I$, where $\tilde{y}_i = 0$ if $i \notin I_F$ and $\tilde{y}_i = y_i$ if $i \in I_F$. Similarly, we think of $\mathfrak{D}_F := \overrightarrow{\mathrm{aff}(\Delta_{I_F})}$ as a subspace of $\mathfrak{D}$ in the same way.

Proposition 4.14. *$\mathcal{A}_F$ and $D\mathcal{A}_F$ are the restrictions of $\mathcal{A}$ and $D\mathcal{A}$ respectively:*

$$\begin{array}{ccc} \Delta_F & \xrightarrow{\mathcal{A}_F} & \mathrm{conv}(\Omega_F) \\ \downarrow & & \downarrow \\ \Delta & \xrightarrow[\mathcal{A}]{} & \mathrm{conv}(\Omega) \end{array} \qquad \begin{array}{ccc} \mathfrak{D}_F & \xrightarrow{D\mathcal{A}_F} & \overrightarrow{\mathrm{aff}(\Omega_F)} \\ \downarrow & & \downarrow \\ \mathfrak{D} & \xrightarrow[D\mathcal{A}]{} & \overrightarrow{\mathrm{aff}(\Omega)} \end{array} \tag{4.18}$$

Proof. This is obvious from the definitions of $\mathcal{A}$ and $\mathcal{A}_F$. □

Corollary 4.15. $\mathfrak{N}_F = \mathfrak{N} \cap \mathfrak{D}_F$.

Proposition 4.16.

(1) $\dim \mathfrak{N} = \dim \mathcal{H} - \dim \mathrm{conv}(\Omega) - 1$

(2) $\dim \mathfrak{N}_F = \dim \mathcal{H}_F - \dim F - 1$

Proof. (1) follows from applying the rank-nullity theorem to $D\mathcal{A}$. (2) follows from (1) applied to the facet functional theory. □

Pick a complement $\mathfrak{N}'$ of $\mathfrak{N}_F$ in $\mathfrak{N}$. For example, we could take $\mathfrak{N}'$ to be the orthogonal complement of $\mathfrak{N}_F$. Roughly speaking, we want to use $\mathfrak{N}'$ to parametrize the ways a classical state can distribute its coefficients (interpreted as probabilities) in the off-facet weights, while $\mathfrak{N}_F$ parametrizes how the coefficients within the facet F are distributed. This idea is made precise in the following proposition:

Proposition 4.17. *Let ρ_* be in the relative interior of a facet $F \subset \mathrm{conv}(\Omega)$, let $\eta \in \overrightarrow{\mathrm{aff}(\Omega)}$ be an inward-pointing vector (relative to $\mathrm{conv}(\Omega)$, see Fig. 4.4), and fix an $l \in \mathfrak{D}$ that solves $D\mathcal{A}(l) = \eta$. Then for any $\epsilon \geq 0$ such that $\rho_* + \epsilon\eta \in \mathrm{conv}(\Omega)$, the fiber of the classical density map $\mathcal{A}$ above $\rho_* + \epsilon\eta$, i.e., the set of all $y \in \Delta_I$ that solve $\mathcal{A}(y) = \rho_* + \epsilon\eta$, is given by*

$$\mathcal{A}^{-1}(\rho_* + \epsilon\eta) = \left\{ y + \epsilon q \middle| q \in \mathbf{P}', y \in \mathbf{P}^{\parallel}(\epsilon q) \right\}, \tag{4.19}$$

where

$$\begin{aligned} \mathbf{P}' &:= \{q \in l + \mathfrak{N}' \mid \forall i \in I \setminus I_F : q_i \geq 0\} \\ \mathbf{P}^{\parallel}(\epsilon q) &:= \{y \in \mathrm{aff}(\Delta_{I_F}) \mid \bar{\mathcal{A}}(y) = \rho_*, \forall i \in I_F : y_i \geq -\epsilon q_i\}. \end{aligned} \tag{4.20}$$

Additionally, we have $\mathbf{P}^{\parallel}(0) = \mathcal{A}^{-1}(\rho_*)$.

Proof. $\mathbf{P}^{\parallel}(0) = \{y \in \mathrm{aff}(\Delta_{I_F}) \mid \bar{\mathcal{A}}(y) = \rho_*\} \cap \Delta_{I_F} = \mathcal{A}^{-1}(\rho_*)$. This proves the last assertion.

If $\epsilon = 0$, the right hand side of Eq. (4.19) reduces to $\mathbf{P}^{\parallel}(0) = \mathcal{A}^{-1}(\rho_*)$, which is the same as the left hand side. Thus we assume $\epsilon > 0$ in the following.

Suppose $z \in \Delta_I$ is a classical state which solves $\mathcal{A}(z) = \rho_* + \epsilon\eta$. Take any $y' \in \Delta_{I_F}$ such that $\mathcal{A}(y') = \rho_*$. Then $D\mathcal{A}(z - y' - \epsilon l) = 0$, so $z = y' + \epsilon l + s + t$, where $s \in \mathfrak{N}_F$ and $t \in \mathfrak{N}'$. Thus $z = y + \epsilon q$ with $y = y' + s$ and $q = l + \epsilon^{-1} q$. Because $z \in \Delta_I$, we have $y_i + \epsilon q_i \geq 0$ for all $i \in I$. If $i \in I \setminus I_F$, this implies $q_i \geq 0$. This shows "$\subset$" in Eq. (4.19).

Conversely, take $z = y + \epsilon q$ with $q \in \mathbf{P}'$ and $y \in \mathbf{P}^{\parallel}(\epsilon q)$. Then

$$\mathcal{A}(z) = \mathcal{A}(y) + \epsilon D\mathcal{A}(q) = \rho_* + \epsilon\eta.$$

Clearly $z \in \Delta_I$ (the inequalities in Eq. (4.20) ensure that $z_i \geq 0$ for all $i \in I$). It follows that $z \in \mathcal{A}^{-1}(\rho_* + \epsilon\eta)$. This shows "$\supset$". □

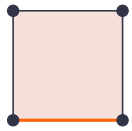
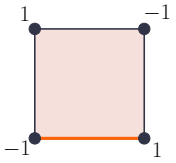
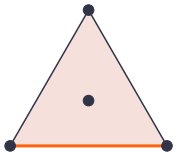
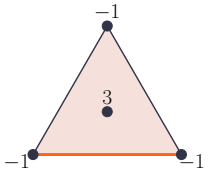
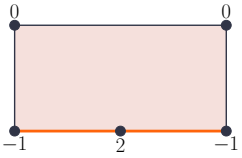
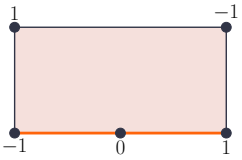
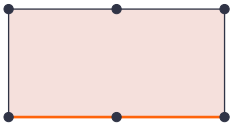
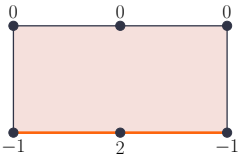
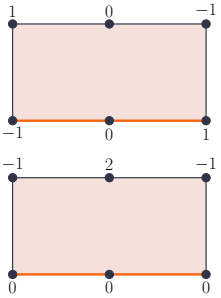
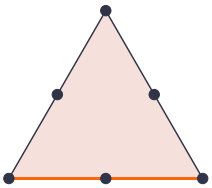
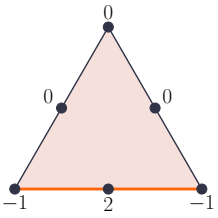
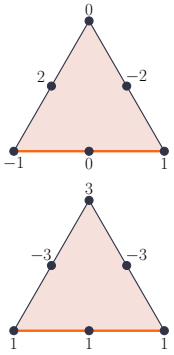
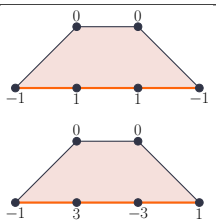
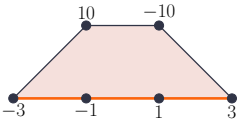
$\Omega, \mathrm{conv}(\Omega), F$	$\dim \mathcal{H}$	$\dim \mathcal{H}_F$	$\mathfrak{N}_F$	$\mathfrak{N}'$
	4	2		
	4	2		
	5	3		
	6	3		
	6	3		
	6	4		

Table 4.1 The subspaces $\mathfrak{N}_F$ and $\mathfrak{N}'$ for several sets of weights Ω and choices of facet $F \subset \mathrm{conv}(\Omega)$. For simplicity, the multiplicity of each weight space is always one, and we take $\mathfrak{N}' = \mathfrak{N}_F^{\perp}$. We present $\mathfrak{N}_F$ and $\mathfrak{N}'$ by listing a set of spanning vectors, which all belong to $\mathfrak{N} \subset \mathfrak{D} := \{t \in \mathbb{R}^I \mid \sum_{i \in I} t_i = 0\}$.

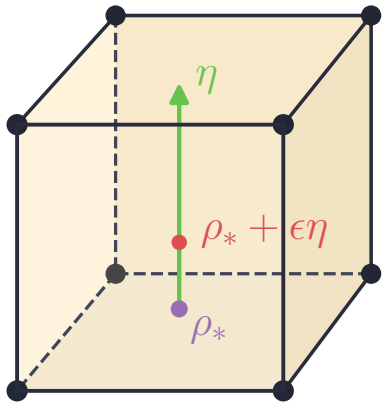


Figure 4.4 Starting with a density ρ_* on a facet (in this case the bottom face of the cube), we pick an inward direction η and study the fiber of $\mathcal{A}$ over $\rho_* + \epsilon\eta$.

Proposition 4.18.

(1) $\mathbf{P}'$ *is a convex polytope.*

(2) $\mathbf{P}^{\parallel}(\epsilon q)$ *is a convex polytope for all* $\epsilon \geq 0$ *and* $q \in \mathbf{P}'$.

(3) If $\rho_* \in \operatorname{relint}(F)$, *then* $\dim \mathbf{P}^{\parallel}(0) = \dim \mathfrak{N}_F$.

Proof. Since $\mathbf{P}'$ is a subset of $l + \mathfrak{N}'$ defined as an intersection of half spaces, it suffices to show that $\mathbf{P}'$ does not contain a ray. For the sake of contradiction, suppose $x + \mathbb{R}_{\geq 0} q$, $q \in \mathfrak{N}' \setminus \{0\}$, were such a ray contained in $\mathbf{P}'$. Let $v \in V$ be a vector (external potential) such that $\langle \omega_i, v \rangle \geq c$ for all $i \in I$, with equality if and only if $i \in I_F$. Then

$$\begin{aligned} 0 = \langle D\mathcal{A}(q), v \rangle &= \left\langle \sum_{i \in I} q_i \omega_i, v \right\rangle = \sum_{i \in I_F} q_i \langle \omega_i, v \rangle + \sum_{i \in I \setminus I_F} q_i \langle \omega_i, v \rangle \\ &= \sum_{i \in I_F} q_i (\langle \omega_i, v \rangle - c) + \sum_{i \in I \setminus I_F} q_i (\langle \omega_i, v \rangle - c) \\ &= \sum_{i \in I \setminus I_F} q_i \underbrace{(\langle \omega_i, v \rangle - c)}_{>0}, \end{aligned}$$

where we have used $\sum_i q_i = 0$. Since $q \in \mathfrak{N}' \setminus \{0\}$, $q_j \neq 0$ for at least one $j \in I$. It follows from the above equation that $q_i < 0$ for at least one $i \in I \setminus I_F$. We can then find $x + \lambda q$ on the ray satisfying $(x + \lambda q)_i < 0$, contradicting the defining inequalities of $\mathbf{P}'$.

The same argument applies to $\mathbf{P}^{\parallel}(\epsilon q)$: if $y + \lambda p \in \mathbf{P}^{\parallel}(\epsilon q)$ for all $\lambda \geq 0$, where $p \in \ker D\mathcal{A} \cap \mathfrak{D}_F = \mathfrak{N}_F$ and $p \neq 0$, then $p_i < 0$ for some $i \in I_F$, so for λ large enough we get $y_i + \lambda p_i + \epsilon q_i < 0$.

(Alternatively, if $\mathbf{P}'$ or $\mathbf{P}^{\parallel}(\epsilon q)$ contained a ray, then so would $\mathcal{A}^{-1}(\rho_* + \epsilon\eta)$, which is a convex polytope.)

(3) follows from applying Proposition 4.11 to $\rho_* \in \operatorname{relint}(F)$. □

Let us recapitulate what we have done so far. The objective is to study the behavior of the pure functional $\mathcal{F}_p$ near one of the facets of its domain, which is $\operatorname{conv}(\Omega)$. To do so, we pick a

density ρ_* in the relative interior of a facet $F \subset \mathrm{conv}(\Omega)$, and walk inwards, in the direction of the fixed vector $\eta \in \overrightarrow{\mathrm{aff}(\Omega)}$. The value of the pure functional at $\rho_* + \epsilon\eta$ is defined as a minimization of the expectation value $\langle\Psi|W|\Psi\rangle$ over $|\Psi\rangle\langle\Psi| \in (\iota^*)^{-1}(\rho_* + \epsilon\eta)$. To understand the fiber, we observe that, with respect to the suitably chosen basis $(|E_i\rangle)_{i\in I}$ for the Hilbert space $\mathcal{H}$, the density map is, up to phases, an affine map $\mathcal{A}$ sending the squared moduli of the coefficients, which we interpret as a "classical state" or probability distribution, to a density in $\mathrm{conv}(\Omega)$. Thus, the fibers of $\mathcal{A}$ contain almost all the information needed to reconstruct the fibers of $\iota^*|_{\mathcal{P}}$. We then find out, in Propositions 4.17 and 4.18, that each element in fiber $\mathcal{A}^{-1}(\rho_* + \epsilon\eta)$ can be written as $p + \epsilon q$, where q is parametrized by a fixed convex polytope, namely $\mathbf{P}'$, and p is parametrized by a convex polytope which depends on ϵq, namely $\mathbf{P}^{\|}(\epsilon q)$. The intuition is that $\epsilon q \in \mathbf{P}'$ parametrizes "how we deviate from a (classical) state mapping to the facet", whereas $p \in \mathbf{P}^{\|}(\epsilon q)$ is supposed to be a "reference state" mapping to ρ_*. We see that this splitting works exactly because any state whose density lies close to the facet must be itself also close to a state which maps to the facet.

However, this parametrization is imperfect: it would be nice if $\mathbf{P}^{\|}(\epsilon q)$ did not depend on ϵq. On the other hand, $\mathbf{P}^{\|}(\epsilon q)$ is "almost" the same as $\mathbf{P}(0)$ as long as ϵ is small, which is clear from Eq. 4.20. Hence, we might hope to replace $\mathbf{P}^{\|}(\epsilon q)$ by $\mathbf{P}(0)$ as the parameter space for p, which can be done if we can find a surjective map $\mathbf{P}(0) \to \mathbf{P}^{\|}(\epsilon q)$ for all ϵ (small enough) and $q \in \mathbf{P}'$. In order to accomplish this, we need the following two lemmas:

Lemma 4.19. *Let P be any convex polytope and let $v^{(1)}, \cdots, v^{(n)}$ be its vertices. There exists a continuous map $\sigma : P \to \Delta_{n-1}$ (where Δ_{n-1} is the standard $(n-1)$-simplex) such that $p = \sum_{k=1}^{n} \sigma(p)_k v^{(k)}$ for all $p \in P$. Furthermore, any such map has the following property: if $f \in P$ is a face and $p \in f$, then $\sigma(p)_k = 0$ whenever $v^{(k)} \notin f$.*

Proof. Triangulate P by simplices whose vertices are vertices of P and glue together the barycentric coordinates of each simplex. This is well-defined on regions where the simplices overlap because barycentric coordinates for a simplex are unique.

The second assertion is a direct consequence of the defining property of σ. □

Remark 4.5. More succinctly, we can summarize the statement of the above lemma by saying that the "weighted average" map $\Delta_{n-1} \to P$, $t \mapsto \sum_{k=1}^{n} t_k v^{(k)}$, admits a continuous section. Such a map is called a system of *generalized barycentric coordinates*. For our purpose below, continuity is enough. Alternatively, Lemma 4.19 can be shown using Michael continuous selection theorem [106], making use of the fact that the surjection $\Delta_{n-1} \to P$ is lower hemicontinuous.

Recall that two convex polytopes P, Q are said to be *combinatorially equivalent* if their face lattices are isomorphic. In other words, there exists a bijection $\phi : \mathrm{Faces}(P) \to \mathrm{Faces}(Q)$ preserving inclusion.

Lemma 4.20. *Let P, Q be combinatorially equivalent convex polytopes and let $\phi : \mathrm{Faces}(P) \to \mathrm{Faces}(Q)$ be a lattice isomorphism. Let $v^{(1)}, \cdots, v^{(n)} \in P$ be the vertices of P. Let $\sigma : P \to \Delta_{n-1}$ be a system of generalized barycentric coordinates. Define a map*

$$\begin{aligned} R : P &\to Q \\ p &\mapsto \sum_{k=1}^{n} \sigma(p)_k \phi(v^{(k)}). \end{aligned}$$

Then R maps each face $f \subset P$ surjectively onto $\phi(f)$. In particular, R is surjective.

Proof. Let $f \subset P$ be any face. Take any point $p \in f$, then $R(p) = \sum_{k=1}^{n} \sigma(p)_k \phi(v^{(k)}) \in \mathrm{conv}(\{\phi(v^{(k)}) \mid v^{(k)} \in f\}) = \phi(f)$ due to the second part of Lemma 4.19. This shows $R(f) \subset \phi(f)$. Applying the same argument to the facets of f, we get $R(\partial f) \subset \partial\phi(f)$. In the following, we will prove that the map $R|_{\partial f} : \partial f \to \partial\phi(f)$ has degree 1 over $\mathbb{Z}/2\mathbb{Z}$ by induction on the dimension of the faces of P. Once this is shown, we can prove surjectivity of $R|_f$ onto $\phi(f)$ for any k-face by contradiction: suppose $y \in \phi(f)$ is not in the image of $R|_f$, then we can construct a continuous map $b : f \to \partial\phi(f)$ which sends $x \in f$ to the intersection of $\partial\phi(f)$ with the ray from y through $f(x)$. Clearly, $b|_{\partial f} \equiv R|_{\partial f}$. Hence, $R|_{\partial f}$ is equal to the composition

$$\partial f \hookrightarrow f \xrightarrow{b} \partial\phi(f).$$

But f is contractible, so the induced map $H_{k-1}(\partial f) \to H_{k-1}(\partial\phi(f))$ must be trivial, a contradiction.

Base case:
First, note that $R(v^{(k)}) = \sum_{k'=1}^{n} \phi(v^{(k')})\delta_{kk'} = \phi(v^{(k)})$, where again we have used the second assertion of Lemma 4.19. Let $f \subset P$ be any 1-face. Its boundary is $\{v^{(k)}, v^{(l)}\} \cong S^0$ with $k \neq l$. Similarly, $\partial\phi(f) = \{\phi(v^{(k)}), \phi(v^{(l)})\} \cong S^0$. Since $R(v^{(k)}) = \phi(v^{(k)})$ and $R(v^{(l)}) = \phi(v^{(l)})$, the map $R|_{\partial f} : \partial f \to \partial\phi(f)$ has degree 1 over $\mathbb{Z}/2\mathbb{Z}$. (We choose to work over $\mathbb{Z}/2\mathbb{Z}$ so we can ignore orientations. Over $\mathbb{Z}$, we would get ± 1 depending on the chosen orientations. Hereafter in this proof, we shall omit "over $\mathbb{Z}/2\mathbb{Z}$".)

Inductive step:
Suppose $R|_{\partial s} : \partial s \to \partial\phi(s)$ has degree 1 for all faces $s \subset P$ of dimension at most m. Take any $(m+1)$-face $f \subset P$.

Let $X_\bullet$, resp. $Y_\bullet$, denote the regular CW complexes constructed from the proper faces of f, resp. $\phi(f)$. That is, X_0 is the vertex set of f and X_k is obtained from X_{k-1} by gluing one open k-disk to X_{k-1} for each k-face of f. For each face $s \subset f$, we will denote the corresponding k-cell also by s. We can identify X_k with the union of all faces of f of dimension at most k.

Since $R(s) \subset \phi(s)$ for all faces s, the map $R : P \to Q$ is a cellular map, i.e., it satisfies $R(X_k) \subset Y_k$ for all $k = 0, \cdots, m$. Indeed, any point $x \in X_k$ is contained in a k-face $s \subset f$. So $R(x) \in R(s) \subset \phi(s) \subset Y_k$ because $\phi(s)$ is a k-face of $\phi(f)$. Therefore, R induces a map between the cellular homology groups. We will now show that $R_* : H_\bullet^{\mathrm{cell}}(X_\bullet) \to H_\bullet^{\mathrm{cell}}(Y_\bullet)$ is an isomorphism (remember that we work over $\mathbb{Z}/2\mathbb{Z}$).

Let $R_\sharp : H_k(X_k, X_{k-1}) \to H_k(Y_k, Y_{k-1})$ denote the induced map between the cellular chain complexes. The group $H_k(X_k, X_{k-1})$, resp. $H_k(Y_k, Y_{k-1})$, is free abelian over the set of k-faces of f, resp. $\phi(f)$. Take any k-face $s \subset f$. Let $\tilde{s}$ be a singular k-chain which represents s, with the latter viewed as an element of $H_k(X_k, X_{k-1})$. That is, $\tilde{s}$ is a relative cycle in (X_k, X_{k-1}) which is not a relative boundary. Then $R(\tilde{s})$ is a singular k-chain which is a relative cycle in (Y_k, Y_{k-1}) since $\partial R(\tilde{s}) = R(\partial\tilde{s}) \subset \partial\phi(s)$. Suppose $R(\tilde{s})$ were a relative boundary, then $R(\tilde{s}) = \partial c$ for some $(k+1)$-chain c in Y_k. But then $\partial R(\tilde{s}) = R(\partial\tilde{s}) = 0$, which contradicts the induction hypothesis that $R|_{\partial s} : \partial s \to \phi(\partial s)$ has $\mathbb{Z}/2\mathbb{Z}$-degree 1 because $\partial\tilde{s}$ represents the generator of $H_{k-1}(\partial s) \cong \mathbb{Z}_2$. It follows that $R_\sharp(s) = \phi(s) \in H_k(Y_k, Y_{k-1})$. Since s was an arbitrary k-face of f, and the set of all k-faces of f (resp. $\phi(f)$) generates $H_k(X_k, X_{k-1})$ (resp. $H_k(Y_k, Y_{k-1})$) freely, we have shown that $R_\sharp : H_k(X_k, X_{k-1}) \to H_{k-1}(Y_k, Y_{k-1})$ is an isomorphism.

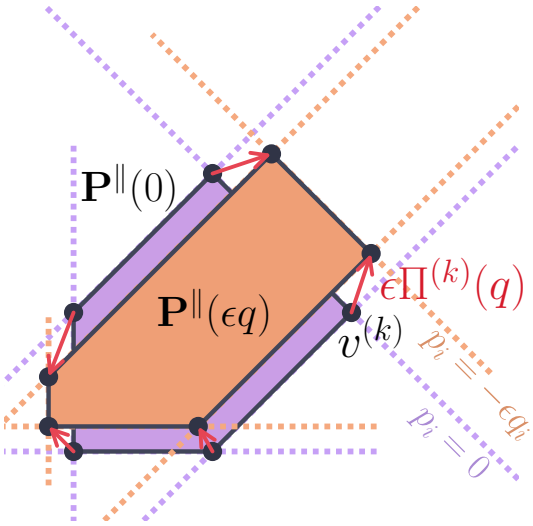


Figure 4.5 Schematic illustration of the relation between $\mathbf{P}^{\|}(0)$ (purple) and $\mathbf{P}^{\|}(\epsilon q)$ (orange) for a fixed q when ϵ is small. The dashed lines represent bounding hyperplanes corresponding to the inequalities $p_i \geq \epsilon q_i$. If ρ_* is a regular value (for the facet density map), each vertex of the polytope $\mathbf{P}^{\|}(0)$ lies in exactly $\dim \mathfrak{N}_F$, in this case two, hyperplanes. As long as ϵ is small enough, the polytope $\mathbf{P}^{\|}(\epsilon q)$ will be bounded by the translations of the same set of hyperplanes, and there will be no "new vertices".

Therefore, $R_* : H_\bullet^{\text{cell}}(X_\bullet) \to H_\bullet^{\text{cell}}(Y_\bullet)$ is an isomorphism. In particular,

$$R_* : H_m^{\text{cell}}(X_\bullet) \cong \mathbb{Z}/2\mathbb{Z} \to H_m^{\text{cell}}(Y_\bullet) \cong \mathbb{Z}/2\mathbb{Z}$$

is an isomorphism. □

Now we will apply Lemma 4.19 and Lemma 4.20 with $P = \mathbf{P}^{\|}(0)$ and $Q = \mathbf{P}'(\epsilon q)$. Before continuing, it is important to remind ourselves that many objects here depend on various previous choices, and the dependence is often suppressed. For example, the classical density map $\mathcal{A}$ is constructed from a chosen orthonormal basis $(|E_i\rangle)_{i\in I}$, and the vector spaces $\mathfrak{N}^{\|}, \mathfrak{N}'$ depend on the facet F. The convex polytopes $\mathbf{P}', \mathbf{P}^{\|}(\epsilon q)$ depend on not only ρ_*, η, but also on the choice of a particular solution $l \in (D\mathcal{A})^{-1}(\eta)$.

Proposition 4.21. *Suppose $\rho_* \in F$ is a regular value of the facet density map. Let $v^{(1)}, \cdots, v^{(n)}$ be the vertices of $\mathbf{P}^{\|}(0)$. Then there exists an $\bar{\epsilon} > 0$ and continuous maps $\Pi^{(k)} : \mathbf{P}' \to \mathfrak{N}^{\|}$, $k = 1, \cdots, n$, such that $v^{(k)} \mapsto v^{(k)} + \epsilon\Pi^{(k)}(q)$ gives a face lattice isomorphism between $\mathbf{P}^{\|}(0)$ and $\mathbf{P}^{\|}(\epsilon q)$ for all $\epsilon \in [0, \bar{\epsilon}]$ and $q \in \mathbf{P}'$ (see Fig. 4.5).*

Proof. The intuition is very simple: the convex polytope $\mathbf{P}^{\|}(\epsilon q)$ is defined to be the subset of $\bar{\mathcal{A}}_F^{-1}(\rho_*)$ bounded by the $|I_F|$ inequalities given by $y_i \geq -\epsilon q_i$, where $i \in I_F$. When we perturb this system of inequalities by slowly increasing ϵ from zero, we are effectively translating the bounding hyperplanes of the polytope. If the initial ($\epsilon = 0$) arrangement of the hyperplanes is not too "degenerate", then we expect the vertices of $\mathbf{P}^{\|}(\epsilon q)$ to be "stable" for small ϵ. We will make this more precise in the following.

For $i \in I_F$, define

$$H_i(\epsilon q) := \{y \in \bar{\mathcal{A}}_F^{-1}(\rho_*) \mid y_i = -\epsilon q_i\} \subset \bar{\mathcal{A}}_F^{-1}(\rho_*).$$

Intuitively, each $H_i(\epsilon q)$ should be a hyperplane in $\bar{\mathcal{A}}_F^{-1}(\rho_*)$, but a priori it can happen that $H_i(\epsilon q) = \varnothing$ or $H_i(\epsilon q) = \bar{\mathcal{A}}_F^{-1}(\rho_*)$. Nevertheless, the latter is impossible at $\epsilon = 0$: if $H_i(0)$ were all of $\bar{\mathcal{A}}_F^{-1}(\rho_*)$ for some $i \in I_F$, then $\bar{\mathcal{A}}_F^{-1}(\rho_*)$ is disjoint from $\mathrm{relint}(\Delta_{I_F})$, which is ruled out since $\rho_* \in \mathrm{relint}(F)$ (see the proof of Proposition 4.18). Thus, each $H_i(0)$ is either empty or a hyperplane in $\bar{\mathcal{A}}_F^{-1}(\rho_*)$. By compactness of $\mathbf{P}'$ (Proposition 4.18), this continues to be true for $H_i(\epsilon q)$ for all $q \in \mathbf{P}'$ as long as ϵ is small enough, say $\epsilon \in [0, \bar{\epsilon}_1]$ (where $\bar{\epsilon}_1 > 0$).

For each $k \in \{1, \cdots, n\}$, let $I_k \subset I_F$ be the index set of hyperplanes $H_i(0)$ that contain $v^{(k)}$. That is, $I_k := \{i \in I \mid v_i^{(k)} = 0\}$. We claim that $|I_k| = \dim \mathfrak{N}_F$ for all k. Indeed, suppose $|I_k| > \dim \mathfrak{N}_F$, then $\rho_* = \sum_{i \in I_F} v_i^{(k)} \omega_i$ lies in the convex hull of $|I_F| - |I_k| < \dim \mathcal{H}_F - \dim \mathfrak{N}_F = \dim \mathrm{conv}(\Omega_F) + 1$ weights, which implies that ρ_* is a critical value (Proposition 4.6), contradicting our assumption. On the other hand, suppose $|I_k| < \dim \mathfrak{N}_F$. We can then find a nonzero $\delta \in \mathfrak{N}_F$ such that $v^{(k)} + \mathbb{R}\delta \subset \bigcap_{i \in I_k} H_i(0)$. But $v_i^{(k)} > 0$ for all $i \in I_F \setminus I_k$, so we can find an open neighborhood U of $0 \in \mathbb{R}$ such that $v^{(k)} + U\delta \subset \mathbf{P}^{\parallel}(0)$, which implies that $v^{(k)}$ is not extremal in $\mathbf{P}^{\parallel}(0)$ and hence not a vertex. So each vertex of $\mathbf{P}^{\parallel}(0)$ is contained in exactly $\dim \mathfrak{N}_F$ bounding hyperplanes. In technical terms, the convex polytope $\mathbf{P}^{\parallel}(0)$ is *simple*.

Let $I_{\mathrm{rel}} := \bigcup_{k=1}^n I_k$ be the set of indices of the "relevant" hyperplanes, for these are the ones that actually bound the polytope $\mathbf{P}^{\parallel}(0)$, and write $I_{\mathrm{irrel}} := I_F \setminus I_{\mathrm{rel}}$. Again by compactness of $\mathbf{P}'$, we can find an $\bar{\epsilon}_2 \in (0, \bar{\epsilon}_1]$ such that for all $\epsilon \in [0, \bar{\epsilon}_2]$, for all $q \in \mathbf{P}'$, for all $k \in \{1, \cdots, n\}$, and for all $j \in I_{\mathrm{irrel}}$, it holds that $H_j(\epsilon q) \cap \bigcap_{i \in I_k} H_i(\epsilon q) = 0$. (In words: as long as ϵ is small enough, the irrelevant hyperplanes will stay irrelevant.)

Now, for each $k \in \{1, \cdots, n\}$ and $q \in \mathbf{P}'$, there is a unique $\Pi^{(k)}(q) \in \mathfrak{N}_F$ that solves $\Pi^{(k)}(q)_i = q_i$ for all $i \in I_k$. It is not hard to see that $\Pi^{(k)}(q)$ depends linearly on q. Then $(v^{(k)} + \epsilon \Pi^{(k)}(q))_{k=1}^n$ are exactly the vertices of the convex polytope $\mathbf{P}^{\parallel}(\epsilon q)$ whenever $\epsilon \in [0, \bar{\epsilon}_3]$ for some $\bar{\epsilon}_3 \in (0, \bar{\epsilon}_2]$.

Finally, we show that the assignment $v^{(k)} \mapsto v^{(k)} + \epsilon \Pi^{(k)}(q)$ gives a lattice isomorphism between $\mathbf{P}^{\parallel}(0)$ and $\mathbf{P}^{\parallel}(\epsilon q)$ whenever $\epsilon \in [0, \bar{\epsilon})$ for some $\bar{\epsilon} \in (0, \bar{\epsilon}_3]$. To be more precise, we define a map $\phi : \mathrm{Faces}(\mathbf{P}^{\parallel}(0)) \to \mathrm{Faces}(\mathbf{P}^{\parallel}(\epsilon q))$ in the following way: if $f = \mathrm{conv}(\{v^{(k_1)}, \cdots v^{(k_m)}\})$ is a face of $\mathbf{P}^{\parallel}(\epsilon q)$, we set

$$\phi(f) = \mathrm{conv}(\{v^{(k_1)} + \epsilon \Pi^{(k_1)}(q), \cdots, v^{(k_m)} + \epsilon \Pi^{(k_m)}(q)\}).$$

It should be clear that $\phi(f) \subset \mathbf{P}^{\parallel}(\epsilon q)$ is a face. Indeed, take $\alpha \in \mathbb{R}^{I_F}$ and $c \in \mathbb{R}$ such that

$$\forall 1 \le k \le n : \begin{cases} v^{(k)} \in f \Rightarrow \langle \alpha, v^{(k)} \rangle = c \\ v^{(k)} \notin f \Rightarrow \langle \alpha, v^{(k)} \rangle > c. \end{cases}$$

Then

$$\langle \alpha, v^{(k)} + \epsilon \Pi^{(k)}(q) \rangle = \langle \alpha, v^{(k)} \rangle + \epsilon \langle \alpha, \Pi^{(k)}(q) \rangle .$$

The second term does not depend on k as long as $v^{(k)} \in f$. Therefore, we have

$$\forall 1 \le k \le n : \begin{cases} v^{(k)} \in f \Rightarrow \langle \alpha, v^{(k)} + \epsilon \Pi^{(k)}(q) \rangle = c' \\ v^{(k)} \notin f \Rightarrow \langle \alpha, v^{(k)} + \epsilon \Pi^{(k)}(q) \rangle > c' \end{cases}$$

for some c' as long as ϵ is small enough, showing that $\phi(f) \subset \mathbf{P}^{\parallel}(\epsilon q)$ is a face. Since $\mathbf{P}'$ is compact and the number of faces is finite, there exists an $\bar{\epsilon} \in (0, \bar{\epsilon}_3]$ such that $\phi : \mathrm{Faces}(\mathbf{P}^{\parallel}(0)) \to \mathrm{Faces}(\mathbf{P}^{\parallel}(\epsilon q))$ is a lattice isomorphism for all $q \in \mathbf{P}'$ and $\epsilon \in [0, \bar{\epsilon}]$. □

Corollary 4.22. *Let $\sigma : \mathbf{P}^{\parallel}(0) \to \Delta_{n-1}$ be any continuous system of generalized barycentric coordinates. Consider the "parametrization map"*

$$\begin{aligned} \Upsilon : [0,\bar{\epsilon}] \times \mathbf{P}' \times \mathbf{P}^{\parallel}(0) &\to \operatorname{aff}(\Delta_{I_F}) \\ (\epsilon, q, p) &\mapsto \sum_{k=1}^{n} \sigma(p)_k (v^{(k)} + \epsilon \Pi^{(k)}(q)) = p + \epsilon \sum_{k=1}^{n} \sigma(p)_k \Pi^{(k)}(q). \end{aligned} \tag{4.21}$$

For any (ϵ, q), the map $\Upsilon(\epsilon, q, \cdot)$ maps $\mathbf{P}^{\parallel}(0)$ surjectively onto $\mathbf{P}^{\parallel}(\epsilon q)$. Moreover, Υ depends continuously on p and q, and affinely on ϵ.

Proof. Apply Lemma 4.20 to the lattice isomorphism given by Proposition 4.21. □

Using Corollary 4.13, we obtain the main result of this section:

Theorem 4.23 (facet–off-facet decomposition of the pure functional)**.** *Let $\rho_* \in \operatorname{relint}(F)$ be a density on the facet which is a regular value and let $\eta \in \overrightarrow{\operatorname{aff}(\Omega)}$ be an inward vector. Choose some $l \in D\mathcal{A}^{-1}(\eta)$, and let $\mathbf{P}'$ and $\mathbf{P}^{\parallel}$ be defined as in Proposition 4.17 according to the choice of l. Let $\bar{\epsilon}$ and $\Pi^{(k)}$ be given by Proposition 4.21. Then*

$$\mathcal{F}_p(\rho_* + \epsilon\eta) = \min_{\xi \in (S^1)^I} \min_{q \in \mathbf{P}'} \min_{p \in \mathbf{P}^{\parallel}(0)} \Big(Q_{ff}(\epsilon; \xi, p, q) + Q_{oo}(\epsilon; \xi, p, q) + Q_{fo}(\epsilon; \xi, p, q) \Big) \tag{4.22}$$

for $\epsilon \in [0, \bar{\epsilon}]$, where the subscripts stand for "facet–facet", "off-facet–off-facet" and "facet–off-facet" respectively. Individually, these terms are (suppressing the dependence on ϵ, ξ, p, q)

$$\begin{aligned} Q_{ff} &:= \sum_{i,j \in I_F} \bar{\xi}_i \xi_j W_{ij} \sqrt{p_i + \epsilon q_i + \epsilon \sum_{k=1}^{n} \sigma(p)_k \Pi^{(k)}(q)_i} \sqrt{p_j + \epsilon q_j + \epsilon \sum_{k=1}^{n} \sigma(p)_k \Pi^{(k)}(q)_j} \\ Q_{fo} &:= 2\mathrm{Re} \sum_{i \in I_F} \sum_{j \in I \setminus I_F} \bar{\xi}_i \xi_j W_{ij} \sqrt{p_i + \epsilon q_i + \epsilon \sum_{k=1}^{n} \sigma(p)_k \Pi^{(k)}(q)_i} \cdot \sqrt{\epsilon} \sqrt{q_j} \\ Q_{oo} &:= \sum_{i,j \in I \setminus I_F} \bar{\xi}_i \xi_j W_{ij} \epsilon \sqrt{q_i q_j}. \end{aligned} \tag{4.23}$$

We will write $Q(\epsilon; \xi, p, q)$ for the sum of the three, so that

$$\mathcal{F}_p(\rho_* + \epsilon\eta) = \min_{\xi \in (S^1)^I} \min_{q \in \mathbf{P}'} \min_{p \in \mathbf{P}^{\parallel}(0)} Q(\epsilon; \eta, p, q). \tag{4.24}$$

Looking at Eqs. (4.22) and (4.23), we can already anticipate where the boundary force might stem from: the facet–off-facet term $Q_{fo}(\epsilon; \xi, p, q)$ contains a factor of $\sqrt{\epsilon}$. If there were no minimizations in Eq. (4.22), we would know immediately at least one source of the boundary force (it is not clear whether the rest of the terms contribute). It then remains to carry out the minimizations and investigate how the $\sqrt{\epsilon}$ behavior survives after doing so.

Step III: Exchanging $\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}$ and min

At first sight, it is not clear whether and how one can differentiate the pure functional $\mathcal{F}_p$, which is defined as a constrained minimization (Definition 2.6). Now that we have cast $\mathcal{F}_p$ into a slightly nicer form, namely Eq. (4.24), it is still not obvious how taking the derivative interacts with minimization.

There is a convenient tool for dealing with such problems, which we have used once in Section 3.5.1. Namely, Danskin's theorem [102] asserts that if $\tilde{Q} : [0, \varepsilon) \times X \to \mathbb{R}$ is a continuous function, where X is a compact metric space, and $\tilde{\mathcal{F}}(\epsilon) := \min_{x \in X} \tilde{Q}(\epsilon; x)$, then $\tilde{\mathcal{F}}'(0_+)$ exists and is equal to $\min_{x \in X_0} \tilde{Q}'(0; x)$ provided that $\tilde{Q}$ is continuously differentiable with respect to ϵ, where $X_0 := \arg\min_{x \in X} \tilde{Q}(0; x)$. In our case, the objective function $Q = Q_{ff} + Q_{fo} + Q_{oo}$ (defined in Eq. (4.23)) is not necessarily continuously differentiable with respect to $\sqrt{\epsilon}$. To see what could go wrong, note that $Q(\epsilon; \xi, p, q)$ contains terms of the form $\sqrt{A + B\epsilon}$, where A, B depend continuously on p and q. Such a function is not continuously differentiable at points where $A = \epsilon = 0$ and $B > 0$. Fortunately, the No Mixing Lemma (Lemma 4.9) will help us conclude that these discontinuities will not matter *at minimizers of the constrained search*. This is the content of Lemma 4.25, which moreover confirms our intuition that only Q_{fo} should matter for the boundary force.

Hereafter, a prime ($'$) on Q, Q_{fo}, Q_{ff}, or Q_{oo} will indicate differentiation with respect to $\sqrt{\epsilon}$. To save space, we will often denote (ξ, p, q) collectively by the variable $x \in (S^1)^I \times \mathbf{P}^{\parallel}(0) \times \mathbf{P}' =: X$. With this notation, Eq. (4.22) becomes

$$\mathcal{F}_p(\rho_* + \epsilon\eta) = \min_{x \in X} Q(\epsilon; x).$$

At each ϵ, the set of minimizers will be denoted by $X_\epsilon \subset X$. That is, $\mathcal{F}_p(\rho_* + \epsilon\eta) = Q(\epsilon; x)$ for all ϵ and $x \in X_\epsilon$.

Lemma 4.24.

$$Q'(0; x) = Q'_{fo}(0; x) = 2 \operatorname{Re} \sum_{i \in I_F} \sum_{j \in I \setminus I_F} \bar{\xi}_i \xi_j W_{ij} \sqrt{p_i} \sqrt{q_j} \tag{4.25}$$

for any $x = (\xi, p, q) \in X_0$.

Proof. We will show $Q'_{ff}(0; x) = 0$ and $Q'_{oo}(0; x) = 0$.

The latter is clear (see Eq. (4.23)):

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}} Q_{oo}(0; x) = \frac{\mathrm{d}}{\mathrm{d}z}\Big|_{z=0} \Big(z^2 \sum_{i,j \in I \setminus I_F} \cdots \Big) = 0 \tag{4.26}$$

The former, on the other hand, requires slightly more work to establish. Observe, with Q_{ff} given in Eq. (4.23), that the potentially nonzero contributions to Q'_{ff} arise only when differentiating a radical which vanishes at $\epsilon = 0$. Let $|\Phi(\epsilon)\rangle$ denote the state corresponding to (ϵ, ξ, p, q), that is,

$$|\Phi(\epsilon)\rangle = \sum_{i \in I_F} \xi_i \sqrt{y_i + q_i + \epsilon \sum_{k=1}^{n} \sigma(p)_k \Pi^{(k)}(q)_i}. \tag{4.27}$$

Then

$$Q'_{ff}(0; \xi, p, q) = 2 \operatorname{Re} \langle \Phi(0) | W | \Phi'(0) \rangle ,$$

where $|\Phi'(\epsilon)\rangle := \frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}|\Phi(\epsilon)\rangle$. Now, for each $i \in I_F$, if $\langle E_i|\Phi(0)\rangle \neq 0$, then $\langle E_i|\Phi'(0)\rangle = 0$ because the function $\epsilon \mapsto \sqrt{A\epsilon + B}$ is smooth for $B > 0$, and consequently the derivative with respect to $\sqrt{\epsilon}$ must vanish. In other words, $|\Phi(0)\rangle$ and $|\Phi'(0)\rangle$ are strongly orthogonal. Since $|\Phi(0)\rangle$ is a minimizer of the constrained search at $\epsilon = 0$ (i.e., at ρ_*), we have $\langle\Phi(0)|W|\Phi'(0)\rangle = 0$ by Lemma 4.10. □

Lemma 4.25. *Let $(\epsilon_\nu)_\nu$ be a sequence converging to 0 with $\epsilon_\nu \in [0, \bar{\epsilon})$. Let $(x_\nu)_\nu = (\xi^\nu, p^\nu, q^\nu)_\nu$ be any sequence in X converging to $x^* = (\xi^*, p^*, q^*) \in X$. If $x^* \in X_0$ (i.e., x^* is a minimizer at $\epsilon = 0$), then*

$$\lim_{\nu\to\infty} Q'(\epsilon_\nu; x_\nu) = Q'_{fo}(0; x^*). \tag{4.28}$$

That is, Q is continuously differentiable at $(0, x^*)$.

Proof. To show Eq. (4.28), note that the only terms that might ruin the equality are those arising from differentiating (with respect to $\sqrt{\epsilon}$) the radical

$$\sqrt{p_j + \epsilon q_j + \epsilon \sum_{k=1}^{n} \sigma(p)_k \Pi^{(k)}(q)_j} \tag{4.29}$$

for some $j \in I_F$ for which the radical itself vanishes in the limit $\nu \to \infty$, since everything else is continuous after differentiation at $\epsilon = 0$ (see Eq. (4.23)). In the rest of this proof, we will show that while the derivative of such a "problematic radical" need not converge to zero in the limit $\nu \to \infty$, the derivative itself never contributes to the limit because it is bounded and it is always multiplied by something that converges to zero.

Fix a $j \in I_F$ such that

$$\underbrace{p_j^\nu + \epsilon_\nu q_j^\nu + \epsilon_\nu \sum_{k=1}^{n} \sigma(p^\nu)_k \Pi^{(k)}(q^\nu)_j}_{=:\epsilon_\nu G_j(p^\nu, q^\nu)} \xrightarrow{\nu\to\infty} 0. \tag{4.30}$$

We want to study the derivative

$$\mathcal{N}_j^\nu := \frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=\epsilon_\nu} \sqrt{p_j^\nu + \epsilon G_j(p^\nu, q^\nu)}, \tag{4.31}$$

where $\mathcal{N}$ stands for "nasty". In $Q'_{ff}(\epsilon_\nu; x_\nu)$, each of these factors always comes multiplied by

$$\sum_{i\in I_F} \bar{\xi}_i^\nu W_{ij} \sqrt{p_i^\nu + \epsilon_\nu q_i^\nu + \epsilon_\nu \sum_{k=1}^{n} \sigma(p^\nu)_k \Pi^{(k)}(q^\nu)_i}$$

(see Eq. (4.23)), which converges to zero by the No Mixing Lemma (Lemma 4.9). In Q'_{fo}, on the other hand, it is always multiplied by $\sum_{i\in I\setminus I_F} \bar{\xi}_i^\nu W_{ij} \sqrt{\epsilon_\nu}\sqrt{q_i^\nu}$, which also converges to zero. Therefore, if we can show that the sequence $(\mathcal{N}_\nu^j)_\nu$ is bounded, all the anomalous terms will vanish.

Now, there are several possibilities for the derivative (4.31). First, suppose $p_j^\nu + \epsilon_\nu G_j(p^\nu, q^\nu) = 0$. If $\epsilon_\nu = 0$, then $p_j^\nu = 0$, which implies $G_j(p^\nu, q^\nu) \geq 0$ (otherwise we would be able make the radicand negative by slightly increasing ϵ_ν). Consequently, we have $\mathcal{N}_j^\nu = \sqrt{G_j(p^\nu, q^\nu)}$,

which is bounded because G_j is continuous and p^ν, q^ν belong to $\mathbf{P}^{\|}(0)$ and $\mathbf{P}'$ respectively, which are both compact. If $\epsilon_\nu \in (0, \bar{\epsilon})$, then $G_j(p^\nu, q^\nu) = 0$ because otherwise the radicand could be made negative by perturbing ϵ_ν. In this case, $\mathcal{N}_j^\nu = 0$.

Next, we deal with the case $p_j^\nu + \epsilon_\nu G_j(p^\nu, q^\nu) > 0$, for which

$$\mathcal{N}_j^\nu = \frac{\sqrt{\epsilon_\nu} G_j(p^\nu, q^\nu)}{\sqrt{p_j^\nu + \epsilon_\nu G_j(p^\nu, q^\nu)}}.$$

If $p_j^\nu = 0$, then $\mathcal{N}_j^\nu$ is simply $\sqrt{G_j(p^\nu, q^\nu)}$. Otherwise, we separate two cases depending on the sign of $G_j(p^\nu, q^\nu)$. If $G_j(p^\nu, q^\nu) \geq 0$, then $0 \leq \mathcal{N}_j^\nu \leq \sqrt{G_j(p^\nu, q^\nu)}$, so $\mathcal{N}_j^\nu$ is bounded. Suppose $G_j(p^\nu, q^\nu) < 0$. Since $p_j^\nu + \bar{\epsilon} G_j(p^\nu, q^\nu) \geq 0$, we have $G_j(p^\nu, q^\nu) \geq -p_j^\nu/\bar{\epsilon}$, which holds for all ν. It follows that

$$\begin{aligned}
0 \leq -\frac{\sqrt{\epsilon_\nu} G_j(p^\nu, q^\nu)}{\sqrt{p_j^\nu + \epsilon_\nu G_j(p^\nu, q^\nu)}} &= \sqrt{-G_j(p^\nu, q^\nu)} \sqrt{\frac{-\epsilon_\nu G_j(p^\nu, q^\nu)}{p_j^\nu + \epsilon_\nu G_j(p^\nu, q^\nu)}} \\
&= \sqrt{-G_j(p^\nu, q^\nu)} \sqrt{-1 + \frac{p_j^\nu}{p_j^\nu + \epsilon_\nu G_j(p^\nu, q^\nu)}} \\
&\leq \sqrt{-G_j(p^\nu, q^\nu)} \sqrt{-1 + \frac{p_j^\nu}{p_j^\nu - \frac{\epsilon_\nu}{\bar{\epsilon}} p_j^\nu}} = \sqrt{-G_j(p^\nu, q^\nu)} \sqrt{-1 + \frac{\bar{\epsilon}}{\bar{\epsilon} - \epsilon_\nu}} \to 0.
\end{aligned}$$

This concludes the proof that none of the anomalous terms survive the limit $\nu \to \infty$. □

The following is adapted from the proof of Danskin's theorem [102].

Theorem 4.26.

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) = \min_{x \in X_0} Q'_{fo}(0; x) \tag{4.32}$$

Proof. Let $(\epsilon_\nu)_\nu$ be a sequence converging to 0 such that $\epsilon_\nu \in (0, \bar{\epsilon})$. For each ν, pick a minimizer $x_\nu \in X_{\epsilon_\nu}$. Also pick $x \in X_0$. Then

$$\begin{aligned}
\frac{\mathcal{F}_p(\rho_* + \epsilon_\nu \eta) - \mathcal{F}_p(\rho_*)}{\sqrt{\epsilon_\nu}} &= \frac{Q(\epsilon_\nu; x_\nu) - Q(0; x)}{\sqrt{\epsilon_\nu}} \\
&= \frac{1}{\sqrt{\epsilon_\nu}} \Big(Q(\epsilon_\nu; x_\nu) - Q(\epsilon_\nu; x) + Q(\epsilon_\nu; x) - Q(0; x) \Big) \\
&\leq \frac{Q(\epsilon_\nu; x) - Q(0; x)}{\sqrt{\epsilon_\nu}} = Q'(b_\nu \epsilon_\nu; x),
\end{aligned} \tag{4.33}$$

where $b_\nu \in [0, 1]$ and we used the Mean Value Theorem in the last equality. Taking $\limsup_{m \to \infty}$ of Eq. (4.33), we get

$$\limsup_{\nu \to \infty} \frac{\mathcal{F}_p(\rho_* + \epsilon_\nu \eta) - \mathcal{F}_p(\rho_*)}{\sqrt{\epsilon_\nu}} \leq Q'(0; x) = Q'_{fo}(0; x), \tag{4.34}$$

where we used the continuity of $\epsilon \mapsto Q'(\epsilon; x)$ and the last equality is due to Lemma 4.24. Note that Eq. (4.34) holds for any $x \in X_0$.

By compactness of X, we extract a subsequence $(x_{\nu_l})_l$ such that $x_{\nu_l} \to x^*$ for some $x^* \in X$. We claim that $x^* \in X_0$. Indeed, for any $y \in X$ it holds that

$$Q(0;y) = \lim_{l\to\infty} Q(\epsilon_{\nu_l};y) \geq \lim_{l\to\infty} Q(\epsilon_{\nu_l};x_{\nu_l}) = Q(0;x^*).$$

By further extracting a subsequence if necessary, we may assume that

$$\lim_{l\to\infty} \frac{Q(\epsilon_{\nu_l};x_{\nu_l}) - Q(0;x^*)}{\sqrt{\epsilon_{\nu_l}}} = \liminf_{\nu\to\infty} \frac{Q(\epsilon_\nu;x_\nu) - Q(0;x^*)}{\sqrt{\epsilon_\nu}}.$$

We now bound the difference quotient from below. We have

$$\begin{aligned} \frac{Q(\epsilon_{\nu_l};x_{\nu_l}) - Q(0;x^*)}{\sqrt{\epsilon_{\nu_l}}} &= \frac{1}{\sqrt{\epsilon_{\nu_l}}}\Big(Q(\epsilon_{\nu_l};x_{\nu_l}) - Q(0;x_{\nu_l}) + Q(0;x_{\nu_l}) - Q(0;x^*)\Big) \\ \geq \frac{Q(\epsilon_{\nu_l};x_{\nu_l}) - Q(0;x_{\nu_l})}{\sqrt{\epsilon_{\nu_l}}} &= Q'(\tilde{b}_l\epsilon_{\nu_l};x_{\nu_l}), \end{aligned} \tag{4.35}$$

for some $(\tilde{b}_l)_l$ with $\tilde{b}_l \in [0,1]$, where again we have used the Mean Value Theorem. Now we apply Lemma 4.25 to the sequences $(\tilde{b}_l\epsilon_{\nu_l})_l$ and $(x_{\nu_l})_l$ to conclude

$$\liminf_{\nu\to\infty} \frac{\mathcal{F}_p(\rho_* + \epsilon_\nu\eta) - \mathcal{F}_p(\rho_*)}{\sqrt{\epsilon_\nu}} \geq \lim_{l\to\infty} Q'(\tilde{b}_l\epsilon_{\nu_l};x_{\nu_l}) = Q'_{fo}(0;x^*). \tag{4.36}$$

This, together with Eq. (4.34), proves the theorem. □

Step IV: Relaxation of P′; Completing the Proof

By Theorem 4.26 and Lemma 4.24, we now have

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) = \min_{(\xi,p,q)\in X_0} \Big(2\,\mathrm{Re}\sum_{i\in I_F}\sum_{j\in I\setminus I_F} \bar{\xi}_i\xi_j W_{ij}\sqrt{p_i}\sqrt{q_j}\Big), \tag{4.37}$$

where X_0 denotes the set of minimizers of $Q(0;x)$:

$$X_0 := \operatorname*{argmin}_{(\xi,p,q)\in X} \left(\sum_{i,j\in I_F} \bar{\xi}_i\xi_j W_{ij}\sqrt{p_i}\sqrt{p_j}\right) \subset (S^1)^I \times \mathbf{P}^{\parallel}(0) \times \mathbf{P}' =: X. \tag{4.38}$$

Since

$$Q(0;\xi,p,q) = Q_{ff}(0;\xi,p,q) = \sum_{i,j\in I_F} \bar{\xi}_i\xi_j W_{ij}\sqrt{p_i}\sqrt{p_j}$$

(see Eq. (4.23)) does not depend on q and ξ_i for $i \in I\setminus I_F$, the set X_0 decomposes as a product $X_0 = \tilde{X}_0 \times \mathbf{P}' \times (S^1)^{I\setminus I_F}$, where $\tilde{X}_0 \subset \mathbf{P}^{\parallel}(0) \times (S^1)^{I_F}$ consists of all pairs (p,ζ) that minimize Q_{ff} at $\epsilon = 0$. In other words, the minimization in Eq. (4.37) can be split into two steps:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) &= \min_{(p,\zeta)\in\tilde{X}_0} \min_{q\in\mathbf{P}',\chi\in(S^1)^{I\setminus I_F}} 2\,\mathrm{Re}\left(\sum_{i\in I_F}\sum_{j\in I\setminus I_F} \bar{\zeta}_i\chi_j W_{ij}\sqrt{p_i}\sqrt{q_j}\right) \\ &= \min_{(p,\zeta)\in\tilde{X}_0} \min_{q\in\mathbf{P}',\chi\in(S^1)^{I\setminus I_F}} 2\,\mathrm{Re}\sum_{j\in I\setminus I_F} \chi_j\sqrt{q_j}\sum_{i\in I_F} \bar{\zeta}_i\sqrt{p_i}W_{ij}, \end{aligned} \tag{4.39}$$

where we have split the collection of phases $\xi = (\xi_i)_{i\in I}$ into the pair (ζ, χ) with $\zeta \in (S^1)^{I_F}$ and $\chi \in (S^1)^{I\setminus I_F}$. We will deal with the second (inner) minimization in Eq. (4.39), that is, the one over $(q, \chi) \in \mathbf{P}' \times (S^1)^{I\setminus I_F}$.

First, note that the minimization over the phases $(\chi_j)_{j\in I\setminus I_F}$ is trivial since they appear independently. Consequently, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0}\mathcal{F}_p(\rho_* + \epsilon\eta) &= -2\max_{(p,\zeta)\in\tilde{X}_0}\max_{q\in\mathbf{P}'}\sum_{j\in I\setminus I_F}\sqrt{q_j}\left|\sum_{i\in I_F}\bar{\zeta}_i\sqrt{p_i}W_{ij}\right| \\ &= -2\max_{(p,\zeta)\in\tilde{X}_0}\max_{q\in\mathbf{P}'}\sum_{j\in I\setminus I_F}\sqrt{q_j}\,|\langle\Phi(p,\zeta)|W|E_j\rangle|\,,\end{aligned} \tag{4.40}$$

where we have defined $|\Phi(p,\zeta)\rangle := \sum_{i\in I_F}\zeta_i\sqrt{p_i}\,|E_i\rangle$. Hence, we are left with the minimization over $q \in \mathbf{P}'$. We noted before (see Proposition 4.18) that $\mathbf{P}'$ is a convex polytope. However, its shape could still be fairly complicated in general, and it is a priori not clear whether it is easy to carry out the minimization over $q \in \mathbf{P}'$ in Eq. (4.40). On the other hand, the objective function depends only on $(q_j)_{j\in I\setminus I_F}$, so not all defining constraints of the polytope $\mathbf{P}'$ are necessarily relevant. To make this more precise, let pr^F denote the projection that sets the entries y_i of a vector $y \in \mathfrak{D} \subset \mathbb{R}^I$ to zero whenever $i \in I_F$. That is,

$$\mathrm{pr}^F(y)_i = \begin{cases} 0 & i \in I_F \\ y_i & i \in I \setminus I_F. \end{cases} \tag{4.41}$$

Then we may replace $\mathbf{P}'$ with $\mathrm{pr}^F(\mathbf{P}')$ in Eq. (4.40). The hope now is that $\mathrm{pr}^F(\mathbf{P}')$ admits a simpler characterization, which is fortunately true thanks to the following lemma:

Lemma 4.27 (Relaxation of the polytope $\mathbf{P}'$)**.** *Let $S \in V$ be an external potential such that $\langle\eta, S\rangle = 1$ and $\langle\overrightarrow{\mathrm{aff}(F)}, S\rangle = 0$. Then $\mathrm{pr}^F(\mathbf{P}') = \mathrm{pr}^F(\mathbf{P}'_{\mathrm{relax}})$, where*

$$\mathbf{P}'_{\mathrm{relax}} := \Big\{q \in \mathfrak{D} \mid \langle D\mathcal{A}(q), S\rangle = 1, \forall i \in I \setminus I_F : q_i \geq 0\Big\}.$$

In other words, we may replace the constraint $D\mathcal{A}(q) = \eta$, which was used to define the set $\mathbf{P}'$, with the (in general) much weaker constraint $\langle D\mathcal{A}(q), S\rangle = 1$, as long as we only care about the image under the projection pr^F.

Proof. Recall that $\mathbf{P}' := \{l + q \mid q \in \mathfrak{N}', \forall i \in I \setminus I_F : q_i \geq 0\}$, where $\mathfrak{N}' \subset \mathfrak{N}$ is a chosen and fixed complement of $\mathfrak{N}_F \subset \mathfrak{N}$. We will prove

$$\mathrm{pr}^F(l + \mathfrak{N}') = \mathrm{pr}^F\underbrace{\Big\{q' \in \mathfrak{D} \mid \big\langle D\mathcal{A}(q'), S\big\rangle = 1\Big\}}_{=:\mathcal{S}}. \tag{4.42}$$

That is, we first ignore the inequalities $q_i \geq 0$ so that both sets become affine spaces in $\mathrm{pr}^F(\mathfrak{D})$. Once Eq. (4.42) is proved, the original statement will follow immediately because $\mathrm{pr}^F(\mathbf{P}')$ and $\mathrm{pr}^F(\mathbf{P}'_{\mathrm{relax}})$ are further cut out from $\mathrm{pr}^F(l + \mathfrak{N}')$ and $\mathrm{pr}^F(\mathcal{S})$ respectively by the same set of inequalities, namely $q_i \geq 0$ for $i \in I \setminus I_F$.

Take any $q \in l + \mathfrak{N}'$. Then $\langle D\mathcal{A}(q), S\rangle = \langle\eta, S\rangle = 1$. This shows $\mathrm{pr}^F(l + \mathfrak{N}') \subset \mathrm{pr}^F(\mathcal{S})$.

The strategy to show the inclusion in the other direction is to show that $\mathrm{pr}^F(l + \mathfrak{N}')$ and $\mathrm{pr}^F(\mathcal{S})$ have the same dimension as affine spaces. Clearly, $\dim(\mathrm{pr}^F(\mathcal{S})) = |I| - |I_F| - 1$.

We claim that pr^F is injective on $\mathfrak{N}'$. Indeed, suppose $\mathrm{pr}^F(q) = 0$ for some $q \in \mathfrak{N}'$, then $q_i \neq 0$ only when $i \in I_F$. But this means $q \in \mathfrak{N}_F$. Since $\mathfrak{N}_F \cap \mathfrak{N}' = 0$, we have $q = 0$. This implies $\dim \mathrm{pr}^F(l + \mathfrak{N}') = \dim \mathrm{pr}^F(\mathfrak{N}') = \dim \mathfrak{N}'$.

It remains to show that $\dim \mathfrak{N}' = |I| - |I_F| - 1$. But this follows directly from Proposition 4.16, which asserts

$$\begin{aligned} \dim \mathfrak{N} &= |I| - \dim \mathrm{conv}(\Omega) - 1 \\ \dim \mathfrak{N}_F &= |I_F| - \dim F - 1. \end{aligned}$$

Taking the difference then gives $\dim \mathfrak{N}' = |I| - |I_F| - 1$. Therefore, $\dim(\mathrm{pr}^F(l + \mathfrak{N}')) = \dim(\mathrm{pr}^F(\mathcal{S}))$. □

Corollary 4.28. $\mathrm{pr}^F(\mathbf{P}') = \left\{ q \in \mathbb{R}^{I \setminus I_F} \mid q_i \geq 0, \sum_{i \in I \setminus I_F} q_i \langle \omega_i - \gamma, S \rangle = 1 \right\}$ *for any* $\gamma \in \mathrm{aff}(F)$.

Proof. By Lemma 4.27, we can replace the left hand side by $\mathrm{pr}^F(\mathbf{P}'_{\mathrm{relax}})$. Take any $q \in \mathbf{P}'_{\mathrm{relax}}$, then $\sum_{i \in I} q_i \langle \omega_i, S \rangle = 1$. Since $\langle \omega_i, S \rangle = \langle \gamma, S \rangle$ for all $i \in I_F$, we have

$$\sum_{i \in I_F} q_i \langle \gamma, S \rangle + \sum_{i \in I \setminus I_F} \langle \omega_i, S \rangle = \sum_{i \in I \setminus I_F} q_i \langle \omega_i - \gamma, S \rangle = 1,$$

showing "$\subset$". Conversely, take any $q \in \mathbb{R}^{I \setminus I_F} \subset \mathbb{R}^I$ satisfying $\sum_{i \in I \setminus I_F} q_i \langle \omega_i - \gamma, S \rangle = 1$. We can find $\tilde{q} \in \mathbb{R}^{I_F}$ such that $q + \tilde{q} \in \mathfrak{D}$. Then $\mathrm{pr}^F(q + \tilde{q}) = q$, and

$$\langle D\mathcal{A}(q + \tilde{q}), S \rangle = \sum_{i \in I \setminus I_F} q_i \langle \omega_i, S \rangle + \sum_{i \in I_F} \tilde{q}_i \underbrace{\langle \omega_i, S \rangle}_{= \langle \gamma, S \rangle} = \sum_{i \in I \setminus I_F} q_i \langle \omega_i - \gamma, S \rangle ,$$

so we also have "$\supset$". □

Finally, we prove Theorem 4.8:

proof of Theorem 4.8. Eq. (4.40) together with Corollary 4.28 gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) &= -2 \max_{(p,\zeta) \in \tilde{X}_0} \max_{q \in \mathbf{P}'_{\mathrm{relax}}} \sum_{j \in I \setminus I_F} \sqrt{q_j} \, |\langle \Phi(p,\zeta) | W | E_j \rangle| \\ &= -2 \max_{(p,\zeta) \in \tilde{X}_0} \max_{\substack{s \in \mathbb{R}^{I \setminus I_F}, s_i \geq 0 \\ \sum_{j \in I \setminus I_F} s_j^2 \langle \omega_j - \gamma, S \rangle = 1}} \sum_{j \in I \setminus I_F} s_j \, |\langle \Phi(p,\zeta) | W | E_j \rangle| . \end{aligned} \tag{4.43}$$

This constrained optimization problem is elementary, with solution given by

$$s_j = \left[\sum_{j' \in I \setminus I_F} \frac{\left| \langle \Phi(p,\zeta) | W | E_{j'} \rangle \right|^2}{\langle \omega_{j'} - \gamma, S \rangle} \right]^{-\frac{1}{2}} \frac{|\langle \Phi(p,\zeta) | W | E_j \rangle|}{\langle \omega_j - \gamma, S \rangle}.$$

Plugging this back into Eq. (4.43) finally yields

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) = -2 \max_{(p,\zeta) \in \tilde{X}_0} \left[\sum_{j \in I \setminus I_F} \frac{|\langle \Phi(p,\zeta) | W | E_j \rangle|^2}{\langle \omega_j - \gamma, S \rangle} \right]^{\frac{1}{2}}.$$

To make the independence of the chosen basis $(|E_i\rangle)_{\in I}$ manifest, note that the denominator of each summand only depends on ω_j, so we can collect the terms corresponding to each weight

$\omega \in \Omega$ so that the individual projectors $|E_j\rangle\langle E_j|$ in the numerators sum to the orthogonal projection Π_ω onto the weight space $\mathcal{H}_\omega$. The equation then becomes

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) = -2 \max_{(p,\zeta)\in\tilde{X}_0} \left[\sum_{\omega\in\Omega\setminus\Omega_F} \frac{\|\Pi_\omega W \,|\Phi(p,\zeta)\rangle\|^2}{\langle \omega - \gamma, S\rangle} \right]^{\frac{1}{2}} .$$

□

Chapter 5

Interlude: Momentum Maps

Although our development of abelian functional theories in Chapter 4 is mathematically satisfactory, interesting functional theories are often not abelian. RDMFT is such an example, where $\iota(V)$ consists of all single-particle operators, which in general do not commute. As another example, consider a system of N two-level spins (Example 2.3), where we allow all local operators to be external potentials. Certainly, the external potentials in this case also do not commute. In Chapter 4, we saw that the properties of the functionals are heavily influenced by the geometry of their domains. Before we can hope to derive any interesting properties of functional theories with non-commuting external potentials, then, it is necessary to first address the representability problem, that is, that of characterizing $\iota^*(\mathcal{P})$ and $\iota^*(\mathcal{E})$. In this scenario, we can no longer apply Theorem 4.1: the notion of "weights" does not even make sense in the first place if $\iota(V)$ is not commutative. As argued already in Chapter 4 , it would be overoptimistic to hope to solve the representability problem in full generality. Therefore, it would be desirable to isolate a class of functional theories which is not too large to involve for example the two-body N-representability problem (see Example 2.4), but still large enough to include functional theories like RDMFT or the one for spin chains with local external potentials. As it turns out, the representability problem can be treated using techniques from symplectic geometry and representation theory whenever the set of external potentials V has the structure of a Lie algebra and the potential map $\iota : V \to i\mathfrak{u}(\mathcal{H})$ is the Lie algebra representation (up to i) corresponding to a unitary representation of a compact Lie group on $\mathcal{H}$. In this case, the pure state density map $\iota^*|_{\mathcal{P}} : \mathcal{P} \to V^*$ will be a so-called *momentum map* (up to i), the image of which has been considerably studied in the symplectic geometry literature.

Before diving into the details, we will briefly summarize the history of the problem. Due to its application in quantum chemistry, quantum many-body theory and quantum information theory, the problem of characterizing $\iota^*(\mathcal{P})$ and $\iota^*(\mathcal{E})$ for a given potential map $\iota : V \to i\mathfrak{u}(\mathcal{H})$, which is the same as that of characterizing the set of all possible combinations of expectation values of a given list of operators, is not new and has been previously studied in the literature. It is a generalization of the *quantum marginal problem*, which is concerned with determining whether a given list of reduced density matrices (marginals) can originate from a pure or ensemble state in a multipartite Hilbert space (see, for example, Refs. [107, 108]). An important special case of the quantum marginal problem is that of characterizing the set of realizable 1RDMs, which is also known as the *N-representability problem*. It has

long been known [109] that for fermions, the necessary and sufficient conditions for a 1RDM to arise from an N-particle ensemble state are the usual Pauli constraints. In other words, the eigenvalues of the 1RDM lie in $[0,1]$, and sum to N. The pure state N-representability problem, on the other hand, is much more complicated. Borland and Dennis [110] found a set of complete constraints on the 1RDM spectrum for $N = 3$ fermions with a $d = 6$-dimensional single-particle Hilbert space, which is the lowest combination of (d, N) for which the N-representability problem is not trivial (if $N = 1$ or $N = d-1$, the problem is completely trivial. If $N = 2$ or $N = d - 2$, the eigenvalues of the 1RDM have to come in pairs, which is the only constraint apart from normalization). As another variant of the quantum marginal problem, Higuchi et al. [111] studied and solved the task of identifying all collections of single-qubit reduced density matrix spectra $(\lambda^{(1)}, \cdots, \lambda^{(N)})$, $\lambda^{(i)} \in \mathbb{R}^2$, which arise from some N-qubit pure state (see also Ref. [112]). Later, Klyachko [113] provided a general solution to the *pure univariant* quantum marginal problem, which includes the preceding two scenarios as special cases (see also Refs. [114–119]), based on previous work by Berenstein and Sjamaar [120], which is in turn built upon previous results on momentum maps [104, 121–125].

In this chapter, we lay out in a self-contained way the mathematical fundamentals of momentum maps, which will be applied in Chapter 6 to our discussion of a large class of nonabelian functional theories. Beyond developing the machinery and language needed for functional theories, the material that follows also serves as a basic introduction to lines of investigation in quantum information theory to understand the geometry of multipartite quantum states using momentum map techniques (see, for example, Refs. [126–128]). It is also of independent interest in other fields of physics, most notably geometric quantization [129]. The contents of this chapter are mainly drawn from Refs. [130–132], with Refs. [133–137] as general references for differential geometry and symplectic geometry and Refs. [138–140] for representation theory. We will keep following our convention that a *thing* always means a *finite-dimensional thing* implicitly, whenever applicable. In particular, "Lie algebra" without modifier will mean "finite-dimensional Lie algebra". Throughout this chapter, we will avoid using the Dirac bra-ket notation, so a vector in $\mathcal{H}$ will simply be written as ψ instead of $|\psi\rangle$.

5.1 Symplectic Geometry

Let M be a smooth manifold. A *symplectic form* on M is a nondegenerate 2-form $\omega \in \Omega^2(M)$ that is closed. The pair (M, ω) is called a *symplectic manifold*. Given a smooth function $f \in C^\infty(M)$, the *symplectic gradient* is the vector field defined by $\operatorname{sgrad} f := (\mathrm{d}f)^\sharp$, where $\sharp : \Omega^1(M) \to \mathfrak{X}(M)$ is the isomorphism induced by ω. In other words, $\operatorname{sgrad} f$ is the unique vector field satisfying

$$\iota_{\operatorname{sgrad} f}\omega = \mathrm{d}f. \tag{5.1}$$

It is common to denote $\operatorname{sgrad} f$ by X_f, but we will avoid this notation. A vector field X is *Hamiltonian* if $X = \operatorname{sgrad} f$ for some smooth function f. If X is Hamiltonian, then

$$L_X\omega = \mathrm{d}\iota_X\omega + \iota_X\mathrm{d}\omega = \mathrm{d}\mathrm{d}f = 0,$$

where $L_X(\cdot)$ and $\iota_X(\cdot)$ denote the Lie derivative along X and contraction with X respectively. A vector field $Y \in \mathfrak{X}(M)$ satisfying $L_Y\omega = 0$ is called *symplectic*. Hence, all Hamiltonian vector fields are symplectic. Conversely, all symplectic vector fields are only locally Hamiltonian, but in general there are symplectic vector fields that are not globally Hamiltonian. Let

$\mathfrak{X}_H(M)$ and $\mathfrak{X}_S(M)$ denote the sets of Hamiltonian and symplectic vector fields respectively, then the situation is as follows:

$$\mathfrak{X}_H(M) \subset \mathfrak{X}_S(M) \subset \mathfrak{X}(M)$$

Proposition 5.1. *Let $X, Y \in \mathfrak{X}_S(M)$ be symplectic vector fields, then*

$$[X, Y] = \operatorname{sgrad}(\omega(Y, X)). \tag{5.2}$$

In particular, $[X, Y]$ is Hamiltonian and therefore $[\mathfrak{X}_S(M), \mathfrak{X}_S(M)] \subset \mathfrak{X}_H(M)$.

Proof. $\iota_{[X,Y]}\omega = L_X\iota_Y\omega - \underbrace{\iota_Y L_X\omega}_{=0} = \mathrm{d}(\iota_X\iota_Y\omega) + \underbrace{\iota_X \mathrm{d}(\iota_Y\omega)}_{=0} = \mathrm{d}(\omega(Y, X))$. □

The symplectic form ω induces a map $\{\cdot,\cdot\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$ defined by $\{f, g\} = \omega(\operatorname{sgrad} f, \operatorname{sgrad} g)$, called the *Poisson bracket*, making $(C^\infty(M), \{\cdot,\cdot\})$ into a Poisson algebra, in particular a Lie algebra.

Proposition 5.2. *The map* $\operatorname{sgrad} : C^\infty(M) \to \mathfrak{X}(M)$ *is a Lie algebra antihomomorphism. That is,*

$$[\operatorname{sgrad} f, \operatorname{sgrad} g] = -\operatorname{sgrad}\{f, g\}. \tag{5.3}$$

Proof. The vector fields $\operatorname{sgrad} f, \operatorname{sgrad} g$ are Hamiltonian, so in particular symplectic. By Proposition 5.1, we have $[\operatorname{sgrad} f, \operatorname{sgrad} g] = \operatorname{sgrad}(\omega(\operatorname{sgrad} g, \operatorname{sgrad} f)) = -\operatorname{sgrad}\{f, g\}$. □

The kernel of the map sgrad is the locally constant functions, so we have an exact sequence of Lie algebras

$$0 \to H^0_{\mathrm{dR}}(M) \to C^\infty(M) \xrightarrow{\operatorname{sgrad}} \mathfrak{X}_H(M) \to 0. \tag{5.4}$$

5.1.1 Momentum Maps

Let G be a Lie group acting on a symplectic manifold (M, ω) on the left. For $g \in G$, let $l_g : M \to M$ denote the diffeomorphism $p \mapsto g \cdot p$. Let $\mathfrak{g}$ denote the Lie algebra of G. The G-action on M defines a Lie algebra antihomomorphism $\phi : \mathfrak{g} \to \mathfrak{X}(M)$, sending an element $A \in \mathfrak{g}$ of the Lie algebra to its corresponding fundamental vector field $A^* \subset \mathfrak{X}(M)$.

Definition 5.1. *The G-action on M is **symplectic** if*

$$l_g^*\omega = \omega \tag{5.5}$$

for all $g \in G$.

If the G-action on M is symplectic, then taking $g = \exp(tA)$ in Eq. (5.5), $A \in \mathfrak{g}$, and differentiating with respect to t gives $L_{A^*}\omega = 0$. That is, the fundamental vector field A^* is symplectic and therefore $\phi(\mathfrak{g}) \subset \mathfrak{X}_S(M)$.

Definition 5.2. *Let G be a connected Lie group acting on a symplectic manifold (M, ω), not necessarily preserving the symplectic form. A* ***comomentum map*** *(for the G-action on M) is a Lie algebra homomorphism $\mu : \mathfrak{g} \to C^\infty(M)$ such that the following diagram commutes:*

$$\begin{array}{ccc} C^\infty(M) & \xrightarrow{\text{sgrad}} & \mathfrak{X}(M) \\ & \nwarrow^{\mu} \quad \uparrow_{\phi} & \\ & \mathfrak{g} & \end{array} \tag{5.6}$$

That is, μ is a lift of $\phi : \mathfrak{g} \to \mathfrak{X}(M)$ to $\mathfrak{g} \to C^\infty(M)$. The corresponding ***momentum map*** *is the map $\Phi : M \to \mathfrak{g}^*$ defined by $\langle \Phi(p), A \rangle = \mu(A)(p)$.*

Remark 5.1. In Diagram 5.6, the maps sgrad and ϕ are antihomomorphisms, so it makes sense to expect μ to be a (usual) homomorphism.

Remark 5.2. There is a more general definition of momentum maps for not necessarily connected Lie groups, but we will always work with connected ones.

Remark 5.3. The image of sgrad in $\mathfrak{X}(M)$ is, by definition, the Hamiltonian vector fields $\mathfrak{X}_H(M) \subset \mathfrak{X}(M)$. It follows that if $\phi(\mathfrak{g}) \not\subset \mathfrak{X}_H(M)$, then there will be no momentum maps. In other words, if there is a momentum map, then $\phi(\mathfrak{g}) \subset \mathfrak{X}_H$, so the G-action must be at least symplectic.

5.1.2 Existence and Uniqueness

In this section, we will discuss the existence and uniqueness of momentum maps given a G-action on M. For applications in functional theories, we will simply show that a given map $M \to \mathfrak{g}^*$ is a momentum map, and we will not be concerned with its uniqueness. Thus, this section is not relevant for Chapter 6.

From now on we will always assume that G is connected. From the definition alone, it is not clear whether a (co)momentum map should exist or, if one exists, whether it is unique. For example, as pointed out in Remark 5.3, if the image of the homomorphism $\phi : \mathfrak{g} \to \mathfrak{X}(M)$ is not contained in $\mathfrak{X}_H(M)$, then there would not exist even a map of *sets* μ that makes Diagram 5.6 commute. If $\phi(\mathfrak{g}) \subset \mathfrak{X}_H(M)$, then we can always find a *linear* map from $\mathfrak{g}$ to $C^\infty(M)$ making Diagram 5.6 commute, but it is not clear whether we can always find a Lie algebra homomorphism.

The conditions for the existence and uniqueness of momentum maps will be phrased in terms of the Lie algebra cohomology $H^\bullet(\mathfrak{g})$ of $\mathfrak{g}$, which is defined as the cohomology of the differential complex

$$0 \xrightarrow{\delta} \mathbb{R} \xrightarrow{\delta} \mathfrak{g}^* \xrightarrow{\delta} \bigwedge^2 \mathfrak{g}^* \xrightarrow{\delta} \bigwedge^3 \mathfrak{g}^* \xrightarrow{\delta} \cdots,$$

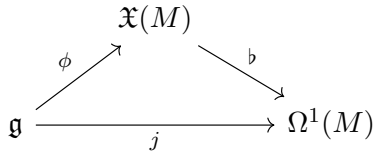
where $\delta : \bigwedge^k \mathfrak{g}^* \to \bigwedge^{k+1} \mathfrak{g}^*$ is defined as

$$\delta\omega(A_1, \cdots, A_{k+1}) = \sum_{1 \le i < j \le k+1} (-1)^{i+j} \omega([A_i, A_j], A_1, \cdots, \widehat{A_i}, \cdots, \widehat{A_j}, \cdots, A_{k+1})$$

and $\delta\omega = 0$ for $\omega \in \bigwedge^0 \mathfrak{g}^* = \mathbb{R}$. An important basic result is *Whitehead's lemma*, which asserts that $H^1(\mathfrak{g}) = H^2(\mathfrak{g}) = 0$ if $\mathfrak{g}$ is semisimple.

Proposition 5.3. *Assume that the G-action on M is symplectic. If $H^1(\mathfrak{g}) = 0$ or $H^1_{\mathrm{dR}}(M) = 0$, then $\phi(\mathfrak{g}) \subset \mathfrak{X}_H(M)$.*

Proof. Let $j : \mathfrak{g} \to \Omega^1(M)$ denote the map $A \mapsto \iota_{A^*}\omega \in \Omega^1(M)$. Then j is the composition of the inverse of $\sharp$, which we denote by $\flat : \mathfrak{X}(M) \to \Omega^1(M)$, and ϕ. Hence, showing that ϕ sends everything to Hamiltonian vector fields is equivalent to showing that j sends everything to exact 1-forms.



Take any $A \in \mathfrak{g}$. By Cartan's magic formula,

$$\mathrm{d}j(A) = \mathrm{d}(\iota_{A^*}\omega) = L_{A^*}\omega - \iota_{A^*}\mathrm{d}\omega = 0,$$

so $j(A)$ is closed. If $H^1_{\mathrm{dR}}(M) = 0$, then $j(A)$ is also exact and we are done.

The above argument shows that we can define a map

$$\begin{aligned} j' : \mathfrak{g} &\to H^1_{\mathrm{dR}}(M) \\ A &\mapsto [j(A)], \end{aligned}$$

where $[j(A)]$ denotes the cohomology class of $j(A)$. Hence, j' can be viewed as an element in $\mathfrak{g}^* \otimes H^1_{\mathrm{dR}}(M)$. For any $A, B \in \mathfrak{g}$, we then have

$$\delta j'(A, B) = -j'([A, B]) = -[j([A, B])] = [\iota_{[A^*, B^*]}\omega] = [\mathrm{d}(\omega(B^*, A^*))] = 0,$$

where we have used Proposition 5.1 in the penultimate equality. It follows that $\delta j' = 0$. Hence, if $H^1(\mathfrak{g}) = 0$, then $j' = 0$. □

Theorem 5.4. *Assume M is connected and $\phi(\mathfrak{g}) \subset \mathfrak{X}_H(M)$ (in particular, the G-action is symplectic).*

(1) If $H^2(\mathfrak{g}) = 0$, then a momentum map exists.

(2) If a momentum map exists, then the set of all momentum maps is an affine space for $H^1(\mathfrak{g})$.

Proof. M being connected implies $H^0_{\mathrm{dR}}(M) = \mathbb{R}$. Hence, the short exact sequence (5.4) becomes

$$0 \to \mathbb{R} \to C^\infty(M) \xrightarrow{\mathrm{sgrad}} \mathfrak{X}_H(M) \to 0,$$

where $\mathbb{R} \to C^\infty(M)$ is the inclusion of constant functions into $C^\infty(M)$. The statements are now consequences of the following lemma. □

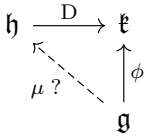
Lemma 5.5. *Let*

$$0 \to \mathbb{R} \to \mathfrak{h} \xrightarrow{\mathrm{D}} \mathfrak{k} \to 0 \tag{5.7}$$

be a central extension of Lie algebras (that is, the sequence is exact and $\mathbb{R} \subset Z(\mathfrak{h})$), with $\mathfrak{h}$ and $\mathfrak{k}$ possibly infinite-dimensional. Let $\mathfrak{g}$ be a Lie algebra and $\phi : \mathfrak{g} \to \mathfrak{k}$ a homomorphism.

(1) If $H^2(\mathfrak{g}) = 0$, then a lift $\mu : \mathfrak{g} \to \mathfrak{h}$ exists.

(2) If at least one lift exists, the space of all lifts $\mathfrak{g} \to \mathfrak{h}$ is an affine space for $H^1(\mathfrak{g})$.



Proof. Since D is surjective, we can find a *linear* map $\nu : \mathfrak{g} \to \mathfrak{h}$ such that $\mathrm{D} \circ \nu = \phi$. In general, however, the map ν will not be a Lie algebra homomorphism, so we might hope to find a "correction" linear map $\beta : \mathfrak{g} \to \mathfrak{h}$ so that $\nu + \beta$ is a Lie algebra homomorphism. If $\mathrm{D} \circ \beta = 0$, then $\nu + \beta$ will still be a lift of ϕ. By the exactness of (5.7), this is equivalent to requiring $\beta(\mathfrak{g}) \subset \mathbb{R}$, i.e., $\beta \in \mathfrak{g}^*$.

The requirement that $\nu + \beta$ be a Lie algebra homomorphism then becomes

$$[(\nu+\beta)(A), (\nu+\beta)(B)] = [\nu(A), \nu(B)] \overset{!}{=} (\nu+\beta)[A,B] \tag{5.8}$$

for all $A, B \in \mathfrak{g}$. Define $\epsilon : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{h}$ by $\epsilon(A,B) := [\nu(A), \nu(B)] - \nu([A,B])$. Then

$$\begin{aligned}
\mathrm{D}(\epsilon(A,B)) &= \mathrm{D}([\nu(A),\nu(B)]) - \mathrm{D}\circ\nu([A,B]) \\
&= [\mathrm{D}\circ\nu(A), \mathrm{D}\circ\nu(B)] - \mathrm{D}\circ\nu([A,B]) \\
&= [\phi(A),\phi(B)] - \phi([A,B]) = 0,
\end{aligned}$$

where we have used the fact that D and ϕ are homomorphisms. By the exactness of (5.7), the image of ϵ is contained in $\mathbb{R} \subset \mathfrak{h}$. Since ϵ is clearly skew-symmetric, we have $\epsilon \in \bigwedge^2 \mathfrak{g}^*$. Condition (5.8) then reads

$$-\delta\beta \overset{!}{=} \epsilon. \tag{5.9}$$

We check that ϵ is closed:

$$\begin{aligned}
-\delta\epsilon(A,B,C) &= \epsilon([A,B],C) + \epsilon([B,C],A) + \epsilon([C,A],B) \\
&= [\nu([A,B]),\nu(C)] - \nu([[A,B],C]) + (\text{cyclic permutations}) \\
&\overset{\text{Jacobi}}{=} [\nu([A,B]),\nu(C)] + (\text{cyclic permutations}) \\
&\overset{\text{Jacobi}}{=} \big[\nu([A,B]) - [\nu(A),\nu(B)],\nu(C)\big] + (\text{cyclic permutations}) \\
&= \big[-\epsilon(A,B),\nu(C)\big] + (\text{cyclic permutations}) = 0,
\end{aligned}$$

where we have used $\mathrm{im}(\epsilon) \subset \mathbb{R}$ in the last line. If $H^2(\mathfrak{g}) = 0$, then ϵ is also exact, so we can find a $\beta \in \mathfrak{g}^*$ satisfying Eq. (5.9), proving (1).

Now suppose $\mu : \mathfrak{g} \to \mathfrak{h}$ is any homomorphism such that $\mathrm{D} \circ \mu = \phi$. Take $\beta \in H^1(\mathfrak{g}) = \ker(\delta : \mathfrak{g}^* \to \bigwedge^2 \mathfrak{g}^*)$. Then $\mathrm{D} \circ \beta = 0$, so $\mathrm{D} \circ (\mu + \beta) = \phi$. Furthermore,

$$[(\mu+\beta)(A),(\mu+\beta)(B)] = (\mu+\beta)([A,B])$$

since $\beta([A,B]) = -\delta\beta(A,B) = 0$ and $\beta(A), \beta(B) \in \mathbb{R}$, showing that $\mu+\beta$ is a homomorphism.

If $\mu' : \mathfrak{g} \to \mathfrak{h}$ is another homomorphism such that $\mathrm{D} \circ \mu' = \phi$, then $\mu' - \mu \in \mathfrak{g}^*$, and

$$\begin{aligned}
(\mu'-\mu)([A,B]) &= [\mu'(A),\mu'(B)] - [\mu(A),\mu(B)] \\
&= [\underbrace{\mu'(A)-\mu(A)}_{\in\mathbb{R}},\mu'(B)] + [\mu(A),\underbrace{\mu'(B)-\mu(B)}_{\in\mathbb{R}}] \\
&= 0,
\end{aligned}$$

so $\delta(\mu' - \mu) = 0$ and therefore $\mu' - \mu \in H^1(\mathfrak{g})$. This proves (2). □

If $\mathfrak{g}$ is semisimple, then $H^1(\mathfrak{g}) = H^2(\mathfrak{g}) = 0$ by Whitehead's lemma. Consequently, we have:

Corollary 5.6. *Let G be a connected Lie group acting on a connected symplectic manifold M. If $\mathfrak{g}$ is semisimple, then there is a unique momentum map for the G-action on M.*

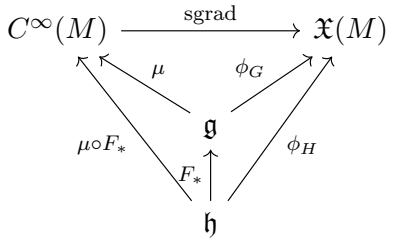
5.1.3 Properties of Momentum Maps

In this section, we collect various results on properties of momentum maps. Among them, Kirwan's theorem (Theorem 5.15) is the most important one and will be relevant for Chapter 6. Although the remaining results will not all be used directly in this thesis, we include them here because we expect them to be useful for future developments of nonabelian functional theories.

Let (M, ω) be a connected symplectic manifold with a symplectic action of a connected Lie group G. Suppose $\Phi : M \to \mathfrak{g}^*$ is a momentum map.

Proposition 5.7. *Let H be a connected Lie group and $F : H \to G$ a Lie group homomorphism, making H act on M as well. Suppose $\Phi : M \to \mathfrak{g}^*$ is a momentum map for the G-action on M. Then $F^* \circ \Phi : M \to \mathfrak{h}^*$ is a momentum map for the H-action.*

Proof. Let $\mu : \mathfrak{g} \to C^\infty(M)$ denote the comomentum map. The following diagram commutes (where $\phi_G(A) = A^*$ etc.):



(5.10)

So $\mu \circ F_*$ is a comomentum map for the H-action. It is easy to check that the momentum map corresponding to $\mu \circ F_*$ is $F^* \circ \Phi$. □

Lemma 5.8. *For all $A, B \in \mathfrak{g}$, $X \in T_pM$:*

(1) $\langle \Phi, [A, B] \rangle = \omega(A^*, B^*)$

(2) $\langle D_p\Phi(X), A \rangle = \langle (A^*)^\flat_p, X \rangle = \omega(A^*_p, X)$.

Proof.
(1) $\langle \Phi, [A, B] \rangle = \{ \langle \Phi, A \rangle, \langle \Phi, B \rangle \} = \omega\big(\operatorname{sgrad} \langle \Phi, A \rangle, \operatorname{sgrad} \langle \Phi, B \rangle \big) = \omega(A^*, B^*)$
(2) $\langle D_p\Phi(X), A \rangle = \mathrm{d}_p(\langle \Phi, A \rangle)(X) = (\operatorname{sgrad} \langle \Phi, A \rangle)^\flat_p(X) = \langle (A^*)^\flat_p, X \rangle = \omega(A^*_p, X)$. □

Corollary 5.9. *The derivative $D_p\Phi : T_pM \to \mathfrak{g}^*$ is the dual map of the map $\mathfrak{g} \to T^*_pM$, $A \mapsto (A^*_p)^\flat$.*

Now, we are ready to establish an important fact about momentum maps: they are G-equivariant. Recall that the Lie group G acts naturally on $\mathfrak{g}^*$ via the *coadjoint action* Ad^*, which is defined by $\langle \mathrm{Ad}^*_g(\alpha), A\rangle = \langle \alpha, \mathrm{Ad}_{g^{-1}}(A)\rangle$ for all $\alpha \in \mathfrak{g}^*$ and $A \in \mathfrak{g}$.

Proposition 5.10. *The map $\Phi : M \to \mathfrak{g}^*$ is G-equivariant with respect to the coadjoint action. That is,*

$$\mu(g \cdot p) = \mathrm{Ad}^*_g \Phi(p). \tag{5.11}$$

Proof. Since G is connected, we can write $g = \exp(tA_1)\exp(tA_2)\cdots\exp(tA_N)$ for $A_i \in \mathfrak{g}$. Hence, it suffices to show Eq. (5.11) for $g = \exp(tA)$.

Take any $A \in \mathfrak{g}$. Define two curves in $\mathfrak{g}^*$ by

$$\gamma_1(t) := \Phi(\exp(tA)p) \qquad \gamma_2(t) := \mathrm{Ad}^*_{\exp(tA)}\Phi(p). \tag{5.12}$$

The goal is then to show $\gamma_1(t)=\gamma_2(t)$ for all t. The trick is to show that γ_1 and γ_2 satisfy the same first-order differential equation, since we already have $\gamma_1(0) = \gamma_2(0) = \Phi(p)$. Take any $B \in \mathfrak{g}$, then

$$\begin{aligned}
&\langle \dot\gamma_1(t_0), B\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle \Phi(\exp(tA)\exp(t_0A)p), B\rangle = \langle D_{\exp(t_0A)p}\Phi(A^*_{\exp(t_0A)p}), B\rangle \\
&\overset{\text{Lem. 5.8}}{=} \omega\left(B^*_{\exp(t_0A)p}, A^*_{\exp(t_0A)p}\right) \overset{\text{Lem. 5.8}}{=} \langle \Phi(\exp(t_0A)p), [B, A]\rangle \\
&= \langle -\gamma_1(t_0)\circ \mathrm{ad}_A, B\rangle
\end{aligned}$$

$$\begin{aligned}
&\langle \dot\gamma_2(t_0), B\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle \mathrm{Ad}^*_{\exp((t_0+t)A)}\,\Phi(p), B\rangle \\
&= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle \mathrm{Ad}^*_{\exp(t_0A)}\,\Phi(p), \mathrm{Ad}_{\exp(-tA)}\,B\rangle = \langle \mathrm{Ad}^*_{\exp(t_0A)}\,\Phi(p), [-A, B]\rangle \\
&= \langle -\gamma_2(t_0)\circ \mathrm{ad}_A, B\rangle\,.
\end{aligned}$$

So both curves satisfy the differential equation $\dot\gamma = -\gamma \circ \mathrm{ad}_A$. □

Recall that if $E \subset \mathfrak{g}$ is any subset, the *annihilator* of E in $\mathfrak{g}^*$ is the set $E^\circ := \{\eta \in \mathfrak{g}^* \mid \forall A \in E : \langle \eta, A\rangle = 0\}$.

Proposition 5.11. *$D_p\Phi(T_pM) = \mathfrak{g}_p^\circ$, where $\mathfrak{g}_p$ is the stabilizer subalgebra of $p \in M$.*

Proof. Take any vector $X \in T_pM$ and $A \in \mathfrak{g}_p$. Then

$$\langle D_p\Phi(X), A\rangle = \omega(A^*_p, X) = 0,$$

since $A^*_p = 0$. This shows $D_p\Phi(T_pM) \subset \mathfrak{g}_p^\circ$. By Corollary 5.9, we have

$$\dim D_p\Phi(T_pM) = \dim \mathrm{im}(A \mapsto A^*_p) = \dim G - \dim \mathfrak{g}_p = \dim \mathfrak{g}_p^\circ,$$

so $D_p\Phi(T_pM) = \mathfrak{g}^\circ$. □

Recall that a smooth map $f : M \to N$ is said to meet a submanifold $Z \subset N$ *transversally* if $D_pf(T_pM) + T_{f(p)}Z = T_{f(p)}N$ for all $p \in f^{-1}(Z)$. A generalization of the regular value theorem asserts that in this case $f^{-1}(Z) \subset M$ is a submanifold.

Let $\mathfrak{t}^*_{>0} \subset \mathfrak{t}^*_+$ denote the open Weyl chamber. The following will be useful for the "Selection Rule" (Theorem 5.19) discussed in the next section.

Proposition 5.12. *The momentum map $\Phi : M \to \mathfrak{g}^*$ meets the open Weyl chamber $\mathfrak{t}^*_{>0}$ transversally.*

Proof. Take $p \in M$ such that $\Phi(p) \in \mathfrak{t}^*_{>0}$. By definition, we need to show

$$D_p\Phi(T_pM) + \mathfrak{t}^* = \mathfrak{g}^*. \tag{5.13}$$

By Proposition 5.11, the first summand is $\mathfrak{g}_p^\circ$, the annihilator of the stabilizer subalgebra of the G-action at $p \in M$.

$\mathfrak{g}_p \subset \mathfrak{g}_{\Phi(p)}$ because Φ is G-equivariant. But $\mathfrak{g}_{\Phi(p)} = \mathfrak{t}$ because $\Phi(p) \in \mathfrak{t}^*_{>0}$. Hence $\mathfrak{g}_p \subset \mathfrak{t}$, or $\mathfrak{g}_p^\circ \supset \mathfrak{t}^\circ$. It follows that

$$D_p\Phi(T_pM) + \mathfrak{t}^* = \mathfrak{g}_p^\circ + \mathfrak{t}^* \supset \mathfrak{t}^\circ + \mathfrak{t}^* = \mathfrak{g}^*.$$

□

Corollary 5.13. *$\Phi^{-1}(\mathfrak{t}^*_{>0})$ is a submanifold of M.*

For our purpose in Chapter 6, the most relevant properties of momentum maps are those of the image $\Phi(M)$. For the following two theorems, we assume that M is compact and connected.

Theorem 5.14 (Atiyah-Guillemin-Sternberg [103, 104])**.** *Suppose a torus $T = (S^1)^n$ acts on M with momentum map $\Phi : M \to \mathfrak{t}^* = \mathbb{R}^n$. Then $\Phi(M)$ is the convex hull of the images of all T-fixed points in M.*

Remark 5.4. When $M = \mathbb{P}(\mathcal{H})$ is the projective Hilbert space with the appropriate symplectic structure (see Section 5.2), we already know this from Theorem 4.1.

If G is not necessarily abelian, there is a weaker convexity result due to Kirwan. In this case, choose any Cartan subalgebra $\mathfrak{t} \subset \mathfrak{g}$ and a Weyl chamber $\mathfrak{t}^*_+ \subset \mathfrak{t}^*$. After fixing a G-invariant inner product on $\mathfrak{g}$, we may think of $\mathfrak{t}^*$ as a subspace of $\mathfrak{g}^*$. Because Φ is G-equivariant, it is enough to know how $\Phi(M)$ intersects with $\mathfrak{t}^*_+$.

Theorem 5.15 (Kirwan [124])**.** *Let G be a compact connected Lie group acting on M with momentum map $\Phi : M \to \mathfrak{g}^*$. The image $\Phi(M)$ intersects $\mathfrak{t}^*_+$ in a convex polytope.*

5.2 The Projective Hilbert Space

Let $\mathcal{H}$ be any complex vector space. Recall that the projective space $\mathbb{P}(\mathcal{H})$ is the complex manifold whose points are equivalence classes $[\psi]$ of nonzero vectors in $\mathcal{H}$, where $\psi \sim \psi'$ if $\psi = \lambda\psi'$ for some $\lambda \in \mathbb{C} \setminus \{0\}$. Any choice of basis for $\mathcal{H}$ induces a biholomorphism $\mathbb{P}(\mathcal{H}) \to \mathbb{CP}^{\dim \mathcal{H}-1}$. If $\mathcal{H}$ is endowed with a Hermitian inner product $\langle \cdot, \cdot \rangle$, it is convenient to embed $\mathbb{P}(\mathcal{H})$ into $i\mathfrak{u}(\mathcal{H})$, the real vector space of Hermitian endomorphisms, via the map

$$\begin{aligned} \tilde{\Phi} : \mathbb{P}(\mathcal{H}) &\to i\mathfrak{u}(\mathcal{H}) \\ [\psi] &\mapsto \frac{\psi \langle \psi, \cdot \rangle}{\langle \psi, \psi \rangle}. \end{aligned} \tag{5.14}$$

It is easy to check that $\tilde{\Phi}$ is indeed an embedding. Under $D_{[\psi]}\tilde{\Phi} : T_{[\psi]}\mathbb{P}(\mathcal{H}) \to T_{\tilde{\Phi}[\psi]} i\mathfrak{u}(\mathcal{H}) = i\mathfrak{u}(\mathcal{H})$, a tangent vector in $T_{[\psi]}\mathbb{P}(\mathcal{H})$ is mapped to a Hermitian endomorphism on $\mathcal{H}$. For notational simplicity, we will denote by $\mathbf{p}_\psi$ the orthogonal projection onto the subspace spanned by ψ for $\psi \neq 0$. That is, $\tilde{\Phi}[\psi] = \mathbf{p}_\psi$. The map $\tilde{\Phi}$ then gives, up to the identification of $i\mathfrak{u}(\mathcal{H})$ with its dual, an identification between $\mathbb{P}(\mathcal{H})$ and the set of pure states $\mathcal{P} \subset i\mathfrak{u}(\mathcal{H})^*$.

5.2.1 Unitary Group Action

There is a natural action of $\mathrm{U}(\mathcal{H})$ on $\mathbb{P}(\mathcal{H})$ given by $g \cdot [\psi] = [g\psi]$. The following proposition gives information about the fundamental vector fields A^* for $A \in \mathfrak{u}(V)$. (Hereafter, we will adopt the convention that we always pick representatives $\psi \in [\psi]$ which are normalized, i.e., $\|\psi\| = 1$.)

Proposition 5.16. *Let $A \in \mathfrak{u}(\mathcal{H})$. Then*

$$D_{[\psi]}\tilde{\Phi}(A^*_{[\psi]}) = A\psi \langle \psi, \cdot \rangle - \psi \langle \psi, A(\cdot) \rangle = [A, \mathbf{p}_\psi], \tag{5.15}$$

where $A^* \in \mathfrak{X}(\mathbb{P}(\mathcal{H}))$ is the fundamental vector field of A for the $\mathrm{U}(\mathcal{H})$-action on $\mathbb{P}(\mathcal{H})$.

Proof.

$$\begin{aligned} D_{[\psi]}\tilde{\Phi}(A^*_{[\psi]}) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \tilde{\Phi}\Big(\exp(tA)[\psi]\Big) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} (\tilde{\Phi} \circ \pi)(\exp(tA)\psi) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \frac{\exp(tA)\psi \langle \exp(tA)\psi, \cdot \rangle}{\langle \exp(tA)\psi, \exp(tA)\psi \rangle} = A\psi \langle \psi, \cdot \rangle + \psi \langle A\psi, \cdot \rangle \\ &= A\psi \langle \psi, \cdot \rangle - \psi \langle \psi, A(\cdot) \rangle , \end{aligned}$$

where $\pi : \mathcal{H} \setminus \{0\} \to \mathbb{P}(\mathcal{H})$ is the canonical projection and we have used the skew-Hermiticity of A in the last line. □

Remark 5.5. In the physics notation, on writes $|\psi\rangle\langle\psi|$ instead of $\psi \langle \psi, \cdot \rangle$ (which is equal to $\mathbf{p}_\psi$ if $\|\psi\| = 1$).

Remark 5.6. At any point $[\psi] \in \mathbb{P}(\mathcal{H})$ in the projective space, the map $\mathfrak{u}(\mathcal{H}) \to T_{[\psi]}\mathbb{P}(\mathcal{H})$, $A \mapsto A^*_{[\psi]}$ gives us a very convenient way of thinking about the tangent space at $[\psi]$. To be specific, since the $\mathrm{U}(\mathcal{H})$-action on $\mathbb{P}(\mathcal{H})$ is transitive, the map $A \mapsto A^*_{[\psi]}$ is surjective onto $T_{[\psi]}\mathbb{P}(\mathcal{H})$. In other words, every tangent vector at $[\psi]$ is the value of some fundamental vector field at $[\psi]$. However, this is not a "perfect" parametrization of the tangent space in the sense that the map $A \mapsto A^*_{[\psi]}$ is not injective.

5.2.2 The Momentum Map

The projective space $\mathbb{P}(\mathcal{H})$ has the structure of a Kähler manifold. In Appendix A, we construct explicitly the Fubini-Study metric h^{FS} and the Fubini-Study 2-form ω^{FS} on $\mathbb{P}(\mathcal{H})$ and show that they are compatible with the almost complex structure J. Of course, what is relevant for us is the symplectic structure, for which we have an explicit formula suitable for computation (Lemma A.5):

$$\omega^{\mathrm{FS}}(A^*_{[\psi]}, B^*_{[\psi]}) = i \langle \psi, [A, B]\psi \rangle \qquad (A, B \in \mathfrak{u}(\mathcal{H})). \tag{5.16}$$

Theorem 5.17. *The map*

$$\begin{aligned} \Phi : \mathbb{P}(\mathcal{H}) &\to \mathfrak{u}(\mathcal{H})^* \\ [\psi] &\mapsto i \operatorname{Tr}(\tilde{\Phi}[\psi](\cdot)) = i \operatorname{Tr}(\mathbf{p}_\psi(\cdot)) = i \langle \psi, (\cdot)\psi \rangle \end{aligned} \tag{5.17}$$

is a momentum map for the $\mathrm{U}(V)$*-action on* $(\mathbb{P}(V), \omega^{\mathrm{FS}})$.

Proof. Define $\mu : \mathfrak{g} \to C^\infty(\mathbb{P}(\mathcal{H}))$ by $\mu(A)[\psi] = \langle \Phi[\psi], A \rangle$. The definition of momentum maps (Definition 5.2) requires us to check two things:

(1) $\operatorname{sgrad} \circ \mu = \phi$

(2) μ is a Lie algebra homomorphism, i.e., $\mu([A, B]) = \{\mu(A), \mu(B)\}$

where $\phi : \mathfrak{g} \to \mathfrak{X}(\mathbb{P}(\mathcal{H}))$ is the map $A \mapsto A^*$.

The first condition is equivalent to $\mathrm{d}\mu(A) = \iota_{A^*}\omega^{\mathrm{FS}}$ for all $A \in \mathfrak{g}$. Take any $B \in \mathfrak{g}$, then

$$\begin{aligned} \langle \mathrm{d}\mu(A), B^*_{[\psi]} \rangle &= \left\langle \mathrm{d}([\psi] \mapsto i \langle \psi, A\psi \rangle), B^*_{[\psi]} \right\rangle \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} i \langle \exp(tB)\psi, A \exp(tB)\psi \rangle \\ &= i \langle \psi, [A, B]\psi \rangle = \omega^{\mathrm{FS}}(A^*_{[\psi]}, B^*_{[\psi]}) \\ &= \left(\iota_{A^*}\omega^{\mathrm{FS}}\right)(B^*_{[\psi]}), \end{aligned}$$

where we have used Eq. (5.16) for the penultimate equality. It follows that $\mathrm{d}\mu(A) = \iota_{A^*}\omega^{\mathrm{FS}}$.

The second condition also holds:

$$\begin{aligned} \{\mu(A), \mu(B)\} &= \omega^{\mathrm{FS}}(\operatorname{sgrad} \mu(A), \operatorname{sgrad} \mu(B)) \\ &= \omega^{\mathrm{FS}}(A^*, B^*) = ([\psi] \mapsto i \langle \psi, [A, B]\psi \rangle) \\ &= \mu([A, B]), \end{aligned}$$

where we have used Eq. (5.16) again. □

Corollary 5.18. *Let* G *be a connected Lie group and* $\tau : G \to \mathrm{U}(V)$ *a unitary representation, so that* G *acts on* $\mathbb{P}(\mathcal{H})$ *by* $(g, [\psi]) \mapsto [\tau(g)\psi]$*. Let* $\tau_* : \mathfrak{g} \to \mathfrak{u}(\mathcal{H})$ *denote the induced Lie algebra homomorphism. Then the map*

$$\begin{aligned} \Phi_G : \mathbb{P}(\mathcal{H}) &\to \mathfrak{g}^* \\ [\psi] &\mapsto i \operatorname{Tr}(\mathbf{p}_\psi \tau_*(\cdot)) = (A \mapsto i \langle \psi, \tau_*(A)\psi \rangle) \end{aligned} \tag{5.18}$$

is a momentum map for the G*-action on* $\mathbb{P}(\mathcal{H})$.

Proof. Clearly $\Phi_G = \tau^* \circ \Phi$. The statement follows from applying Proposition 5.7 to τ. □

Let G be a compact connected Lie group and $\tau : G \to \mathrm{U}(\mathcal{H})$ a unitary representation with momentum map $\Phi_G : \mathbb{P}(\mathcal{H}) \to \mathfrak{g}^*$, $[\psi] \mapsto (A \mapsto i \langle \psi, \tau_*(A)\psi \rangle)$. Pick a Cartan subalgebra $\mathfrak{t} \subset \mathfrak{g}$ and a G-invariant inner product on $\mathfrak{g}$, together with a Weyl chamber $\mathfrak{t}^*_+$.

Kirwan's theorem (Theorem 5.15) states that $\Phi_G(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}^*_+$ is a convex polytope, so it is characterized by a finite set of inequalities $\langle \cdot, S \rangle \geq c$, where $S \in \mathfrak{t}$ and $c \in \mathbb{R}$. In the abelian case, any $[\psi] \in \mathbb{P}(\mathcal{H})$ mapping to a facet must be spanned by the weight vectors whose weights are on the facet. This turns out to be true even when $\mathfrak{g}$ is not abelian, a result called the "Selection Rule" [141] which we now state and prove.

Theorem 5.19 (Selection Rule (Lemma 2.13 of Ref. [105])). *Let $\langle \cdot, S\rangle \geq c$ be an inequality corresponding to a facet of the convex polytope $\Phi_G(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}_+^*$. Suppose $\psi \in \mathcal{H}$ ($\|\psi\| = 1$) is such that $\Phi_G[\psi] \in \mathfrak{t}_{>0}^*$ and $\langle \Phi_G[\psi], S\rangle = c$ (i.e., $\Phi_G[\psi]$ is lying on the facet). Then $i\tau_*(S)\psi = c\psi$ and $\mathrm{d}_{[\psi]}\langle \Phi_G, S\rangle = 0$.*

Proof. Take any $Y \in D_{[\psi]}\Phi_G(T_{[\psi]}\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}^*$. Since $\Phi_G^{-1}(\mathfrak{t}_{>0}^*) \subset \mathbb{P}(\mathcal{H})$ is a submanifold (Corollary 5.13), we can find a smooth curve $\gamma : (-\epsilon, \epsilon) \to \Phi_G^{-1}(\mathfrak{t}_{>0}^*)$ such that $\gamma(0) = [\psi]$ and $D_{[\psi]}\Phi_G(\dot\gamma(0)) = Y$. Then

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle \Phi_G(\gamma(t)), S\rangle = \langle Y, S\rangle .$$

But $\langle \Phi_G(\gamma(t)), S\rangle \geq c$ is smooth and satisfies $\langle \Phi_G(\gamma(0)), S\rangle = \langle \Phi_G[\psi], S\rangle = c$, so the derivative at $t = 0$ must be zero, implying $\langle Y, S\rangle = 0$.

It follows that $\mathfrak{g}_{[\psi]}^\circ \cap \mathfrak{t}^* \subset S^\circ$, so $\mathbb{R}S \subset \mathfrak{g}_{[\psi]} + \mathfrak{t}^\perp$. We must have $S \in \mathfrak{g}_{[\psi]}$ because $S \notin \mathfrak{t}^\perp$, implying

$$\exp(t\tau_*(S))[\psi] = [\psi] \tag{5.19}$$

for all t. So $\exp(t\tau_*(S))\psi = k(t)\psi$ for some smooth function $k(t)$. Differentiating with respect to t at $t = 0$ yields $\tau_*(S)\psi = \dot k(0)\psi$. But, by assumption,

$$c = \langle \Phi_G[\psi], S\rangle = i\,\langle \psi, \tau_*(S)\psi\rangle = i\dot k(0),$$

so $\dot k(0) = -ic$. This proves the first part of the theorem.

To show the second part, note that by definition $(\mathrm{sgrad}\,\langle \Phi_G, S\rangle)_{[\psi]} = S^*_{[\psi]}$, but Eq. (5.19) implies $S^*_{[\psi]} = 0$, so $(\mathrm{sgrad}\,\langle \Phi_G, S\rangle)_{[\psi]} = (\mathrm{d}_{[\psi]}\langle \Phi_G, S\rangle)^\sharp = 0$. Since $\sharp$ is an isomorphism, we conclude $\mathrm{d}_{[\psi]}\langle \Phi_G, S\rangle = 0$. □

Finally, we will quote a result from Ref. [105] for determining the facets of the convex polytope $\Phi_G(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}_+^*$. We need some preparation for the statement of the theorem. Let G be a maximal compact subgroup of a complex reductive algebraic group G', so that $\mathfrak{g}' = \mathfrak{g}_{\mathbb{C}}$, with a representation $\tau' : G' \to \mathrm{GL}(\mathcal{H})$ such that $\tau'(G) \subset \mathrm{U}(\mathcal{H})$. Let

$$\mathfrak{g}' = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}'_\alpha$$

be the root space decomposition of $\mathfrak{g}'$, where $R \subset i\mathfrak{t}^*$ is the set of roots. Pick a base for R, which fixes the positive Weyl chamber $\mathfrak{t}_+^*$ and the decomposition of R into the disjoint union of positive and negative roots, which we denote by R_+ and R_- respectively. For $S \in \mathfrak{t}, c \in \mathbb{R}$, define

$$\begin{aligned}
\mathcal{H}(S < c) &:= \bigoplus_{\omega \in \Omega : i\langle \omega, S\rangle < c} \mathcal{H}_\omega \\
\mathcal{H}(S = c) &:= \bigoplus_{\omega \in \Omega : i\langle \omega, S\rangle = c} \mathcal{H}_\omega \\
\mathfrak{n}_-(S < 0) &:= \bigoplus_{\alpha \in R_- : i\langle \alpha, S\rangle < 0} \mathfrak{g}'_\alpha,
\end{aligned}$$

where $\Omega \subset i\mathfrak{t}^*$ is the set of weights of the representation τ' relative to $\mathfrak{t}_{\mathbb{C}}$. Note that $\tau'(\mathfrak{n}_-(S < 0))\mathcal{H}(S = c) \subset \mathcal{H}(S < c)$.

Theorem 5.20 (Theorem 3.14 of Ref. [105]). *If the convex polytope $\Phi_G(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}_+^*$ has full dimension in $\mathfrak{t}^*$, then it is characterized by (in addition to the inequalities defining the positive Weyl chamber)*

$$\langle \cdot, S \rangle \geq c,$$

for all $S \in \mathfrak{t}$, $c \in \mathbb{R}$ such that

(1) the hyperplane $\langle \cdot, S \rangle = c$ is the affine hull of a subset of $i\Omega$, and

(2) there exists a $\psi \in \mathcal{H}(S = c)$ for which the map

$$\begin{aligned} \mathfrak{n}_-(S < 0) &\to \mathcal{H}(S < c) \\ A &\mapsto \tau'_*(A)\psi \end{aligned} \tag{5.20}$$

is an isomorphism.

Example 5.1. Take $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^3$, $G = \mathrm{SU}(2) \times \mathrm{SU}(2)$, $G' = \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C})$, with G' acting on $\mathcal{H}$ by the tensor product of the fundamental representation $\tau^{(2)}$ and the adjoint representation $\tau^{(3)}$ of $\mathrm{SL}(2,\mathbb{C})$. The momentum map is

$$\begin{aligned} \Phi : \mathbb{P}(\mathbb{C}^2 \otimes \mathbb{C}^3) &\to \mathfrak{su}(2)^* \oplus \mathfrak{su}(2)^* \\ [\psi \otimes \chi] &\mapsto (A \mapsto i \langle \psi, \tau_*^{(2)}(A)\psi \rangle , B \mapsto i \langle \chi, \tau_*^{(3)}(A)\chi \rangle). \end{aligned}$$

For the Cartan subalgebra, we choose

$$\mathfrak{t}_{\mathbb{C}} = \mathbb{C}Z_1 \oplus \mathbb{C}Z_2,$$

where $Z_1 = (Z, 0)$ and $Z_2 = (0, Z)$ with $Z = \mathrm{diag}(1, -1)$. Relative to $\mathfrak{t}_{\mathbb{C}}$, the four roots are

$$\begin{aligned} \alpha &= (Z \mapsto 2, 0) \\ -\alpha &= (Z \mapsto -2, 0) \\ \beta &= (0, Z \mapsto 2) \\ -\beta &= (0, Z \mapsto 2). \end{aligned}$$

Similarly, the weights of the representation $\tau^{(2)} \otimes \tau^{(3)}$ are

$$\begin{aligned} \omega_{-1,2} &= (Z \mapsto -1, Z \mapsto 2) & \omega_{1,2} &= (Z \mapsto 1, Z \mapsto 2) \\ \omega_{-1,0} &= (Z \mapsto -1, Z \mapsto 0) & \omega_{1,0} &= (Z \mapsto 1, Z \mapsto 0) \\ \omega_{-1,-2} &= (Z \mapsto -1, Z \mapsto -2) & \omega_{1,-2} &= (Z \mapsto 1, Z \mapsto -2), \end{aligned}$$

each with multiplicity one. Note that all the roots and weights belong to $i\mathfrak{t}^*$. We will choose α and β to be the positive simple roots, yielding the positive Weyl chamber

$$\mathfrak{t}_+^* = \{\lambda i\alpha + \mu i\beta \mid \lambda, \mu \geq 0\} \subset \mathfrak{t}^*.$$

We will now use Theorem 5.20 to determine the polytope $\Lambda := \Phi_G(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}_+^*$.

First we show that Λ satisfies the hypothesis that $\dim \Lambda = \dim \mathfrak{t}^* = 2$. To see this, consider the images of the vectors $\phi_{1,2}, \phi_{1,0}, \psi := \frac{1}{\sqrt{2}}(\phi_{1,2} + \phi_{-1,-2})$, where we label the (unit-norm) weight vectors by the eigenvalues of Z_1 and Z_2. Clearly, we have

$$\Phi[\phi_{1,2}] = i\omega_{1,2} \quad \Phi[\phi_{1,0}] = i\omega_{1,0} \qquad \Phi[\psi] = 0.$$

The points $i\omega_{1,2}, i\omega_{1,0}, 0$ all belong to $\mathfrak{t}_+^*$ and are not collinear, so $\dim \Lambda$ must be two and the hypothesis of Theorem 5.20 is satisfied.

To obtain the inequalities, we only have to check lines (codimension one affine spaces) $\langle \cdot, S\rangle = c$ passing through (at least) two weights. Looking at Fig. 5.1, we see that there are twelve possibilities (six lines in total):

line #	S	c
1	$\pm iZ_2$	∓ 2
2	$\pm(iZ_1 - iZ_2)$	± 1
3	$\pm(2iZ_1 - iZ_2)$	0
4	$\pm iZ_1$	∓ 1
5	$\pm(iZ_1 + iZ_2)$	∓ 1
6	$\pm(iZ_1 - iZ_2)$	∓ 1

For each candidate (S, c), we have to show whether there exists a $\psi \in \mathcal{H}(S = c)$ such that the map (5.20) is an isomorphism. This already rules out most candidates due to dimension reasons: only those (S, c) for which $\dim \mathfrak{n}_-(S < 0) = \dim \mathcal{H}(S < c)$ have to be considered.

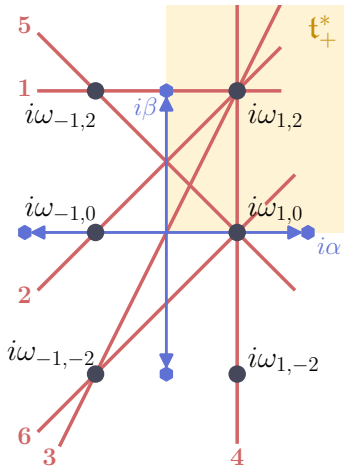


Figure 5.1 Candidate inequalities.

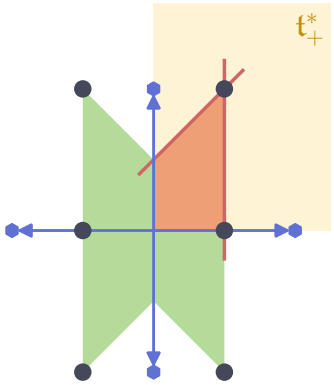


Figure 5.2 $\Phi(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}^*$ (green and orange) and $\Lambda = \Phi(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}_+^*$ (orange).

In Table 5.1, we list all twelve candidate inequalities, together with the corresponding subspaces $\mathfrak{n}_-(S < 0)$, $\mathcal{H}(S = c)$, and $\mathcal{H}(S < c)$, where a filled hexagon or circle indicates that the corresponding root space or weight space is included in the respective subspaces. Each inequality is labeled by a pair (m, s), where m is the index of the line on which the inequality is saturated, and s is the choice of sign (choice of one of the two half spaces bounded by the line). For example, inequality $(1, +)$ is $\langle \cdot, iZ\rangle \geq -2$. For inequalities $(1,-), (2,+), (3,+), (3,-), (4,-), (5,+), (5,-), (6,-)$, the dimensions of $\mathfrak{n}_-(S < 0)$ and $\mathcal{H}(S < c)$ differ, so they cannot be valid inequalities.

This leaves us with inequalities $(1,+), (2,-), (4,+), (6,+)$. For $(1,+)$ and $(4,+)$, the isomorphism criterion (5.20) is trivially satisfied since $\dim \mathfrak{n}_-(S < 0) = \dim \mathcal{H}(S < c) = 0$, so both are valid inequalities. Moreover, these already imply $(6,+)$, so we only have to check

criterion (5.20) for $(2,-)$, in which case

$$\begin{aligned}\mathfrak{n}_-(S<0) &= \mathfrak{g}'_{-\alpha} = \mathbb{C}(X_1 - iY_1)\\ \mathcal{H}(S=c) &= \mathbb{C}\phi_{-1,0} + \mathbb{C}\phi_{1,2}\\ \mathcal{H}(S<c) &= \mathbb{C}\phi_{-1,2}.\end{aligned}$$

Since $\tau'_*(X_1 - iY_1)\phi_{1,2} = \lambda\phi_{-1,2}$, $\lambda \neq 0$, the map $\mathfrak{n}_-(S<0) \to \mathcal{H}(S<0), A \mapsto \tau'_*(A)\phi_{1,2}$ is bijective. It follows that $\langle \cdot, -iZ_1 + iZ_2\rangle \geq -1$ is a bounding inequality of the polytope Λ.

We conclude that the polytope Λ is bounded by the inequalities $(1,+), (4,+)$, and $(2,-)$ (see Fig. 5.2).

Example 5.2. Take $\mathcal{H} = \mathfrak{sl}(3,\mathbb{C})$, $G = \mathrm{SU}(3)$, $G' = \mathrm{SL}(3,\mathbb{C})$, with G' acting on $\mathcal{H}$ by the adjoint action $\tau' = \mathrm{Ad}$. We choose the subalgebra of diagonal matrices to be the Cartan subalgebra $\mathfrak{t}_\mathbb{C}$, which has a basis given by $Z := \mathrm{diag}(1,-1,0)$, $Z' := \mathrm{diag}(0,1,-1)$. The six roots are

$$\begin{array}{lll}\alpha_1 = (Z \mapsto 2; Z' \mapsto -1) & \alpha_2 = (Z \mapsto -1; Z' \mapsto 2) & \alpha_3 = (Z \mapsto 1; Z' \mapsto 1)\\ -\alpha_1 = (Z \mapsto -2; Z' \mapsto 1) & -\alpha_2 = (Z \mapsto 1; Z' \mapsto -2) & -\alpha_3 = (Z \mapsto -1; Z' \mapsto -1),\end{array}$$

which, together with $0 \in i\mathfrak{t}^*$ with multiplicity two, are the weights of τ'. After carrying out the same procedure as in Example 5.1, one obtains the polytope $\Phi(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}^*_+$, which is shown in Fig. 5.3. In this case, there are two inequalities, which are given by

$$\begin{aligned}\langle \cdot, iZ\rangle &\geq -1\\ \langle \cdot, iZ'\rangle &\geq -1.\end{aligned}$$

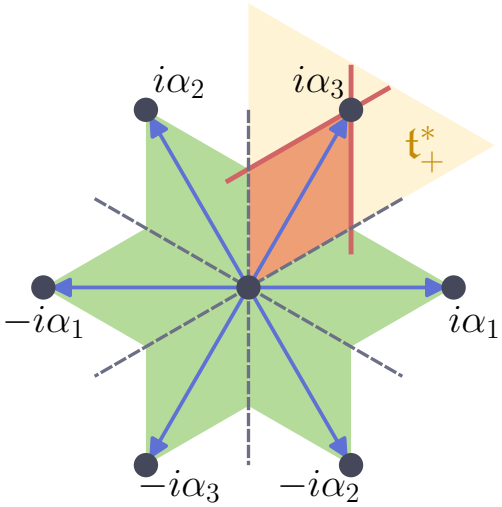


Figure 5.3 The image of the momentum map for the adjoint action of $\mathrm{SU}(3)$ on $\mathfrak{sl}(3,\mathbb{C})$ intersected with $\mathfrak{t}^*$ (green and orange) and with $\mathfrak{t}^*_+$ (orange). The gray dashed lines are the reflection hyperplanes of the Weyl group. The two red lines are the bounding hyperplanes of $\Phi(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}^*_+$.

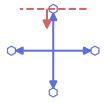
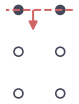
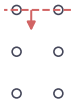
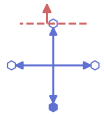
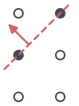
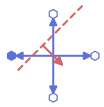
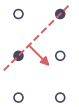
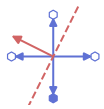
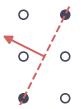
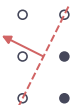
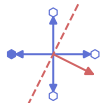
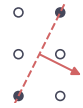
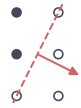
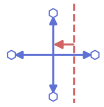
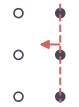
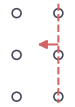
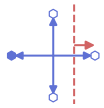
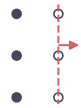
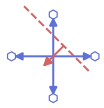
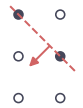
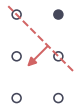
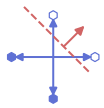
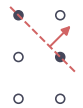
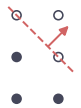
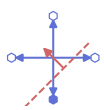
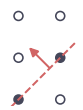
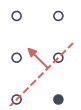
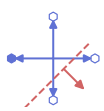
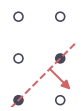
candidate inequality (line #, sign)	S	c	$\mathfrak{n}_-(S<0)$	$\mathcal{H}(S=c)$	$\mathcal{H}(S<c)$
$1,+$	iZ_2	-2			
$1,-$	$-iZ_2$	2			
$2,+$	iZ_1-iZ_2	1			
$2,-$	$-iZ_1+iZ_2$	-1			
$3,+$	$2iZ_1-iZ_2$	0			
$3,-$	$-2iZ_1+iZ_2$	0			
$4,+$	iZ_1	-1			
$4,-$	$-iZ_1$	1			
$5,+$	iZ_1+iZ_2	-1			
$5,-$	$-iZ_1-iZ_2$	$+1$			
$6,+$	iZ_1-iZ_2	-1			
$6,-$	$-iZ_1+iZ_2$	1			

Table 5.1 Candidate inequalities for the action of $\mathrm{SU}(2)\times\mathrm{SU}(2)$ on $\mathbb{C}^2\otimes\mathbb{C}^3$ (Example 5.1).

Chapter 6

Boundary Forces in Nonabelian Functional Theories

6.1 Definitions

As remarked in the beginning of Chapter 5, investigating all nonabelian functional theories in the most general setting might be overly ambitious. Therefore, we will limit ourselves to the case where the potential map comes from a unitary representation of a compact Lie group on the Hilbert space. This motivates the following definition:

Definition 6.1. *A **momentum map functional theory** is a tuple $(G, \mathcal{H}, \tau, W)$, where G is a compact connected Lie group, $\mathcal{H}$ is a Hilbert space, $\tau : G \to \mathrm{U}(\mathcal{H})$ is a representation, and $W \in i\mathfrak{u}(\mathcal{H})$ is any Hermitian operator.*

Let $\mathfrak{g}$ denote the Lie algebra of G and $\tau_* : \mathfrak{g} \to \mathfrak{u}(\mathcal{H})$ the Lie algebra representation. Then the tuple $(i\mathfrak{g}, \mathcal{H}, \iota, W)$ is a generalized functional theory in the sense of Definition 2.1, where the potential map $\iota : i\mathfrak{g} \to i\mathfrak{u}(\mathcal{H})$ is defined by $\iota(iA) = i\tau_*(A)$. Instead of writing ι, we will often simply use τ_* to denote the potential map whenever there is no ambiguity. To put it another way, the potential map is the restriction of the complexified representation $(\tau_*)_{\mathbb{C}} : \mathfrak{g}_{\mathbb{C}} \to \mathfrak{gl}(\mathcal{H})$ to $i\mathfrak{g} \subset \mathfrak{g}_{\mathbb{C}}$.

The density map is then the restriction of the dual map $\tau^* : i\mathfrak{u}(\mathcal{H})^* \to i\mathfrak{g}^*$ to the set of ensemble states $\mathcal{E} \subset i\mathfrak{u}(\mathcal{H})^*$ (recall our convention of viewing an ensemble state as a linear functional on $i\mathfrak{u}(\mathcal{H})$). As before, we will not indicate the restriction to $\mathcal{E}$ explicitly, so the density map is simply denoted as τ^*.

Example 6.1. RDMFT is a momentum map functional theory. Indeed, let $\mathcal{H}_1$ be any Hilbert space (interpreted as the single-particle Hilbert space), and let $\mathcal{H} = \bigwedge^N \mathcal{H}_1$ or $\mathrm{Sym}^N \mathcal{H}_1$. There is a natural homomorphism $\tau : \mathrm{U}(\mathcal{H}_1) \to \mathrm{U}(\mathcal{H})$, and the corresponding potential map is

$$\begin{aligned} \tau_* : i\mathfrak{u}(\mathcal{H}_1) &\to i\mathfrak{u}(\mathcal{H}) \\ h &\mapsto h \otimes \mathbb{1}^{\otimes(N-1)} + \cdots + \mathbb{1}^{\otimes(N-1)} \otimes h. \end{aligned}$$

The density map $\tau^* : i\mathfrak{u}(\mathcal{H})^* \to i\mathfrak{u}(\mathcal{H}_1)^*$ is just N times the partial trace over $N-1$ particles.

Example 6.2. Let $\mathcal{H} = (\mathbb{C}^d)^{\otimes N}$ be the Hilbert space of N qudits. Let $G = \prod_{i=1}^N \mathrm{U}(d)$, and

$$\begin{aligned} \tau : G &\to \mathrm{U}(\mathcal{H}) \\ (g_1, \cdots, g_N) &\mapsto g_1 \otimes \cdots \otimes g_N. \end{aligned}$$

So $\tau(G) \subset \mathrm{U}(\mathcal{H})$ consists of the local unitaries. The Lie algebra is $\mathfrak{g} = \bigoplus_{i=1}^N \mathfrak{u}(d)$, and $\tau_*(i\mathfrak{g}) \subset i\mathfrak{u}(\mathcal{H})$ is the Lie algebra of local observables. That is, $\tau_*(\mathfrak{g}) \subset \mathfrak{u}(\mathcal{H})$ consists of operators of the form $h_1 \otimes \mathbb{1}^{\otimes(N-1)} + \cdots + \mathbb{1}^{\otimes(N-1)} \otimes h_N$, where $h_i \in \mathfrak{u}(d)$. The density map $\tau^* : i\mathfrak{u}(\mathcal{H})^* \to i\mathfrak{g}^*$ sends a state $\Gamma \in i\mathfrak{u}(\mathcal{H})^*$ to the list of partial traces

$$(\mathrm{Tr}_{2,\cdots,N}\,\Gamma, \mathrm{Tr}_{1,3,\cdots,N}\,\Gamma, \cdots, \mathrm{Tr}_{1,\cdots,N-1}\,\Gamma).$$

The advantage of studying momentum map functional theories is that the pure and ensemble state representability problems are not completely out of reach as compared to, for example, the two-body problem (see Example 2.4). This is because we can apply the machinery developed in Chapter 5 to investigate the density map τ^*:

Proposition 6.1. *The composition*

$$\mathbb{P}(\mathcal{H}) \to \mathcal{P} \xrightarrow{\tau^*} i\mathfrak{g}^* \xrightarrow{\times i} \mathfrak{g}^*$$

is a momentum map.

Proof. See Corollary 5.18. □

Let $(G, \mathcal{H}, \tau, W)$ be a momentum map functional theory. Fix a Cartan subalgebra $\mathfrak{t} \subset \mathfrak{g}$ and a G-invariant inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$, so that we can think of $\mathfrak{t}^*$ as a subspace of $\mathfrak{g}^*$. Pick a Weyl chamber $\mathfrak{t}_+^* \subset \mathfrak{g}^*$. Let $\Omega \subset i\mathfrak{t}^*$ denote the set of weights of the representation $\tau : \mathfrak{g} \to \mathfrak{u}(\mathcal{H})$ relative to $\mathfrak{t}$. For each weight $\omega \in \Omega$, let $\mathcal{H}_\omega$ stand for the weight space, which is defined as $\{|\Psi\rangle \in \mathcal{H} \mid \forall A \in \mathfrak{t} : \tau(A)\,|\Psi\rangle = \langle \omega, A \rangle\,|\Psi\rangle\}$. Then we have the weight space decomposition

$$\mathcal{H} = \bigoplus_{\omega \in \Omega} \mathcal{H}_\omega.$$

Proposition 6.2. *The set of pure state representable densities $\tau^*(\mathcal{P}) \subset i\mathfrak{g}^*$ intersects $i\mathfrak{t}_+^*$ in a convex polytope Λ, which we will refer to from now on as the* ***Kirwan polytope****.*

Proof. This is a straightforward corollary to Proposition 6.1 and Kirwan's theorem (Theorem 5.15). □

Example 6.3. In the case of RDMFT, this means that the spectra of the 1RDMs which come from N-particle pure states form a convex polytope.

The ensemble state representability problem in this case is still easy:

Proposition 6.3. *The intersection of $i\mathfrak{t}^*$ and the set of ensemble state representable densities $\tau^*(\mathcal{E})$ is exactly* $\mathrm{conv}(\Omega)$.

Proof. Clearly, $\tau^*(\mathcal{E}) \cap i\mathfrak{t}^*$ is convex and contains Ω, so $\tau^*(\mathcal{E}) \cap i\mathfrak{t}^* \supset \mathrm{conv}(\Omega)$. Conversely, suppose

$$\rho = \tau^* \left(\sum_i t_i \, |\Psi_i\rangle\langle\Psi_i| \right) \in i\mathfrak{t}^*.$$

For each i, decompose $|\Psi_i\rangle = \sum_{\omega \in \Omega} |\Psi_i^\omega\rangle$, then

$$\langle \rho, v \rangle = \sum_i \sum_{\omega,\omega' \in \Omega} t_i \, \langle \Psi_i^\omega | \tau(v) | \Psi_i^{\omega'} \rangle = \sum_i \sum_{\omega \in \Omega} t_i \|\Psi_i^\omega\|^2 \, \langle \omega, v \rangle$$

for any $v \in i\mathfrak{t}$, so $\rho \in \mathrm{conv}(\Omega)$. Hence $\tau^*(\mathcal{E}) \cap i\mathfrak{t}^* \subset \mathrm{conv}(\Omega)$. □

In the following, we will always assume that the Kirwan polytope $\Lambda := \tau^*(\mathcal{P}) \cap i\mathfrak{t}_+^*$ has full dimension in $i\mathfrak{t}^*$, which implies that $\mathrm{conv}(\Omega)$ has full dimension in $i\mathfrak{t}^*$. Note that the converse is not true in general. Indeed, consider the "trivial" functional theory on $\mathcal{H} = \mathbb{C}^N$ (the interaction W is irrelevant here), where $G = \mathrm{SU}(N)$ and $\tau : \mathrm{SU}(N) \hookrightarrow \mathrm{U}(N)$ is the inclusion. A Cartan subalgebra for $\mathfrak{g} = \mathfrak{su}(N)$ is the set $\mathfrak{t}$ of imaginary diagonal traceless matrices, and we can take $\mathfrak{t}_+^*$ to be the subset of matrices whose diagonal entries times i are in nonincreasing order. Then

$$\Lambda = \tau^*(\mathcal{P}) \cap i\mathfrak{t}_+^* = \left\{ \mathrm{diag}\left(\frac{N-1}{N}, -\frac{1}{N}, \cdots, -\frac{1}{N} \right) \right\}$$

is zero-dimensional, yet $\mathrm{conv}(\Omega)$ clearly has full dimension in $i\mathfrak{t}^*$.

Let $R \subset i\mathfrak{t}^*$ be the set of roots, for which we have the root space decomposition

$$\mathfrak{g}_{\mathbb{C}} = \mathfrak{t}_{\mathbb{C}} \oplus \bigoplus_{\alpha \in R} \mathfrak{g}'_\alpha.$$

For each $\alpha \in R$, let $(L_\alpha^+, L_\alpha^-, H_\alpha)$ be an $\mathfrak{sl}(2,\mathbb{C})$-tuple, meaning in particular that $L_\alpha^\pm \in \mathfrak{g}'_{\pm\alpha}$, $H_\alpha \in \mathfrak{t}_{\mathbb{C}}$, and $[L_\alpha^+, L_\alpha^-] = H_\alpha$. Let R_+ be the set of positive roots relative to $\mathfrak{t}_+^*$, and write

$$\mathfrak{g} = \mathfrak{t} \oplus \bigoplus_{\alpha \in R_+} (i\mathbb{R}X_\alpha) \oplus \bigoplus_{\alpha \in R_+} (i\mathbb{R}Y_\alpha),$$

where $X_\alpha := L_\alpha^+ + L_\alpha^-$ and $Y_\alpha := -iL_\alpha^+ + iL_\alpha^-$.

As in the abelian case, we would like to understand the behavior of the pure functional near the boundary of its domain. Ideally, we would prove a generalization of Theorem 4.8, but it is not at all clear how one would adapt the proof we presented in Chapter 4 for nonabelian $\mathfrak{g}$. For example, it is not obvious anymore whether the fiber of the density map can be understood by first studying the classical density map (Definition 4.5), as we did in Chapter 4 for the proof of the boundary force using constrained search. To see what kind of complications might arise, consider a state $|\Psi\rangle = \sum_{\omega \in \Omega} c_\omega \, |E_\omega\rangle$, where $(|E_\omega\rangle)_{\omega \in \Omega}$ is a basis for $\mathcal{H}$ (assuming that all weights have multiplicity one for simplicity), satisfying $\sum_{\omega \in \Omega} |c_\omega|^2 \omega = \rho \in i\mathfrak{t}^*$. We still have $\langle \tau^*(|\Psi\rangle\langle\Psi|), v \rangle = \langle \rho, v \rangle$ for all $v \in i\mathfrak{t}$. However, we cannot conclude immediately anymore that $\tau^*(|\Psi\rangle\langle\Psi|) = \rho$ as in the abelian setting, because $\tau^*(|\Psi\rangle\langle\Psi|)$ might not vanish on X_α or Y_α. In other words, $\tau^*(|\Psi\rangle\langle\Psi|)$ might not even be in $i\mathfrak{t}^*$!

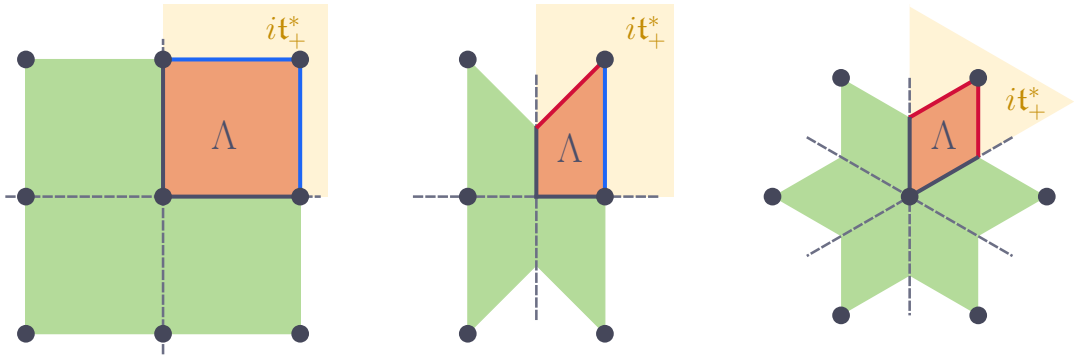
Nevertheless, it is still possible to obtain a formula analogous to Eq. (4.4) using perturbation theory, in a manner similar to the argument presented in Section 4.3.1 before the rigorous

proof in Section 4.3.2. As we will see, however, the perturbative method works only for the so-called *nice* facets, which we will now define.

First, note that we should not expect the boundary force to be present near all facets of the convex polytope $\Lambda \subset i\mathfrak{t}_+^*$. Intuitively, a facet that is entirely contained in $i\mathfrak{t}_+^* \setminus i\mathfrak{t}_{>0}^*$ (the latter being the open Weyl chamber) has nothing to do with the "true boundary" of the domain of the pure functional. Following Ref. [105], we will call such a facet $F \subset \Lambda$ *trivial*.

Definition 6.2. *A nontrivial facet $F \subset \Lambda$ is said to be **nice** if F is contained in a facet of* $\mathrm{conv}(\Omega)$.

Fig. 6.1 shows several examples of trivial, nontrivial and nice, and nontrivial and not nice facets. Informally speaking, a nice facet of Λ is one which remains after convexifying $\tau^*(\mathcal{P}) \cap i\mathfrak{t}^*$, which is the union of Λ and its images under the Weyl group action. If a facet F' of $\mathrm{conv}(\Omega)$ is the union of a nice facet of Λ and its images under the Weyl group, we will also call F' nice.



(a) **(b)** **(c)**

Figure 6.1 Trivial (black), nice (blue), and nontrivial not nice (red) facets. (a): $G = \mathrm{SU}(2) \times \mathrm{SU}(2)$, $\mathcal{H} = \mathbb{C}^3 \otimes \mathbb{C}^3$. (b): $G = \mathrm{SU}(2) \times \mathrm{SU}(2)$, $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^3$. (c): $G = \mathrm{SU}(3)$, $\mathcal{H} = \mathfrak{sl}(3, \mathbb{C}) \cong \mathbb{C}^8$. All representations are the obvious ones. The gray dashed lines are the Weyl group reflection hyperplanes.

If $F \subset \Lambda$ is a nice facet, we will let Ω_F denote the subset of weights that lie on the affine hull of F. Thus, in general, F will be a proper subset of $\mathrm{conv}(\Omega_F)$. Similarly, $\mathcal{H}_F \subset \mathcal{H}$ denotes the subspace spanned by the weight spaces of weights in Ω_F, and $\Pi_F : \mathcal{H} \to \mathcal{H}_F$ is the orthogonal projection.

As in the case of abelian functional theories, a notion of the restricted functional theory to a facet F will be useful for the boundary force derivation (see Definition 4.3 for the definition in the abelian case), although it is not immediately clear what the correct definition should be when $\mathfrak{g}$ is not abelian. In the abelian setting, all we had to do was restrict the Hilbert space to the subspace spanned by the weight spaces on the facet. However, doing so is not good enough for nonabelian functional theories: for arbitrary $g \in G$, the operator $\Pi_F \tau(g) \Pi_F^\dagger \in \mathrm{End}(\mathcal{H}_F)$ might fail to be unitary. In fact, $g \mapsto \Pi_F \tau(g) \Pi_F^\dagger$ might fail to be a representation at all! Hence, naively applying Definition 4.3 will in general result in a functional theory which is not a momentum map functional theory anymore. What follows is a definition which turns out to play well with boundary force calculations presented in Section 6.3.

Definition 6.3. *Let $F \subset \mathrm{conv}(\Omega)$ be a nice facet, and let $\langle \cdot, S\rangle \geq c$ be a corresponding inequality, where $S \in i\mathfrak{t}$ and $c \in \mathbb{R}$. The* ***restricted functional theory on the facet*** *F (or simply* ***facet functional theory****) is the momentum map functional theory given by $(G_F, \mathcal{H}_F, \tau_F, \Pi_F W \Pi_F^\dagger)$, where G_F is the identity component of $C_G(S)$ (the centralizer) and*

$$\mathcal{H}_F := \bigoplus_{\omega \in \Omega_F} \mathcal{H}_\omega \qquad \tau_F(g) := \Pi_F \tau(g) \Pi_F^\dagger,$$

Note: (1) G_F is compact because $C_G(S)$ is; (2) If $|\Psi\rangle \in \mathcal{H}_F$, then $\tau_*(S)\tau(g)\,|\Psi\rangle = \tau(g)\tau_*(S)\,|\Psi\rangle = c\tau(g)\,|\Psi\rangle$ for all $g \in G_F$, so $\tau(g)\,|\Psi\rangle \in \mathcal{H}_F$. This shows that $\mathcal{H}_F$ is stable under the G-action.

Example 6.4. If $\mathfrak{g}$ is abelian, then $G_F = G$ because G is connected and compact, so we recover Definition 4.3.

We will also define

$$\mathfrak{g}^{\nparallel} := \bigoplus_{\alpha \in R_+ : \langle \alpha, S\rangle \neq 0} (i\mathbb{R}X_\alpha \oplus i\mathbb{R}Y_\alpha)$$
$$\mathfrak{g}^{\parallel} := \bigoplus_{\alpha \in R_+ : \langle \alpha, S\rangle = 0} (i\mathbb{R}X_\alpha \oplus i\mathbb{R}Y_\alpha).$$

Note that $\mathfrak{g}^{\nparallel} = 0$ if $\mathfrak{g}$ is abelian.

Proposition 6.4. *$\mathfrak{t} \oplus \mathfrak{g}^{\parallel} = \mathfrak{g}_F$, where $\mathfrak{g}_F$ is the Lie algebra of G_F.*

Proof. Suppose $\alpha \in R$ is parallel to the facet, i.e., $\langle \alpha, S\rangle = 0$. Then $[L_\alpha^\pm, S] = \mp \langle \alpha, S\rangle L_\alpha^\pm = 0$, so $L_\alpha^\pm \in C_{\mathfrak{g}}(S) = \mathfrak{g}_F$ and consequently $iX_\alpha, iY_\alpha \in \mathfrak{g}_F$. On the other hand, if $B \in \mathfrak{t}$, then $[B, S] = 0$ trivially. This shows "$\subset$".

Now take any $B \in \mathfrak{g}$ such that $[B, S] = 0$. Decompose B in $\mathfrak{g}' = \mathfrak{g}_{\mathbb{C}}$ into $B = B_0 + \sum_{\alpha \in R} c_\alpha L_\alpha^+$ with $B_0 \in \mathfrak{t}$. Then $[B, S] = 0$ implies $\sum_{\alpha \in R} c_\alpha \langle \alpha, S\rangle L_\alpha^+ = 0$. That is, $c_\alpha = 0$ whenever $\langle \alpha, S\rangle \neq 0$. This shows $B \in \mathfrak{t} \oplus \mathfrak{g}^{\parallel}$. □

Example 6.5. Consider the action of $\mathrm{SU}(2) \times \mathrm{SU}(2)$ on $\mathbb{C}^2 \otimes \mathbb{C}^3$ as in Example 5.1. We will take F to be the only nice facet of Λ (see Fig. 6.1b), for which S can be taken to be $-Z_1 = (-Z, 0) \in i\mathfrak{su}(2) \oplus i\mathfrak{su}(2)$. Then

$$\begin{aligned}\mathfrak{g}_F &= C_{\mathfrak{g}}(-Z_1) = \mathrm{span}_{\mathbb{R}}\{iZ_1, iX_2, iY_2, iZ_2\}\\ \mathfrak{g}^{\nparallel} &= \mathrm{span}_{\mathbb{R}}\{iX_1, iY_1\}\\ \mathfrak{g}^{\parallel} &= \mathrm{span}_{\mathbb{R}}\{iX_2, iY_2\}.\end{aligned}$$

The connected Lie group of the facet functional theory is then $G_F = \mathrm{U}(1) \times \mathrm{SU}(2)$, where $\mathrm{U}(1)$ is embedded into $\mathrm{SU}(2)$ by $g \mapsto \mathrm{diag}(g, g^{-1})$. The weights lying on the affine hull of F are

$$\begin{aligned}\omega_{1,2} &= (Z \mapsto 1, Z \mapsto 2)\\ \omega_{1,0} &= (Z \mapsto 1, 0)\\ \omega_{1,-2} &= (Z \mapsto 1, Z \mapsto -2),\end{aligned}$$

so the restricted Hilbert space $\mathcal{H}_F$ is the subspace of $\mathcal{H}$ spanned by $|1,2\rangle, |1,0\rangle$, and $|1,-2\rangle$ (where $|m_1, m_2\rangle$ is what we called ϕ_{m_1,m_2} in Example 5.1). Hence, we end up with an action of $\mathrm{U}(1) \times \mathrm{SU}(2)$ on $\mathcal{H}_F \cong \mathbb{C}^3$, with $\mathrm{U}(1)$ acting by $e^{i\varphi} \mapsto e^{i\varphi}$ and $\mathrm{SU}(2)$ by the three-dimensional irreducible representation.

For the nice facets, we have a boundary force formula:

Conjecture 6.1. *Let $F \subset \mathrm{conv}(\Omega)$ be a nice facet with a corresponding inequality $\langle \cdot, S\rangle \geq c$. Let $\rho_* \in F$ be a density which is "sufficiently nice". Let $\eta \in i\mathfrak{t}^*$ be inward-pointing with $\langle \eta, S\rangle = 1$. Then*

$$\frac{\mathrm{d}}{\mathrm{d}\sqrt{\epsilon}}\Big|_{\epsilon=0} \mathcal{F}_p(\rho_* + \epsilon\eta) = -2 \min_{v \in \mathfrak{g}^{\not\parallel}} \left(\sum_{\omega \in \Omega \setminus \Omega_F} \frac{\|\Pi_\omega(\tau_*(v) + W)\,|\Phi\rangle\|^2}{\langle \omega, S\rangle - c} \right)^{\frac{1}{2}}. \tag{6.1}$$

This result is labeled as a conjecture because we only have a perturbative derivation, which relies heavily on various strong assumptions, just like what was necessary in Section 4.3.1 for the perturbative derivation of the abelian boundary force (Eq. (4.4)). Currently, we do not have a proof of Conjecture 6.1 with constrained search. In the statement of Conjecture 6.1, a density $\rho_* \in F$ being "nice" is an unspecified condition that we hope to make more precise in the process of finding a rigorous proof. Intuitively, this condition should at least resemble the regularity condition in the abelian case (see the statement of Theorem 4.8).

The reason for isolating the nice facets for boundary forces is that the bad facets of Λ are exactly those which cannot be reached, in loose terms, by "sending the external potential to infinity in the correct direction". It is conceivable that the pure functional might behave quite differently near bad facets, although there are analytical arguments (see the second half of Chapter 7) which suggest that a boundary force may still exist.

In the rest of this chapter, we will first show a concrete derivation of the boundary force for the bosonic Hubbard dimer in the next section, followed by a perturbative calculation in the general setting for obtaining Eq. (6.1) in the last section.

6.2 The Bosonic Hubbard Dimer

Before showing Eq. (6.1) for general momentum map functional theories, we will work out in this section the boundary force of the bosonic Hubbard dimer, for which a derivation can be found in Ref. [25]. However, the approach here will be easier to adapt to general functional theories. The reason for discussing the dimer first is that it captures one of the essential differences between abelian and nonabelian functional theories, namely that $\mathfrak{g}^{\not\parallel}$ is not trivial, resulting in, as we will see, an additional optimization over the subspace spanned by the nonparallel root vectors. Nevertheless, the dimer is still slightly too simple to be representative of all momentum map functional theories. For one, the polytope Λ is one-dimensional, and has only one trivial facet and one nice facet. More importantly, the facet functional theory has a one-dimensional Hilbert space, which eliminates some conceptual difficulties encountered in the general derivation presented in Section 6.3.

The bosonic Hubbard dimer is a system consisting of N identical (spinless) bosons on two lattice sites interacting via the Hubbard interaction. To be more precise, the single-particle Hilbert space is $\mathbb{C}^2$, and the N-particle Hilbert space is $\mathcal{H} = \mathrm{Sym}^N(\mathbb{C}^2) \cong \mathbb{C}^{N+1}$, which has an orthonormal basis given by

$$|N,0\rangle, |N-1,1\rangle, \cdots, |0,N\rangle,$$

with the state $|m, N-m\rangle$ having m bosons on the first site and $N-m$ on the second one. The boson-boson interaction is taken to be

$$W = n_1^2 + n_2^2 = (b_1^\dagger b_1)^2 + (b_2^\dagger b_2)^2,$$

where $n_i, b_i, b_i^\dagger$ are the number operator, annihilation operator, and creation operator of the i-th site respectively.

To define a momentum map functional theory, we need to specify a compact connected Lie group G and a unitary representation $\tau : G \to \mathrm{U}(\mathcal{H})$. We take $G = \mathrm{SU}(2)$, with $\tau : \mathrm{SU}(2) \to \mathrm{U}(\mathcal{H})$, $g \mapsto g \otimes \cdots \otimes g$, identifying $\mathrm{Sym}^N(\mathbb{C}^2)$ with a subspace of $(\mathbb{C}^2)^{\otimes N}$. The potential map is then

$$\begin{aligned} \tau_* : i\mathfrak{su}(2) &\to i\mathfrak{u}(\mathcal{H}) \\ v &\mapsto v \otimes \mathbb{1}^{\otimes(N-1)} + \cdots + \mathbb{1}^{\otimes(N-1)} \otimes v. \end{aligned}$$

In the following, we will identify $i\mathfrak{su}(2)$ with $\mathbb{R}^3$ via the basis consisting of the Pauli matrices X, Y, Z, which also gives an identification $i\mathfrak{su}(2)^* \subset \mathbb{R}^3$, so a density $\rho \in i\mathfrak{su}(2)^*$ will be thought of as a vector in $\mathbb{R}^3$. One should keep in mind that the Hilbert-Schmidt inner product on $i\mathfrak{su}(2)$ is twice the standard inner product on $\mathbb{R}^3$, and we still use the latter to identify $(\mathbb{R}^3)^* \cong \mathbb{R}^3$. Under this identification, the potential and density maps are explicitly given by

$$\begin{aligned} &\tau_* : \mathbb{R}^3 \to i\mathfrak{u}(\mathcal{H}) \\ &v \mapsto v_1(b_1^\dagger b_2 + b_2^\dagger b_1) + v_2(-ib_1^\dagger b_2 + ib_2^\dagger b_1) + v_3(b_1^\dagger b_1 - b_2^\dagger b_2) \end{aligned} \tag{6.2}$$

$$\begin{aligned} &\tau^* : \mathcal{E} \to \mathbb{R}^3 \\ &\Gamma \mapsto \left(\langle \Gamma, b_1^\dagger b_2 + b_2^\dagger b_1 \rangle, \langle \Gamma, -ib_1^\dagger b_2 + ib_2^\dagger b_1 \rangle, \langle \Gamma, b_1^\dagger b_1 - b_2^\dagger b_2 \rangle \right). \end{aligned} \tag{6.3}$$

If $N = 1$, then $\mathcal{H} = \mathbb{C}^2$ and $b_1^\dagger b_2 + b_2^\dagger b_1 = X$, $-ib_1^\dagger b_2 + ib_2^\dagger b_1 = Y$, $b_1^\dagger b_1 - b_2^\dagger b_2 = Z$. Consequently, $\tau^*(\mathcal{P}) \subset \mathbb{R}^3$ is just the unit sphere (Bloch sphere). If $N > 1$, consider the state

$$|\Psi\rangle := \frac{1}{\sqrt{2}}(|N, 0\rangle + |0, N\rangle),$$

for which $\langle\Psi|b_1^\dagger b_2|\Psi\rangle = \langle\Psi|b_2^\dagger b_1|\Psi\rangle = \langle\Psi|b_1^\dagger b_1 - b_2^\dagger b_2|\Psi\rangle = 0$, so $\tau^*(|\Psi\rangle\langle\Psi|) = 0$. Since $\tau^*(|N, 0\rangle\langle N, 0|) = (0, 0, N)$, we have shown that both 0 and $N\alpha_3$ belong to $\tau^*(\mathcal{P})$, where $(\alpha_i)_{i=1}^3$ is the standard basis on $\mathbb{R}^3$. By Kirwan's theorem, $\tau^*(\mathcal{P}) \cap \mathbb{R}_{\geq 0}\alpha_3$ is a convex polytope, so

$$\tau^*(\mathcal{P}) \cap \mathbb{R}_{\geq 0}\alpha_3 \supset \{t\alpha_3 \mid t \in [0, N]\}.$$

It is easy to show the inclusion in the other direction. By the SU(2)-equivariance of τ^*, then, we have $\tau^*(\mathcal{P}) = \{\sum_{i=1}^3 t_i\alpha_i \mid \sum_{i=1} t_i^2 \leq N^2\}$.

To summarize the discussion above, the solution to the pure state representability problem for the bosonic dimer is given by

$$\operatorname{dom} \mathcal{F}_p = \tau^*(\mathcal{P}) = \begin{cases} \text{sphere of radius } 1 & N = 1 \\ \text{ball of radius } N & N > 1. \end{cases} \tag{6.4}$$

From now on, we will assume $N > 1$. The boundary of $\operatorname{dom} \mathcal{F}_p$ is then the sphere of radius N.

Fix an angle θ. Consider the density

$$\rho_* = (N \sin\theta, 0, N\cos\theta) \tag{6.5}$$

and the basis vectors for $\mathbb{R}^3$ (see Fig. 6.2)

$$\begin{aligned} e_0 &:= (\sin\theta, 0, \cos\theta) \\ e_1 &:= (\cos\theta, 0, -\sin\theta) \\ e_2 &:= (0, 1, 0). \end{aligned} \tag{6.6}$$

Let $\varepsilon_0, \varepsilon_1, \varepsilon_2 \in (\mathbb{R}^3)^*$ be the corresponding dual basis (which are the same as the e_i).

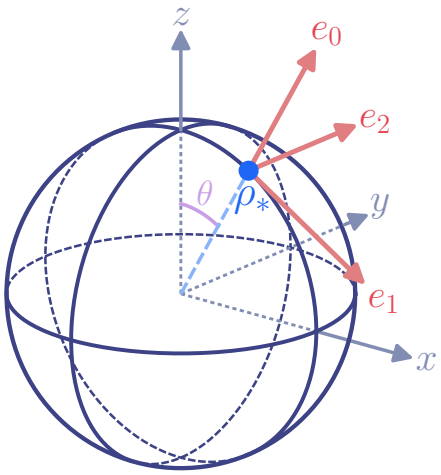


Figure 6.2 The chosen density ρ_* (blue) on the boundary and the corresponding basis vectors e_i (red). The dark blue sphere represents the boundary of $\operatorname{dom}\mathcal{F}_p$, which is the ball of radius N.

For fixed $x_1, x_2 \in \mathbb{R}$, consider a one-parameter family of Hamiltonians

$$H_g := \tau_*(-e_0) + g(\tau_*(x_1 e_1 + x_2 e_2) + W). \tag{6.7}$$

It is easy to check that $\tau_*(-e_0) = -\sin\theta(b_1^\dagger b_2 + b_2^\dagger b_1) - \cos\theta(b_1^\dagger b_1 - b_2^\dagger b_2)$ has a nondegenerate ground state given by

$$|\Phi_0\rangle = \frac{1}{\sqrt{N!}}(a^\dagger)^N |\rangle\,, \tag{6.8}$$

where

$$a^\dagger := b_1^\dagger \cos\frac{\theta}{2} + b_2^\dagger \sin\frac{\theta}{2} \qquad a_\perp^\dagger := b_1^\dagger \sin\frac{\theta}{2} - b_2^\dagger \cos\frac{\theta}{2} \tag{6.9}$$

and $|\rangle$ is the zero-boson state in the Fock space. In general, the m-th excited state of $\tau_*(-e_0)$ is

$$|\Phi_m\rangle = \frac{1}{\sqrt{(N-m)!}}\frac{1}{\sqrt{m!}}(a^\dagger)^{N-m}(a_\perp^\dagger)^m |\rangle\,, \tag{6.10}$$

with eigenvalue $\lambda_m = \langle\Phi_m|\tau_*(-e_0)|\Phi_m\rangle = 2m - N$. In the language of representation theory, the vectors $|\Phi_m\rangle$ are the weight vectors of the representation τ_* relative to the one-dimensional Cartan subalgebra $\mathfrak{t}$ spanned by $e_0 \in i\mathfrak{su}(2)$. The appropriate Weyl chamber here is $i\mathfrak{t}_+^* =$

$\mathbb{R}_{\geq 0}\varepsilon_0 \subset i\mathfrak{t}^*$, and the Kirwan polytope $\Lambda := \tau^*(\mathcal{P}) \cap i\mathfrak{t}^*_+$ is the line segment $[0, N] \cdot \varepsilon_0$, which has 0 as the only trivial facet and ρ_* as the only nice facet. The images of $|\Phi_m\rangle\langle\Phi_m|$ under the density map τ^* are shown in Fig. 6.3.

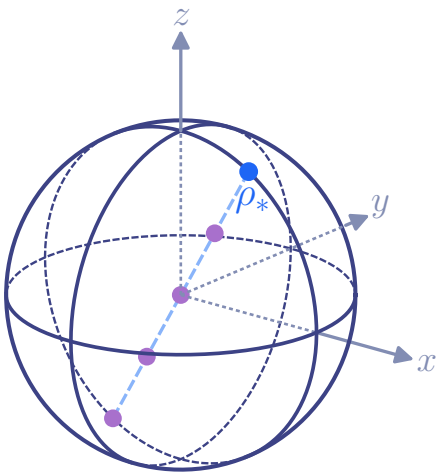


Figure 6.3 Images of the pure states $|\Phi_m\rangle\langle\Phi_m|$ under τ^* (purple: $m > 0$, blue: $m = 0$).

For small $|g|$, the lowest eigenvalue of (6.7) up to second order in g is

$$\lambda_0(g) \approx \lambda_0 + gE_1(x_1, x_2) + g^2 E_2(x_1, x_2), \tag{6.11}$$

with

$$\begin{aligned} E_1(x_1, x_2) &= \langle\Phi_0|\tau_*(x_1 e_1 + x_2 e_2) + W|\Phi_0\rangle \\ E_2(x_1, x_2) &= \sum_{m=1}^{N} \frac{|\langle\Phi_m|\tau_*(x_1 e_1 + x_2 e_2) + W|\Phi_0\rangle|^2}{\lambda_0 - \lambda_m} = -\sum_{m=1}^{N} \frac{|\langle\Phi_m|\tau_*(x_1 e_1 + x_2 e_2) + W|\Phi_0\rangle|^2}{2m}. \end{aligned} \tag{6.12}$$

In other words, the value of the energy function E at $v \in \mathbb{R}^3$, which is by definition the ground state energy of the Hamiltonian

$$\begin{aligned} H(v) &= \tau_*(v) + W = \sum_{i=1}^{3} \langle\varepsilon_i, v\rangle\, \tau_*(e_i) + W \\ &= -\langle\varepsilon_0, v\rangle \left[-\tau_*(e_0) + \frac{1}{-\langle\varepsilon_0, v\rangle} \Big(\tau_*(\langle\varepsilon_1, v\rangle\, e_1 + \langle\varepsilon_2, v\rangle\, e_2) + W \Big) \right], \end{aligned} \tag{6.13}$$

is given by, due to Eq. (6.11),

$$\begin{aligned} E(h) \approx -\langle\varepsilon_0, v\rangle \Bigg[\lambda_0 &+ \frac{1}{-\langle\varepsilon_0, v\rangle} \langle\Phi_0|\tau_*\left(\langle\varepsilon_1, v\rangle\, e_1 + \langle\varepsilon_2, v\rangle\, e_2\right) + W|\Phi_0\rangle \\ &+ \left(\frac{1}{-\langle\varepsilon_0, v\rangle}\right)^2 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) \Bigg]. \end{aligned} \tag{6.14}$$

Rearranging, we get

$$\begin{aligned} E(v) &\approx -\langle\varepsilon_0, v\rangle\,\lambda_0 + \langle\Phi_0|\tau_*(\langle\varepsilon_1, v\rangle\,e_1 + \langle\varepsilon_2, v\rangle\,e_2) + W|\Phi_0\rangle - \frac{1}{\langle\varepsilon_0, v\rangle}E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) \\ &= \langle\Phi_0|\tau_*(v) + W|\Phi_0\rangle - \frac{1}{\langle\varepsilon_0, v\rangle}E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle). \end{aligned} \tag{6.15}$$

It is important to remember that Eq. (6.15) only holds when $\langle\varepsilon_0, v\rangle < 0$ and $|\langle\varepsilon_0, v\rangle|$ is large enough. The derivative is

$$\begin{aligned} \mathrm{d}_v E \approx \tau^*(|\Phi_0\rangle\langle\Phi_0|) &+ \frac{\varepsilon_0}{\langle\varepsilon_0, v\rangle^2}E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) \\ &- \frac{1}{\langle\varepsilon_0, v\rangle}\Big(\varepsilon_1\partial_1 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) + \varepsilon_2\partial_2 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle)\Big). \end{aligned} \tag{6.16}$$

Note that the first term $\tau^*(|\Phi_0\rangle\langle\Phi_0|)$ is nothing but ρ_*. Now, we are interested in the derivative of the functional at $\rho_* - \epsilon\varepsilon_0$ with ϵ small. The plan is the following: if we can find v so that $\mathrm{d}_v E = \rho_* - \epsilon\varepsilon_0$, then we will have $\mathrm{d}_{\rho_*-\epsilon\varepsilon_0}\mathcal{F} = -v$. The condition $\mathrm{d}_v E = \rho_* - \epsilon\varepsilon_0$ gives

$$\begin{cases} -\epsilon = \frac{1}{\langle\varepsilon_0,v\rangle^2}E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) \\ 0 = -\frac{1}{\langle\varepsilon_0,v\rangle}\partial_1 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) \\ 0 = -\frac{1}{\langle\varepsilon_0,v\rangle}\partial_2 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) \end{cases} \Rightarrow \begin{cases} \langle\varepsilon_0, v\rangle = -\frac{1}{\sqrt{\epsilon}}\sqrt{-E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle)} \\ \partial_1 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) = 0 \\ \partial_2 E_2(\langle\varepsilon_1, v\rangle, \langle\varepsilon_2, v\rangle) = 0, \end{cases} \tag{6.17}$$

where we ignored the higher order terms altogether. To proceed, note that

$$\tau_*(e_1) = a_\perp^\dagger a + a^\dagger a_\perp \qquad \tau_*(e_2) = -ia_\perp^\dagger a + ia^\dagger a_\perp. \tag{6.18}$$

Therefore, $\langle\Phi_m|\tau_*(e_1)|\Phi_0\rangle$ is real and $\langle\Phi_m|\tau_*(e_2)|\Phi_0\rangle$ is imaginary. Moreover, it is easy to see that $\langle\Phi_m|W|\Phi_0\rangle$ is real. For Eq. (6.17) to hold, then, it must be that $\langle\varepsilon_2, v\rangle = 0$. We work out explicitly the matrix elements of $\tau_*(e_1)$ and W:

$$\langle\Phi_m|\tau_*(e_1)|\Phi_0\rangle = \langle\Phi_m|a_\perp^\dagger a + a^\dagger a_\perp|\Phi_0\rangle = \langle\Phi_m|a_\perp^\dagger a|\Phi_0\rangle = \delta_{m,1}\sqrt{N} \tag{6.19}$$

$$\begin{aligned} \langle\Phi_1|W|\Phi_0\rangle &= (N-1)\sqrt{N}\sin\theta\cos\theta \\ \langle\Phi_2|W|\Phi_0\rangle &= \frac{\sqrt{N(N-1)}}{\sqrt{2}}\sin^2\theta \\ \langle\Phi_{m>2}|W|\Phi_0\rangle &= 0. \end{aligned} \tag{6.20}$$

Therefore

$$E_2(\langle\varepsilon_1, v\rangle, 0) = -\frac{1}{2}\left|\langle\varepsilon_1, v\rangle\sqrt{N} + (N-1)\sqrt{N}\sin\theta\cos\theta\right|^2 - \frac{1}{4}\left|\frac{\sqrt{N(N-1)}}{\sqrt{2}}\sin^2\theta\right|^2. \tag{6.21}$$

According to Eq. (6.17), the value of $\langle\varepsilon_1, v\rangle$ is fixed by requiring that E_2 be stationary. Clearly, then, we should choose $\langle\varepsilon_1, v\rangle$ so that the first term of Eq. (6.21) vanishes, so the stationary value of E_2 is $-\frac{1}{8}N(N-1)\sin^4\theta$. The derivative of the functional is then

$$\begin{aligned} \mathrm{d}_{\rho_*-\epsilon\varepsilon_0}\mathcal{F} &= -v = -\langle\varepsilon_0, v\rangle\,e_0 - \langle\varepsilon_1, v\rangle\,e_1 \\ &\approx \frac{e_0}{\sqrt{\epsilon}}\sqrt{\frac{1}{8}N(N-1)\sin^4\theta} + e_1(N-1)\sin\theta\cos\theta \\ &= \frac{e_0}{\sqrt{\epsilon}}\frac{\sqrt{N(N-1)}}{2\sqrt{2}}\sin^2\theta + e_1\frac{N-1}{2}\sin 2\theta. \end{aligned} \tag{6.22}$$

Finally, we obtain the directional derivative along $-\varepsilon_0$:

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* - \epsilon\varepsilon_0) = \langle -\varepsilon_0, \mathrm{d}_{\rho_* - \epsilon\varepsilon_0}\mathcal{F}\rangle \approx -\frac{1}{\sqrt{\epsilon}}\frac{\sqrt{N(N-1)}}{2\sqrt{2}}\sin^2\theta. \tag{6.23}$$

This agrees with Eq. (12) of Ref. [25].

6.3 Perturbation Theory in General Momentum Map Functional Theories

We will now derive Eq. (6.1) by perturbation theory. After making all the necessary niceness assumptions about the functional theory as we did in Section 4.3.1 (see the actual assumptions therein), we will not distinguish between $\mathcal{F}_p$ and $\mathcal{F}_{HK}$, and will use $\mathcal{F}$ to denote both in the following.

Let $(G, \mathcal{H}, \tau, W)$ be a momentum map functional theory with a chosen Cartan subalgebra $\mathfrak{t}$ and Weyl chamber $i\mathfrak{t}^*_+$ such that the convex polytope $\Lambda = \tau^*(\mathcal{P}) \cap i\mathfrak{t}^*_+$ has full dimension in $i\mathfrak{t}^*$. Let $\Omega \subset i\mathfrak{t}^*$ denote the set of weights of the representation $\tau_* : \mathfrak{g} \to \mathfrak{u}(\mathcal{H})$. Fix a nice facet $F \subset \mathrm{conv}(\Omega)$, with associated inequality $\langle \cdot, S\rangle \geq c$, where $S \in i\mathfrak{t}$ and $c \in \mathbb{R}$, along with a density $\rho_* \in \mathrm{relint}(F)$ and an inward vector $\eta \in i\mathfrak{t}^*$. We can scale S and c by a positive number so that $\langle \eta, S\rangle = 1$.

Recall from Section 6.1 that the facet F induces a decomposition of vector spaces

$$\mathfrak{g} = \mathfrak{t} \oplus \mathfrak{g}^{\parallel} \oplus \mathfrak{g}^{\nparallel},$$

where $\mathfrak{t} \oplus \mathfrak{g}^{\parallel} = C_{\mathfrak{g}}(S) = \mathfrak{g}_F$ is the space of external potentials (up to i) of the facet functional theory. Define $U := \ker(\eta) \oplus i\mathfrak{g}^{\parallel}$, where $\ker(\eta) := \{w \in i\mathfrak{t} \mid \langle \eta, w\rangle = 0\} \subset i\mathfrak{t}$. We then have the decomposition

$$i\mathfrak{g} = \underbrace{\mathbb{R}S \oplus U}_{i\mathfrak{g}_F} \oplus i\mathfrak{g}^{\nparallel}.$$

With respect to this decomposition, define projections $\pi_S, \pi_U, \pi_{\nparallel} : i\mathfrak{g} \to i\mathfrak{g}$. That is, π_S is identity on $\mathbb{R}S$ and zero on $U \oplus \mathfrak{g}^{\nparallel}$ etc. These satisfy $\pi_S + \pi_U + \pi_{\nparallel} = \mathrm{Id}_{i\mathfrak{g}}$. Furthermore, it is easy to check that $\pi_S(\cdot) = \langle \eta, \cdot\rangle\, S$. We will also denote the projection onto the space $i\mathfrak{g}_F$ of external potentials of the facet functional theory by $\pi_F := \pi_S + \pi_U$.

Now consider the family of Hamiltonians

$$H(v) = \tau_*(v) + W = \langle \eta, v\rangle\, \tau_*(v) + \tau_*(\pi_U(v)) + \tau_*(\pi_{\nparallel}(v)) + W. \tag{6.24}$$

We will compute the ground state energy $E(v)$ in the limit $\langle \eta, v\rangle \to \infty$ using perturbation theory. Let us write Eq. (6.24) in a more suggestive form:

$$H(v) = \langle \eta, v\rangle \left[\tau_*(S) + \frac{1}{\langle \eta, v\rangle}\left(\tau_*(\pi_U(v)) + \tau_*(\pi_{\nparallel}(v)) + W\right)\right] \tag{6.25}$$

We treat everything inside of $\frac{1}{\langle \eta, v\rangle}(\cdots)$ as a small perturbation, and $\frac{1}{\langle \eta, v\rangle}$ will be the expansion parameter.

Since F is a nice facet, all eigenvalues of $\tau_*(S) - c\mathbb{1}$ are nonnegative, and the zero eigenspace is just $\mathcal{H}_F$, the Hilbert space of the facet functional theory, which is spanned by the weight vectors whose weights are on F. Therefore, at zeroth order, any state in $\mathcal{H}_F$ is a ground state of $H(v)$, with energy $c\,\langle \eta, v\rangle$.

The energy up to second order in $\langle \eta, v\rangle^{-1}$ is then

$$E(v) \approx \langle \eta, v\rangle \left[c + \frac{1}{\langle \eta, v\rangle} E_1(v) + \frac{1}{\langle \eta, v\rangle^2} E_2(v) \right], \tag{6.26}$$

where E_1 and E_2 are the first and second order energy corrections respectively. Although we write $E_1(v), E_2(v)$, each of these corrections actually does not depend on all of v, but only on certain components of v selected by some of the projections. For example, the first order energy correction is

$$E_1(v) = \langle \Phi(v) | \tau_*(\pi_U(v) + \pi_{\not\parallel}(v)) + W | \Phi(v)\rangle\,, \tag{6.27}$$

where the term coming from $\pi_{\not\parallel}(v)$ vanishes because $\langle \Phi(v)|\tau_*(\pi_{\not\parallel}(v))|\Phi(v)\rangle = 0$. In Eq. (6.27), $|\Phi(v)\rangle$ is the ground state in $\mathcal{H}_F$ of the Hamiltonian

$$\Pi_F \big(\langle \eta, v\rangle\, \tau_*(S) + \tau_*(\pi_U(v)) + W \big) \Pi_F^\dagger,$$

which is nothing but the Hamiltonian $H_F(v)$ of the facet functional theory (see Definition 6.3). This implies that the present derivation of the boundary force only works near those points on F that are pure state v-representable in the facet functional theory.

For each $\omega \in \Omega$, let $\Pi_\omega : \mathcal{H} \to \mathcal{H}_\omega$ be the orthogonal projection onto the weight space $\mathcal{H}_\omega$. The second order correction is then

$$E_2(v) = - \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\| \Pi_\omega (\tau_*(\pi_{\not\parallel}(v)) + W)) \, |\Phi(v)\rangle \|^2}{\langle \omega, S\rangle - c}, \tag{6.28}$$

where we used $\Pi_\omega \tau_*(\pi_U(v)) \, |\Phi(v)\rangle = 0$.

Note that $E_2(v)$ depends on v in two ways: through $\pi_{\not\parallel}(v)$ explicitly, and through $|\Phi(v)\rangle$, which depends only on $\pi_U(v)$.

We proceed by employing the same method as in Section 4.3.1 to compute $\mathrm{d}_{\rho_* + \epsilon\eta}\mathcal{F}$: we first determine $\mathrm{d}_v E$, and then solve $\mathrm{d}_v E = \rho_* + \epsilon\eta$ for v. Collecting everything, we get

$$\begin{aligned} E(v) \approx \langle \eta, v\rangle \Bigg[c \; &+ \frac{1}{\langle \eta, v\rangle} \langle \Phi(v) | \tau_*(\pi_U(v)) + W | \Phi(v)\rangle \\ &- \frac{1}{\langle \eta, v\rangle^2} \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\| \Pi_\omega (\tau_*(\pi_{\not\parallel}(v)) + W) \, |\Phi(v)\rangle \|^2}{\langle \omega, S\rangle - c} \Bigg] \\ = E_F(\underbrace{\pi_S(v) + \pi_U(v)}_{= \pi_F(v) \in i\mathfrak{g}_F}) &- \frac{1}{\langle \eta, v\rangle} \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\| \Pi_\omega (\tau_*(\pi_{\not\parallel}(v)) + W) \, |\Phi(v)\rangle \|^2}{\langle \omega, S\rangle - c}, \end{aligned} \tag{6.29}$$

where $E_F : i\mathfrak{g}_F \to \mathbb{R}$ is the ground state energy function of the facet functional theory. Hence

$$\begin{aligned} \mathrm{d}_v E \approx \mathrm{d}_{\pi_F(v)} E_F &+ \frac{\eta}{\langle \eta, v\rangle^2} \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\| \Pi_\omega (\tau_*(\pi_{\not\parallel}(v)) + W) \, |\Phi(v)\rangle \|^2}{\langle \omega, S\rangle - c} \\ &- \frac{1}{\langle \eta, v\rangle} \mathrm{d}_v \left\{ \sum_{\omega \in \Omega \setminus \Omega_F} \frac{\| \Pi_\omega (\tau_*(\pi_{\not\parallel}(v)) + W) \, |\Phi(v)\rangle \|^2}{\langle \omega, S\rangle - c} \right\}. \end{aligned} \tag{6.30}$$

Note that $\mathrm{d}_{\pi_F(v)}E_F$ is the ground state density corresponding to $\pi_F(v) \in i\mathfrak{g}_F = iC_{\mathfrak{g}}(S)$ in the facet functional theory. We will assume $\mathrm{d}_{\pi_F(v)}E_F \in F$. Fix a density $\rho_* \in F$, which we assume to be pure state v-representable. We now want to find a path $v(\epsilon) =: v_\epsilon$ such that $\mathrm{d}_{v_\epsilon}E = \rho_* + \epsilon\eta$. Using Eq. (6.30), this condition reads

$$\begin{aligned}
&\rho_* + \epsilon\eta \\
&\approx \mathrm{d}_{\pi_F(v)}E_F + \frac{\eta}{\langle \eta, v_\epsilon\rangle^2}\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega(\tau_*(\pi_{\nparallel}(v_\epsilon)) + W)\,|\Phi(v_\epsilon)\rangle\|^2}{\langle\omega, S\rangle - c} \\
&\qquad - \frac{1}{\langle\eta, v_\epsilon\rangle}\mathrm{d}_{v_\epsilon}\left\{\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega(\tau_*(\pi_{\nparallel}(v_\epsilon)) + W)\,|\Phi(v_\epsilon)\rangle\|^2}{\langle\omega, S\rangle - c}\right\}.
\end{aligned} \tag{6.31}$$

Evaluate both sides on S to get

$$\epsilon \approx \frac{1}{\langle\eta, v_\epsilon\rangle^2}\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega(\tau_*(\pi_{\nparallel}(v_\epsilon)) + W)\,|\Phi(v_\epsilon)\rangle\|^2}{\langle\omega, S\rangle - c},$$

where we have used $\langle\rho_*, S\rangle = \langle \mathrm{d}_{\pi_F(v)}E_F, S\rangle$, which holds because both ρ_* and $\mathrm{d}_{\pi_F(v)}E_F$ are densities on the facet F. The derivative of the functional is then

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* + \epsilon\eta) = \langle\eta, -v_\epsilon\rangle \approx -\frac{1}{\sqrt{\epsilon}}\left[\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega(\tau_*(\pi_{\nparallel}(v_\epsilon)) + W)\,|\Phi(v_\epsilon)\rangle\|^2}{\langle\omega, S\rangle - c}\right]^{\frac{1}{2}}. \tag{6.32}$$

We may replace $|\Phi(v_\epsilon)\rangle$ with $|\Phi\rangle = |\Phi(v_0)\rangle$, which results in a higher-order error. To determine $\pi_{\nparallel}(v_\epsilon)$, evaluate Eq. (6.31) on any $w \in i\mathfrak{g}^{\nparallel}$ to get

$$0 = \left\langle \mathrm{d}_{v_\epsilon}\left\{\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega(\tau_*(\pi_{\nparallel}(v_\epsilon)) + W)\,|\Phi\rangle\|^2}{\langle\omega, S\rangle - c}\right\}, w\right\rangle, \tag{6.33}$$

which means that $\pi_{\nparallel}(v_\epsilon)$ is a stationary point. Upon closer examination, one finds that $\pi_{\nparallel}(v_\epsilon)$ minimizes the square root in Eq. (6.32). It follows that

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* + \epsilon\eta) \approx -\frac{1}{\sqrt{\epsilon}}\min_{v\in i\mathfrak{g}^{\nparallel}}\left[\sum_{\omega\in\Omega\setminus\Omega_F}\frac{\|\Pi_\omega(\tau_*(v) + W)\,|\Phi\rangle\|^2}{\langle\omega, S\rangle - c}\right]^{\frac{1}{2}}, \tag{6.34}$$

which is the claimed boundary force formula Eq. (6.1).

Example 6.6. We will apply Eq. (6.34) to the functional theory defined by

$$(G = \mathrm{SU}(2)\times\mathrm{SU}(2), \tau = \tau^{(2)}\otimes\tau^{(3)}, \mathbb{C}^2\otimes\mathbb{C}^3, W),$$

where $\tau^{(2)}$ and $\tau^{(3)}$ are the two- and three-dimensional irreducible representations of SU(2) respectively.

In Example 5.1, we worked out the set of pure state representable densities $\iota^*(\mathcal{P})$, whose

intersection with $i\mathfrak{t}_+^*$ is shown in Fig. 6.1b. For the interaction W, we will take

$$\begin{aligned}
&\langle 1,2|W|-1,2\rangle = \langle -1,2|W|1,2\rangle = u_1\\
&\langle 1,0|W|-1,0\rangle = \langle -1,0|W|1,0\rangle = u_2\\
&\langle 1,-2|W|-1,-2\rangle = \langle -1,-2|W|1,-2\rangle = u_3\\
&\langle 1,2|W|1,2\rangle = k_1\\
&\langle 1,-2|W|1,-2\rangle = k_3\\
&\langle 1,2|W|1,0\rangle = \langle 1,0|W|1,2\rangle = 1\\
&\langle 1,-2|W|1,0\rangle = \langle 1,0|W|1,-2\rangle = 1,
\end{aligned}$$

where u_1, u_2, u_3, k_1, k_3 are real numbers, and $\langle i,j|W|k,l\rangle = 0$ for all matrix elements not listed above (see Fig. 6.4a).

Take $\rho_* = (Z \mapsto 1, Z \mapsto 1) \in F$, where $F \subset \Lambda$ is the unique nice facet. The inward-pointing vector $\eta \in i\mathfrak{t}^*$ and the normal vector $S \in i\mathfrak{t}$ are chosen as

$$\eta = (Z \mapsto -1, 0) \quad S = -Z_1,$$

which satisfy $\langle \eta, S\rangle = 1$.

First, we try to understand the preimage of the density ρ_* under the density map τ^*. If $\tau^*(|\Phi\rangle\langle\Phi|) = \rho_*$, then we must have

$$|\Phi\rangle = a\,|1,2\rangle + b\,|1,0\rangle + c\,|1,-2\rangle\,,$$

where b is real and nonnegative without loss of generality. The condition that $\tau^*(|\Phi\rangle\langle\Phi|) \in i\mathfrak{t}^*$ is equivalent to

$$\begin{aligned}
&\langle\Phi|\tau_*(X_1)|\Phi\rangle = \langle\Phi|\tau_*(Y_1)|\Phi\rangle = 0\\
&\langle\Phi|\tau_*(X_2)|\Phi\rangle = \langle\Phi|\tau_*(Y_2)|\Phi\rangle = 0.
\end{aligned}$$

The first condition is trivially satisfied, while the second one is equivalent to $\bar{a}b + bc = 0$. Hence, $b = 0$ or $a = -\bar{c}$. If $a = -\bar{c}$, then $\tau^*(|\Phi\rangle\langle\Phi|)$ would be $(Z \mapsto 1, 0)$, which is not equal to ρ_*. Thus, we have $b = 0$. It follows that

$$|\Phi\rangle = \frac{\sqrt{3}}{2}\,|1,2\rangle + \frac{1}{2}e^{i\phi}\,|1,-2\rangle\,,$$

where $e^{i\phi}$ is any phase. Clearly, $|\Phi\rangle$ is a minimizer of the constrained search for any ϕ.

Eq. (6.34) now reads

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* + \epsilon\eta) \approx$$
$$\frac{-1}{\sqrt{\epsilon}} \min_{v\in i\mathfrak{g}^{\mathbb{H}}} \left[\frac{|\langle -1,2|(\tau_*(v)+W)|\Phi\rangle|^2}{2} + \frac{|\langle -1,0|(\tau_*(v)+W)|\Phi\rangle|^2}{2} + \frac{|\langle -1,-2|(\tau_*(v)+W)|\Phi\rangle|^2}{2}\right]^{\frac{1}{2}}.$$

Now, $i\mathfrak{g}^{\natural} = \mathrm{span}\{X_1, Y_1\}$ (see Example 6.5), so

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathcal{F}(\rho_* + \epsilon\eta) &\approx -\frac{1}{\sqrt{\epsilon}} \min_{v \in i\mathfrak{g}^{\natural}} \left[\frac{3\,|\langle -1,2|(\tau_*(v)+W)|1,2\rangle|^2}{8} + \frac{|\langle -1,-2|(\tau_*(v)+W)|1,-2\rangle|^2}{8}\right]^{\frac{1}{2}} \\ &= -\frac{1}{2\sqrt{2}\sqrt{\epsilon}} \min_{x,y\in\mathbb{R}} \sqrt{3\,|u_1 + x + iy|^2 + |u_3 + x + iy|^2} \\ &= -\frac{1}{2\sqrt{2}\sqrt{\epsilon}} \min_{x\in\mathbb{R}} \sqrt{3(u_1+x)^2 + (u_3+x)^2} \\ &= -\frac{1}{2\sqrt{2}\sqrt{\epsilon}} \sqrt{\frac{3}{4}(u_1-u_3)^2} = -\frac{1}{\sqrt{\epsilon}} \cdot \frac{\sqrt{6}}{8}\,|u_1 - u_3|\,.\end{aligned} \tag{6.35}$$

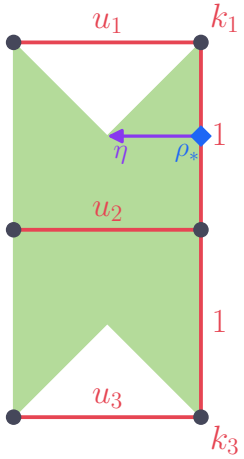
Finally, we integrate Eq. (6.35) to get an approximation to the functional itself:

$$\mathcal{F}(\rho_* + \epsilon\eta) \approx \mathcal{F}(\rho_*) - \int_0^{\epsilon} \frac{1}{\sqrt{\epsilon'}} \cdot \frac{\sqrt{6}}{8}\,|u_1 - u_3|\,\mathrm{d}\epsilon' = \mathcal{F}(\rho_*) - \sqrt{\epsilon}\frac{\sqrt{6}}{4}\,|u_1 - u_3|\,, \tag{6.36}$$

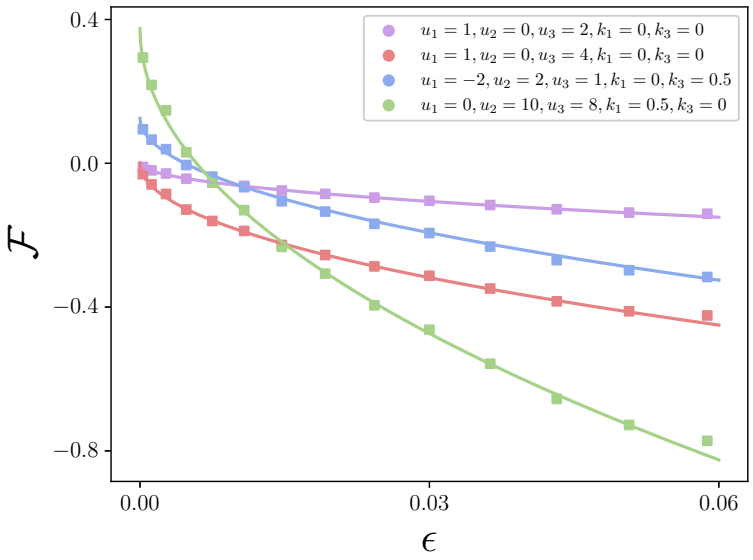
where

$$\mathcal{F}(\rho_*) = \langle\Phi|W|\Phi\rangle = \frac{3}{4}k_1 + \frac{1}{4}k_3.$$

In Fig. 6.4b, we compare Eq. (6.36) with numerically obtained values of the pure functional $\mathcal{F}_p$. The latter is achieved by executing the pure state constrained search in the following way: 2×10^6 states in the Hilbert space are sampled in a specific way so that their images under the density map all lie in $i\mathfrak{t}^*$ and are close to ρ_*. For each ϵ, the expectation value $\langle\Psi|W|\Psi\rangle$ is minimized over all states $|\Psi\rangle$ mapping sufficiently close to $\rho_* + \epsilon\eta$.



(a) Schematic illustration of the matrix elements of the interaction W. Two weights (black filled circles) are connected by a red line if the matrix element of W between the respective weight vectors is not zero.



(b) Comparison of numerical values of $\mathcal{F}_p$ (square dots) with Eq. (6.36) (solid lines).

Figure 6.4

Chapter 7

Conclusion and Outlook

In this thesis, inspired by recent observations in the RDMFT literature (namely Refs. [25, 37–39]) regarding the behavior of the universal functional near the boundary of its domain, we carried out a systematic investigation of the interplay between the functional and its domain's geometry in a way agnostic to physical details of individual systems, facilitated by the abstract notion of a generalized functional theory defined in Chapter 2. We model a functional theory by the following data: a Hilbert space $\mathcal{H}$, a fixed interaction operator W, and a linear parametrization map $\iota : V \to i\mathfrak{u}(\mathcal{H})$ called the *potential map*, the dual of which is the *density map*. In Chapter 2, we showed that these data alone already result in rich mathematical structure, in particular allowing us to establish the weak Hohenberg-Kohn theorem (Theorem 2.5) and to define the universal functionals (see Section 2.4).

Following the introduction of bosonic momentum space functional theories in Chapter 3, where a concrete functional form and a boundary force formula were derived informally, we identified in Chapter 4 a special class of generalized functional theories, called *abelian functional theories*, for which many results from Chapter 3 could be generalized. An abelian functional theory is one whose external potentials all commute, permitting a straightforward characterization of the set of representable densities in terms of *weights*, which are given by the eigenvalues of the external potentials (see Definition 4.2). In particular, we showed (Theorem 4.1) that the domains of the pure and ensemble functionals both coincide with the convex hull of all weights, which is a convex polytope. Our most important finding in Chapter 4 is Theorem 4.8, which gives an explicit formula (Eq. (4.4)) for calculating the prefactor of the diverging boundary force, strongly constraining the behavior of the functional near a facet of the polytope. Unlike previous works on this topic, we are the first to write down such a general formula and, moreover, to provide a rigorous proof based on the Levy-Lieb constrained search, which we elaborated in Section 4.3.2.

Finally, we generalized some of our results on abelian functional theories from Chapter 4 to the nonabelian setting. For the representability problem to be tractable, we restricted ourselves to a class of generalized functional theories called *momentum map functional theories* (Definition 6.1), which still include many functional theories of physical interest such as DFT, RDMFT, and spin chains with local observables as external potentials. Roughly speaking, a momentum map functional theory is one for which the space of external potentials is a Lie algebra and the density map is a momentum map, a mathematical object whose properties were discussed in Chapter 5. Because of Kirwan's theorem (Theorem 5.15), we know that

the domain of the pure functional intersects any Weyl chamber of any Cartan subalgebra in a convex polytope, for which Theorem 5.20 gives a complete characterization. We applied these mathematical statements in Chapter 6, in particular to formulate Conjecture 6.1, which is the principal result of the last part of the thesis. Conjecture 6.1 claims that for a *nice facet* of the Kirwan polytope (Definition 6.2) the counterpart of Eq. (4.4), the abelian boundary force formula, is given by Eq. (6.1), which involves an additional optimization over the span of the root spaces whose roots are not parallel to the facet. Although we do not have a rigorous proof at the moment, a derivation of Eq. (6.1) based on perturbation theory was given in Section 6.3.

Towards a Proof of Conjecture 6.1 and Its Generalization Beyond Nice Facets

Looking ahead, we outline a few ideas that could potentially lead to a rigorous proof of Conjecture 6.1 by constrained search, which may work equally well also for nontrivial facets that are not nice.

A natural strategy would be to mimic what we did for abelian functional theories in Section 4.3.2. Recall the setting: we choose a density ρ_* on a facet F of the convex polytope $\mathrm{conv}(\Omega)$, where Ω is the set of weights, and an inward-pointing vector $\eta \in \overrightarrow{\mathrm{aff}(\Omega)}$. The goal is to understand how $\mathcal{F}_p(\rho_* + \epsilon\eta)$ depends on ϵ for small ϵ. A key ingredient was a simple description of the fiber of the density map over $\rho_* + \epsilon\eta$ in terms of the fiber over ρ_*, which was achieved by Proposition 4.17. Roughly speaking, we were able to show that, up to phases, the fiber over $\rho_* + \epsilon\eta$ is almost the product of the fiber over ρ_* and a certain convex polytope, with the latter parametrizing the different ways a state can deviate in order to change the density by $\epsilon\eta$.

For nonabelian momentum map functional theories, we then have to solve the following problem about momentum maps:

Problem 7.1. *Let G be a compact connected Lie group acting unitarily on a complex inner product space $\mathcal{H}$, and let $\Phi_G : \mathbb{P}(\mathcal{H}) \to \mathfrak{g}^*$ be the natural momentum map (see Corollary 5.18). Pick a Cartan subalgebra $\mathfrak{t}$ and Weyl chamber $\mathfrak{t}_+^*$, and let $\Lambda := \Phi_G(\mathbb{P}(\mathcal{H})) \cap \mathfrak{t}_+^*$ be the Kirwan polytope, which we assume to have full dimension in $\mathfrak{t}^*$. Pick a nontrivial facet $F \subset \Lambda$, a point $\rho_* \in \mathrm{relint}(F)$, and an inward-pointing vector $\eta \in i\mathfrak{t}^*$. For small ϵ, describe the fiber $\Phi_G^{-1}(\rho_* + \epsilon\eta)$, preferably in terms of $\Phi_G^{-1}(\rho_*)$.*

We already know something useful about $\Phi_G^{-1}(\rho_*)$: by the Selection Rule (Theorem 5.19), $\Phi_G^{-1}(\rho_*)$ must be contained in $\mathbb{P}(\mathcal{H}_F)$. Ideally, we would like to say that the fiber $\Phi_G^{-1}(\rho_* + \epsilon\eta)$ is not "too different" from $\Phi_G^{-1}(\rho_*)$, in the sense that every element in the former lies within distance $\delta(\epsilon)$ of some element in the latter with respect to, say, the trace norm. A corollary of Lemma 2.14 of Ref. [105], which is essentially a consequence of the compactness of $\mathbb{P}(\mathcal{H})$ and the Selection Rule, provides a partial confirmation:

Lemma 7.1. *There exists a function $\delta(\epsilon)$ such that*

$$\forall \epsilon : \forall [\psi] \in \Phi_G^{-1}(\rho_* + \epsilon\eta) : \\ \min_{[\phi] \in \mathbb{P}(\mathcal{H}_F)} \mathrm{Tr}(|\mathbf{p}_\phi - \mathbf{p}_\psi|) \leq \delta(\epsilon)$$

and $\lim_{\epsilon \to 0} \delta(\epsilon) = 0$.

(Recall that $\mathbf{p}_\psi := \psi \langle \psi, \cdot \rangle$ denotes the projector onto $\mathbb{C}\psi$.)

However, this is too weak for the boundary force for two reasons. First, we do not have control over the minimizer $[\phi] \in \mathbb{P}(\mathcal{H}_F)$. Second, if we want to conclude $\mathcal{F}_p(\rho_* + \epsilon\eta) \approx \mathcal{F}_p(\rho_*) - \mathcal{G}\sqrt{\epsilon}$, which we believe to hold for nice facets for good reasons (with $-\mathcal{G}$ given by Eq. (6.1)), we expect $\delta(\epsilon) \sim \sqrt{\epsilon}$, which is clearly true if $\mathfrak{g}$ is abelian. Again, partial evidence for the latter claim can be found in Ref. [142], which deals with the case of fermionic RDMFT, for which $\mathcal{H} = \bigwedge^N \mathcal{H}_1$, $G = \mathrm{U}(\mathcal{H}_1)$, and Φ_G is $N\mathrm{Tr}_{N-1}$ up to a factor of i. Fix an ordered basis for $\mathcal{H}_1$, which induces a choice of Cartan subalgebra $\mathfrak{t} \subset \mathfrak{u}(\mathcal{H})$ and a Weyl chamber $\mathfrak{t}^*_+ \subset \mathfrak{t}^*$.

Theorem 7.2. *Let $\langle \cdot, S \rangle \geq c$ be an inequality corresponding to a nontrivial facet F of the convex polytope $N\mathrm{Tr}_{N-1}(\mathcal{P}) \cap i\mathfrak{t}^*_+$. There exists some constant $C > 0$ such that for $\epsilon > 0$ small enough and for any N-particle state $|\Psi\rangle$ for which $\rho := N\mathrm{Tr}_{N-1}(|\Psi\rangle\langle\Psi|) \in i\mathfrak{t}^*_{>0}$, $\langle \rho, S \rangle = \epsilon$, and ρ is not too close to $i\mathfrak{t}^*_+ \setminus i\mathfrak{t}^*_{>0}$, there exists an N-particle state $|\Phi\rangle$ such that the 1RDM $N\mathrm{Tr}_{N-1}(|\Phi\rangle\langle\Phi|)$, after conjugating by a suitable single-particle unitary, lies on F, and*

$$\|\Phi - \Psi\| \leq C\sqrt{\epsilon}.$$

However, with this theorem we do not have control over whether $N\mathrm{Tr}_{N-1}(|\Phi\rangle\langle\Phi|)$ actually lies on the facet F, which is what we need. In the quantum chemistry language, one says that the states $|\Phi\rangle$ and $|\Psi\rangle$ may not share the same natural orbitals.

Appendix A

Kähler Structure of $\mathbb{P}(\mathcal{H})$

This appendix will be dedicated to giving $\mathbb{P}(\mathcal{H})$ the structure of a Kähler manifold, in particular making $\mathbb{P}(\mathcal{H})$ into a symplectic manifold.

Recall that a Kähler manifold is a Riemannian, symplectic and complex manifold such that the three structures are compatible with each other. More precisely, a complex manifold M is Kähler if it is endowed with a Riemannian metric h and a symplectic form ω such that $h(\cdot,\cdot) = \omega(\cdot, J(\cdot))$, where $J : T_pM \to T_pM$ is the almost complex structure.

In our case, it turns out to be easier to first construct a Riemannian metric h and then (try to) define from it a symplectic form. In this case, the following lemma is useful:

Lemma A.1. *Let M be a complex manifold and let $J : TM \to TM$ denote the almost complex structure. Let h be a Riemannian metric on M and define $\omega \in \Gamma(T^*M \otimes T^*M)$ by $\omega(X,Y) = h(JX,Y)$ for tangent vectors X, Y. Then the following are equivalent:*

(1) $h(\cdot,\cdot) = \omega(\cdot, J(\cdot))$

(2) ω *is skew-symmetric, i.e.,* $\omega(X,Y) = -\omega(Y,X)$

(3) h *is invariant under* J*, i.e.,* $h(J(\cdot), J(\cdot)) = h(\cdot,\cdot)$

(4) ω *is invariant under* J*, i.e.,* $\omega(J(\cdot), J(\cdot)) = \omega(\cdot,\cdot)$

Proof. (1) $\Rightarrow$ (2): $\omega(X,Y) = h(JX,Y) = h(Y,JX) = \omega(Y,JJX) = -\omega(Y,X)$.
(2) $\Rightarrow$ (3): $h(JX,JY) = \omega(X,JY) = -\omega(JY,X) = -h(JJY,X) = h(X,Y)$.
(3) $\Rightarrow$ (4): $\omega(JX,JY) = h(JJX,JY) = h(JX,Y) = \omega(X,Y)$.
(4) $\Rightarrow$ (1): $\omega(X,JY) = \omega(JX,JJY) = h(JJX,JJY) = h(X,Y)$. □

Our plan of giving $\mathbb{P}(\mathcal{H})$ a Kähler structure is as follows:

(1) Understand the almost complex structure J on $\mathbb{P}(\mathcal{H})$.

(2) Give $\mathbb{P}(\mathcal{H})$ a Riemannian metric, called the *Fubini-Study metric* h^{FS}.

(3) With J and h^{FS}, define $\omega^{\mathrm{FS}}(\cdot,\cdot) := h^{\mathrm{FS}}(J(\cdot),\cdot)$.

(4) Show that ω^{FS} is skew-symmetric, and hence defines a 2-form.

(5) Show that ω^{FS} is closed, thus concluding that $(\mathbb{P}(\mathcal{H}), h^{\mathrm{FS}}, \omega^{\mathrm{FS}})$ is a Kähler manifold.

A.1 Almost Complex Structure

Recall the general strategy we employ to understand the tangent space $T_{[\psi]}\mathbb{P}(\mathcal{H})$ (see Remark 5.6) using the map $A \mapsto A^*_{[\psi]}$. For the almost complex structure $J : T_{[\psi]}\mathbb{P}(\mathcal{H}) \to T_{[\psi]}\mathbb{P}(\mathcal{H})$, this means that it suffices to describe $JA^*_{[\psi]}$ for all $A \in \mathfrak{u}(\mathcal{H})$. Since $D_{[\psi]}\tilde{\Phi} : T_{[\psi]}\mathbb{P}(\mathcal{H}) \to i\mathfrak{u}(\mathcal{H})$ is injective, a formula for $D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]})$ would be a satisfactory description of $JA^*_{[\psi]}$.

Proposition A.2. *Let $[\psi] \in \mathbb{P}(\mathcal{H})$ and take any normalized representative $\psi \in [\psi]$. Then*

$$D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]}) = i(A\mathbf{p}_\psi + \mathbf{p}_\psi A) - 2\langle \psi, iA\psi\rangle \mathbf{p}_\psi = i(A\mathbf{p}_\psi + \mathbf{p}_\psi A) - 2i\mathbf{p}_\psi A\mathbf{p}_\psi \tag{A.1}$$

for all $A \in \mathfrak{u}(\mathcal{H})$.

Proof. Let $\tilde{A} \in \mathfrak{X}(\mathcal{H})$ denote the fundamental vector field of the action of $\mathrm{U}(\mathcal{H})$ on $\mathcal{H}$. That is,

$$\tilde{A}_\psi = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \exp(tA)\psi = A\psi \in T_\psi\mathcal{H} = \mathcal{H}. \tag{A.2}$$

Clearly, we have $D_\psi\pi(\tilde{A}_\psi) = A^*_{[\psi]}$, where $\pi : \mathcal{H}\setminus\{0\} \to \mathbb{P}(\mathcal{H})$ is the canonical projection. The map π is holomorphic, so

$$JA^*_{[\psi]} = JD_\psi\pi\tilde{A}_\psi = D_\psi\pi J\tilde{A}_\psi = D_\psi\pi(i\tilde{A}_\psi) = D_\psi\pi(iA\psi).$$

It follows that

$$D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]}) = D_{[\psi]}\tilde{\Phi}\circ D_\psi\pi(iA\psi) = D_\psi(\tilde{\Phi}\circ\pi)(iA\psi).$$

But the composition $\tilde{\Phi}\circ\pi$ is nothing but the map $\phi \mapsto \phi\langle\phi,\cdot\rangle / \langle\phi,\phi\rangle$. Therefore

$$\begin{aligned} D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]}) &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \frac{(\psi + iA\psi t)\langle\psi + iA\psi t, \cdot\rangle}{\langle\psi + iA\psi t, \psi + iA\psi t\rangle} \\ &= iA\psi\langle\psi,\cdot\rangle + \psi\langle iA\psi,\cdot\rangle - (\langle iA\psi,\psi\rangle + \langle\psi, iA\psi\rangle)\psi\langle\psi,\cdot\rangle \\ &= i(A\psi\langle\psi,\cdot\rangle + \psi\langle\psi,\cdot\rangle A) - 2\langle\psi, iA\psi\rangle\psi\langle\psi,\cdot\rangle. \end{aligned}$$

□

Example A.1. Sanity check: take $A = -i\mathbb{1} \in \mathfrak{u}(\mathcal{H})$, then $A^*_{[\psi]} = 0$ and so we expect $D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]}) = 0$. Indeed, the right hand side of Eq. (A.1) gives $\mathbf{p}_\psi + \mathbf{p}_\psi - 2\mathbf{p}_\psi = 0$.

Corollary A.3.

$$D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]}) = i\left[D_{[\psi]}\tilde{\Phi}(A^*_{[\psi]}), \mathbf{p}_\psi\right] \tag{A.3}$$

Proof. Apply Propositions 5.16 and A.2:

$$\begin{aligned} i\left[D_{[\psi]}\tilde{\Phi}(A^*_{[\psi]}), \mathbf{p}_\psi\right] &= i[[A,\mathbf{p}_\psi],\mathbf{p}_\psi] = i[(A\mathbf{p}_\psi - \mathbf{p}_\psi A),\mathbf{p}_\psi] \\ &= i(A\mathbf{p}_\psi - \mathbf{p}_\psi A\mathbf{p}_\psi - \mathbf{p}_\psi A\mathbf{p}_\psi + \mathbf{p}_\psi A) = i(A\mathbf{p}_\psi + \mathbf{p}_\psi) - 2i\mathbf{p}_\psi A\mathbf{p}_\psi \\ &= D_{[\psi]}\tilde{\Phi}(JA^*_{[\psi]}) \end{aligned}$$

□

A.2 The Fubini-Study Metric

Now we will give $\mathbb{P}(\mathcal{H})$ a Riemannian metric. On the real vector space $i\mathfrak{u}(\mathcal{H})$, there is a natural inner product given by

$$h^{\mathrm{HS}}(X,Y) := \mathrm{Tr}(XY) \tag{A.4}$$

for $X, Y \in i\mathfrak{u}(\mathcal{H})$. Thus h^{HS} defines a Riemannian metric on $i\mathfrak{u}(\mathcal{H})$ via the identification $T_A i\mathfrak{u}(\mathcal{H}) = i\mathfrak{u}(\mathcal{H})$ for all $A \in i\mathfrak{u}(\mathcal{H})$.

Definition A.1. *The **Fubini-Study metric** h^{FS} on $\mathbb{P}(\mathcal{H})$ is the pullback of h^{HS} by $\tilde{\Phi}$. That is,*

$$h^{\mathrm{FS}}_{[\psi]}(X,Y) = \mathrm{Tr}\Big[(D_{[\psi]}\tilde{\Phi}(X)) \circ (D_{[\psi]}\tilde{\Phi}(Y))\Big] \tag{A.5}$$

for all $X, Y \in T_{[\psi]}\mathbb{P}(\mathcal{H})$.

That h^{FS} is a Riemannian metric follows from the fact that $\tilde{\Phi}$ is an embedding. The following proposition relates h^{FS} to the inner product on $\mathcal{H}$:

Proposition A.4. *Take any unit vector $\psi \in \mathcal{H}$ and $X, Y \in T_\psi\mathcal{H} = \mathcal{H}$ such that $\langle \psi, X\rangle = \langle \psi, Y\rangle = 0$. Then*

$$h^{\mathrm{FS}}(D_\psi\pi(X), D_\psi\pi(Y)) = 2\,\mathrm{Re}\,\langle X, Y\rangle\,. \tag{A.6}$$

Proof.

$$\begin{aligned} &h^{\mathrm{FS}}(D_\psi\pi(X), D_\psi\pi(Y)) = h^{\mathrm{HS}}(D_{[\psi]}\tilde{\Phi} \circ D_\psi\pi(X), D_{[\psi]}\tilde{\Phi} \circ D_\psi\pi(Y)) \\ &= h^{\mathrm{HS}}(D_\psi(\tilde{\Phi}\circ\pi)(X), D_\psi(\tilde{\Phi}\circ\pi)(Y)) \end{aligned}$$

We have

$$D_\psi(\tilde{\Phi}\circ\pi)(X) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \frac{(\psi + tX)\,\langle \psi + tX, \cdot\rangle}{\langle \psi + tX, \psi + tX\rangle} = X\,\langle \psi, \cdot\rangle + \psi\,\langle X, \cdot\rangle\,,$$

where we used $\langle \psi, \psi\rangle = 1$ and $\langle \psi, X\rangle = 0$. Therefore,

$$\begin{aligned} &h^{\mathrm{FS}}(D_\psi\pi(X), D_\psi\pi(Y)) \\ &= \mathrm{Tr}\left[(X\,\langle \psi, \cdot\rangle + \psi\,\langle X, \cdot\rangle) \circ (Y\,\langle \psi, \cdot\rangle + \psi\,\langle Y, \cdot\rangle)\right] \\ &= \mathrm{Tr}\Big[X\,\langle Y, \cdot\rangle + \langle X, Y\rangle\,\psi\,\langle \psi, \cdot\rangle\Big] \\ &= \langle Y, X\rangle + \langle X, Y\rangle\,. \end{aligned}$$

□

Example A.2. The complex projective line $\mathbb{CP}^1$ can be identified with the two-sphere S^2 by (think of the Bloch sphere)

$$\begin{aligned} &\alpha : \mathbb{CP}^1 \to S^2 \\ &\quad [z_0 : z_1] \mapsto \frac{1}{|z_0|^2 + |z_1|^2}(2\,\mathrm{Re}(\bar{z}_0 z_1), 2\,\mathrm{Im}(\bar{z}_0 z_1), |z_0|^2 - |z_1|^2). \end{aligned} \tag{A.7}$$

Under this identification, the Fubini-Study metric on $\mathbb{CP}^1$ coincides, up to a factor of $\frac{1}{2}$, with the standard round metric h° on S^2. That is, $h^{\mathrm{FS}} = \frac{1}{2}\alpha^* h^\circ$.

Proof. Consider the map

$$j : S^2 \to i\mathfrak{u}(2)$$
$$(x, y, z) \mapsto \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}.$$

Then the following diagram commutes:

$$\begin{array}{ccc} \mathbb{CP}^1 & \xrightarrow{\tilde{\Phi}} & i\mathfrak{u}(2) \\ & {\scriptstyle\alpha}\searrow \quad \nearrow{\scriptstyle j} & \\ & S^2 & \end{array} \tag{A.8}$$

It is easy to check that $h^\circ = 2j^*h^{\mathrm{HS}}$. Pulling back by α yields $\alpha^*h^\circ = 2\alpha^*j^*h^{\mathrm{HS}} = 2(j \circ \alpha)^*h^{\mathrm{HS}} = 2\tilde{\Phi}^*h^{\mathrm{HS}} = 2h^{\mathrm{FS}}$. □

A.3 The Fubini-Study Symplectic Form

We are now ready to give $\mathbb{P}(\mathcal{H})$ a symplectic structure:

Definition A.2. *The **Fubini-Study 2-form** ω^{FS} on $\mathbb{P}(\mathcal{H})$ is defined by*

$$\omega^{\mathrm{FS}}(X, Y) = h^{\mathrm{FS}}(JX, Y) \tag{A.9}$$

for all $X, Y \in T_{[\psi]}\mathbb{P}(\mathcal{H})$.

Of course, for ω^{FS} to be a symplectic form, we need to check a few things. First of all, for ω^{FS} to be a 2-form at all, it has to be skew-symmetric. This can be shown by, for example, proving that h^{FS} is invariant under J, and applying Lemma A.1. Instead, we will show this by direct computation, which has the advantage of yielding an explicit formula for $\omega^{\mathrm{FS}}(A^*_{[\psi]}, B^*_{[\psi]})$. Secondly, we need to show that ω^{FS} is closed.

Lemma A.5. *Take $A, B \in \mathfrak{u}(\mathcal{H})$ and $[\psi] \in \mathbb{P}(\mathcal{H})$. Choose a normalized representative ψ of $[\psi]$. Then*

$$\omega^{\mathrm{FS}}(A^*_{[\psi]}, B^*_{[\psi]}) = i\,\langle \psi, [A, B]\psi \rangle\,. \tag{A.10}$$

Proof.

$$\begin{aligned}
&\omega^{\mathrm{FS}}(A^*_{[\psi]}, B^*_{[\psi]}) = h^{\mathrm{FS}}(JA^*_{[\psi]}, B^*_{[\psi]}) \\
&= h^{\mathrm{HS}}\Big(D_{[\psi]}\tilde{\Phi}JA^*_{[\psi]}, D_{[\psi]}\tilde{\Phi}B^*_{[\psi]}\Big) \\
&= h^{\mathrm{HS}}\Big(i\big(A\mathbf{p}_\psi + \mathbf{p}_\psi A - 2\mathbf{p}_\psi A\mathbf{p}_\psi\big), [B, \mathbf{p}_\psi]\Big) \\
&= i\mathrm{Tr}\Big[\big(A\mathbf{p}_\psi + \mathbf{p}_\psi A - 2\mathbf{p}_\psi A\mathbf{p}_\psi\big)(B\mathbf{p}_\psi - \mathbf{p}_\psi B)\Big] \\
&= i\mathrm{Tr}\Big[(A\mathbf{p}_\psi + \mathbf{p}_\psi A)(B\mathbf{p}_\psi - \mathbf{p}_\psi B)\Big] \\
&= i\mathrm{Tr}(A\mathbf{p}_\psi B\mathbf{p}_\psi - A\mathbf{p}_\psi B + \mathbf{p}_\psi AB\mathbf{p}_\psi - \mathbf{p}_\psi A\mathbf{p}_\psi B) \\
&= i\mathrm{Tr}(\mathbf{p}_\psi[A, B]) \\
&= i\,\langle \psi, [A, B]\psi \rangle
\end{aligned}$$

□

Remark A.1. $[A, B] \in \mathfrak{u}(\mathcal{H})$ is skew-Hermitian, so $i\langle\psi, [A, B]\psi\rangle$ is real.

Equation (A.10) shows that ω^{FS} is indeed skew-symmetric, and hence a 2-form. Now we need to check that ω^{FS} is closed.

Lemma A.6. *Let $A, B, C \in \mathfrak{u}(\mathcal{H})$ and take $[\psi] \in \mathbb{P}(\mathcal{H})$ with normalized representative $\psi \in \mathcal{H}$. Then*

$$A^*_{[\psi]}\omega^{\mathrm{FS}}(B^*, C^*) = i\langle\psi, [[B, C], A]\psi\rangle\,. \tag{A.11}$$

Note: In Eq. (A.11), B^*, C^* are vector fields, so $\omega^{\mathrm{FS}}(B^*, C^*)$ is a smooth function on $\mathbb{P}(\mathcal{H})$. We have used the standard notation where Xf denotes the Lie derivative of a smooth function f along a vector field X. That is, $Xf = \mathrm{d}f(X)$. If p is a point, then $X_p f = (Xf)_p = \mathrm{d}f(X_p)$.

Proof.

$$\begin{aligned}
A^*_{[\psi]}\omega^{\mathrm{FS}}(B^*, C^*) &= A^*_{[\psi]}\Big([\psi] \mapsto i\langle\psi, [B, C]\psi\rangle\Big) \\
&= i\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}\langle\exp(tA)\psi, [B, C]\exp(tA)\psi\rangle \\
&= i\langle A\psi, [B, C]\psi\rangle + i\langle\psi, [B, C]A\psi\rangle \\
&= i\langle\psi, [[B, C], A]\psi\rangle
\end{aligned}$$

□

Proposition A.7. *The Fubini-Study 2-form ω^{FS} is closed.*

Proof. We will use the following fact: if ω is any 2-form on a manifold M and $X, Y, Z \in \mathfrak{X}(M)$ are vector fields, then

$$\begin{aligned}
&\mathrm{d}\omega(X, Y, Z) \\
&= X\omega(Y, Z) + Y\omega(Z, X) + Z\omega(X, Y) - \omega([X, Y], Z) - \omega([Y, Z], X) - \omega([Z, X], Y).
\end{aligned}$$

We apply this to $\omega = \omega^{\mathrm{FS}}$. Let $A, B, C \in \mathfrak{u}(\mathcal{H})$, then

$$\mathrm{d}\omega^{\mathrm{FS}}(A^*_{[\psi]}, B^*_{[\psi]}, C^*_{[\psi]}) = A^*_{[\psi]}\omega^{\mathrm{FS}}(B^*, C^*) + B^*_{[\psi]}\omega^{\mathrm{FS}}(C^*, A^*) + C^*_{[\psi]}\omega^{\mathrm{FS}}(A^*, B^*),$$

where the last three terms canceled due to Lemma A.5 and the Jacobi identity. Using Lemma A.6, we see that the remaining three terms also cancel due to the Jacobi identity. Since the fundamental vector fields span $T_{[\psi]}\mathbb{P}(\mathcal{H})$ at any $[\psi]$, we conclude $\mathrm{d}\omega^{\mathrm{FS}} = 0$. □

Bibliography

[1] S. R. White, "Density matrix formulation for quantum renormalization groups", Physical Review Letters **69**, 2863–2866 (1992).

[2] K. Hallberg, "New Trends in Density Matrix Renormalization", Advances in Physics **55**, 477–526 (2006).

[3] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states", Annals of physics **326**, 96–192 (2011).

[4] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, "Matrix product states and projected entangled pair states: Concepts, symmetries, theorems", Reviews of Modern Physics **93**, 045003 (2021).

[5] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, "Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions", Reviews of Modern Physics **68**, 13–125 (1996).

[6] D. Zgid and G. K.-L. Chan, "Dynamical mean-field theory from a quantum chemical perspective", The Journal of Chemical Physics **134**, 094115 (2011).

[7] D. Vollhardt, "Dynamical mean-field theory for correlated electrons", Annalen der Physik **524**, 1–19 (2012).

[8] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas", Physical Review **136**, B864–B871 (1964).

[9] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects", Physical Review **140**, A1133–A1138 (1965).

[10] K. Burke, "Perspective on density functional theory", The Journal of Chemical Physics **136**, 150901 (2012).

[11] A. D. Becke, "Perspective: Fifty years of density-functional theory in chemical physics", The Journal of Chemical Physics **140**, 18A301 (2014).

[12] P. J. Hasnip, K. Refson, M. I. J. Probert, J. R. Yates, S. J. Clark, and C. J. Pickard, "Density functional theory in the solid state", Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **372**, 20130270 (2014).

[13] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density", Physical Review. B, Condensed Matter **37**, 785–789 (1988).

[14] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange", The Journal of Chemical Physics **98**, 5648–5652 (1993).

[15] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized Gradient Approximation Made Simple", Physical Review Letters **77**, 3865–3868 (1996).

[16] J. Toulouse, "Review of Approximations for the Exchange-Correlation Energy in Density-Functional Theory", in *Density Functional Theory: Modeling, Mathematical Analysis, Computational Methods, and Applications*, edited by E. Cancès and G. Friesecke (Springer International Publishing, Cham, 2023), pp. 1–90.

[17] T. L. Gilbert, "Hohenberg-Kohn theorem for nonlocal external potentials", Physical Review B **12**, 2111–2120 (1975).

[18] R. A. Donnelly and R. G. Parr, "Elementary properties of an energy functional of the first-order reduced density matrix", The Journal of Chemical Physics **69**, 4431–4439 (1978).

[19] M. Piris, "Natural Orbital Functional Theory", in *Reduced-Density-Matrix Mechanics: With Application to Many-Electron Atoms and Molecules* (John Wiley & Sons, Ltd, 2007) Chap. 14, pp. 385–427.

[20] S. Sharma, J. K. Dewhurst, N. N. Lathiotakis, and E. K. U. Gross, "Reduced density matrix functional for many-electron systems", Physical Review B **78**, 201103 (2008).

[21] K. Pernal and K. J. H. Giesbertz, "Reduced Density Matrix Functional Theory (RDMFT) and Linear Response Time-Dependent RDMFT (TD-RDMFT)", in *Density-Functional Methods for Excited States*, Vol. 368, edited by N. Ferré, M. Filatov, and M. Huix-Rotllant (Springer International Publishing, Cham, 2015), pp. 125–183.

[22] N. N. Lathiotakis, S. Sharma, N. Helbig, J. K. Dewhurst, M. a. L. Marques, F. Eich, T. Baldsiefen, A. Zacarias, and E. K. U. Gross, "Discontinuities of the Chemical Potential in Reduced Density Matrix Functional Theory", Zeitschrift für Physikalische Chemie **224**, 467–480 (2010).

[23] R. López-Sandoval and G. M. Pastor, "Density-matrix functional theory of the Hubbard model: An exact numerical study", Physical Review B **61**, 1764–1772 (2000).

[24] A. J. Cohen and P. Mori-Sánchez, "Landscape of an exact energy functional", Physical Review A **93**, 042511 (2016).

[25] C. L. Benavides-Riveros, J. Wolff, M. A. L. Marques, and C. Schilling, "Reduced Density Matrix Functional Theory for Bosons", Physical Review Letters **124**, 180603 (2020).

[26] S. Goedecker and C. J. Umrigar, "Natural Orbital Functional for the Many-Electron Problem", Physical Review Letters **81**, 866–869 (1998).

[27] M. A. BUIJSE and E. J. BAERENDS, "An approximate exchange-correlation hole density as a functional of the natural orbitals", Molecular Physics **100**, 401–421 (2002).

[28] O. Gritsenko, K. Pernal, and E. J. Baerends, "An improved density matrix functional by physically motivated repulsive corrections", The Journal of Chemical Physics **122**, 204102 (2005).

[29] M. A. L. Marques and N. N. Lathiotakis, "Empirical functionals for reduced-density-matrix-functional theory", Physical Review A **77**, 032509 (2008).

[30] P. Leiva and M. Piris, "Assessment of a new approach for the two-electron cumulant in natural-orbital-functional theory", The Journal of Chemical Physics **123**, 214102 (2005).

[31] M. Piris, "A new approach for the two-electron cumulant in natural orbital functional theory", International Journal of Quantum Chemistry **106**, 1093–1104 (2006).

[32] M. Piris, X. Lopez, F. Ruipérez, J. M. Matxain, and J. M. Ugalde, "A natural orbital functional for multiconfigurational states", The Journal of Chemical Physics **134**, 164102 (2011).

[33] M. Piris, "Global Method for Electron Correlation", Physical Review Letters **119**, 063002 (2017).

[34] M. Piris, "Global Natural Orbital Functional: Towards the Complete Description of the Electron Correlation", Physical Review Letters **127**, 233001 (2021).

[35] M. Piris, *Advances in Approximate Natural Orbital Functionals: From Historical Perspectives to Contemporary Developments*, Dec. 2023.

[36] A. M. K. Müller, "Explicit approximate relation between reduced two- and one-particle density matrices", Physics Letters A **105**, 446–452 (1984).

[37] C. Schilling and R. Schilling, "Diverging Exchange Force and Form of the Exact Density Matrix Functional", Physical Review Letters **122**, 013001 (2019).

[38] J. Liebert and C. Schilling, "Functional theory for Bose-Einstein condensates", Physical Review Research **3**, 013282 (2021).

[39] T. Maciążek, "Repulsively diverging gradient of the density functional in the reduced density matrix functional theory", New Journal of Physics **23**, 113006 (2021).

[40] J. Cioslowski and K. Pernal, "Unoccupied natural orbitals in two-electron Coulombic systems", Chemical Physics Letters **430**, 188–190 (2006).

[41] C. L. Benavides-Riveros, J. M. Gracia-Bondía, and M. Springborg, "Quasipinning and entanglement in the lithium isoelectronic series", Physical Review A **88**, 022508 (2013).

[42] C. Schilling, "Quasipinning and its relevance for N-fermion quantum states", Physical Review A **91**, 022105 (2015).

[43] F. Tennie, D. Ebler, V. Vedral, and C. Schilling, "Pinning of fermionic occupation numbers: General concepts and one spatial dimension", Physical Review A **93**, 042126 (2016).

[44] F. Tennie, V. Vedral, and C. Schilling, "Influence of the fermionic exchange symmetry beyond Pauli's exclusion principle", Physical Review A **95**, 022336 (2017).

[45] C. Schilling, C. L. Benavides-Riveros, A. Lopes, T. Maciążek, and A. Sawicki, "Implications of pinned occupation numbers for natural orbital expansions: I. Generalizing the concept of active spaces", New Journal of Physics **22**, 023001 (2020).

[46] T. Maciążek, A. Sawicki, D. Gross, A. Lopes, and C. Schilling, "Implications of pinned occupation numbers for natural orbital expansions. II: rigorous derivation and extension to non-fermionic systems", New Journal of Physics **22**, 023002 (2020).

[47] R. Van Leeuwen, "Density Functional Approach to the Many-Body Problem: Key Concepts and Exact Functionals", in *Advances in Quantum Chemistry*, Vol. 43 (Elsevier, 2003), pp. 25–94.

[48] M. Lewin, E. H. Lieb, and R. Seiringer, "Universal Functionals in Density Functional Theory", in (2023), pp. 115–182.

[49] N. D. Mermin, "Thermal Properties of the Inhomogeneous Electron Gas", Physical Review **137**, A1441–A1443 (1965).

[50] E. Runge and E. K. U. Gross, "Density-Functional Theory for Time-Dependent Systems", Physical Review Letters **52**, 997–1000 (1984).

[51] M. Levy, "Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem", Proceedings of the National Academy of Sciences **76**, 6062–6065 (1979).

[52] E. H. Lieb, "Density functionals for coulomb systems", International Journal of Quantum Chemistry **24**, 243–277 (1983).

[53] S. M. Valone, "Consequences of extending 1-matrix energy functionals from pure–state representable to all ensemble representable 1 matrices", The Journal of Chemical Physics **73**, 1344–1349 (1980).

[54] P.-O. Löwdin, "Quantum Theory of Many-Particle Systems. I. Physical Interpretations by Means of Density Matrices, Natural Spin-Orbitals, and Convergence Problems in the Method of Configurational Interaction", Physical Review **97**, 1474–1489 (1955).

[55] L. H. Thomas, "The calculation of atomic fields", Mathematical Proceedings of the Cambridge Philosophical Society **23**, 542–548 (1927).

[56] E. Fermi, "Un Metodo Statistico per la Determinazione di alcune Prioprietà dell'Atomo", **6**, 602–607 (1927).

[57] U. V. Barth and L Hedin, "A local exchange-correlation potential for the spin polarized case. i", Journal of Physics C: Solid State Physics **5**, 1629–1642 (1972).

[58] C. A. Ullrich, "Nonuniqueness in spin-density-functional theory on lattices", Physical Review B **72**, 073102 (2005).

[59] L. N. Oliveira, E. K. U. Gross, and W. Kohn, "Density-Functional Theory for Superconductors", Physical Review Letters **60**, 2430–2433 (1988).

[60] M. Lüders, M. A. L. Marques, N. N. Lathiotakis, A. Floris, G. Profeta, L. Fast, A. Continenza, S. Massidda, and E. K. U. Gross, "Ab initio theory of superconductivity. I. Density functional formalism and approximate functionals", Physical Review B **72**, 024545 (2005).

[61] M. A. L. Marques, M. Lüders, N. N. Lathiotakis, G. Profeta, A. Floris, L. Fast, A. Continenza, E. K. U. Gross, and S. Massidda, "*Ab initio* theory of superconductivity. II. Application to elemental metals", Physical Review B **72**, 024546 (2005).

[62] M. Higuchi and K. Higuchi, "Arbitrary choice of basic variables in density functional theory: Formalism", Physical Review B **69**, 035113 (2004).

[63] V. L. Líbero and K. Capelle, "Spin-distribution functionals and correlation energy of the Heisenberg model", Physical Review B **68**, 024423 (2003).

[64] F. C. Alcaraz and K. Capelle, "Density functional formulations for quantum chains", Physical Review B **76**, 035109 (2007).

[65] S. Sachdev, *Quantum Phase Transitions*, 2nd ed. (Cambridge University Press, Cambridge, 2011).

[66] M. Vojta, "Quantum phase transitions", Reports on Progress in Physics **66**, 2069 (2003).

[67] L.-A. Wu, M. S. Sarandy, D. A. Lidar, and L. J. Sham, "Linking entanglement and quantum phase transitions via density-functional theory", Physical Review A **74**, 052335 (2006).

[68] Á. Nagy, M. Calixto, and E. Romera, "A Density Functional Theory View of Quantum Phase Transitions", Journal of Chemical Theory and Computation **9**, 1068–1072 (2013).

[69] Á. Nagy and E. Romera, "Quantum phase transitions via density-functional theory: Extension to the degenerate case", Physical Review A **88**, 042515 (2013).

[70] S.-J. Gu, "Density-Functional Fidelity Approach to Quantum Phase Transitions", Chinese Physics Letters **26**, 026401 (2009).

[71] M. Penz and R. van Leeuwen, "Density-functional theory on graphs", The Journal of Chemical Physics **155**, 244111 (2021).

[72] J. T. Chayes, L. Chayes, and M. B. Ruskai, "Density functional approach to quantum lattice systems", Journal of Statistical Physics **38**, 497–518 (1985).

[73] J. Liebert, A. Y. Chaou, and C. Schilling, "Refining and relating fundamentals of functional theory", The Journal of Chemical Physics **158**, 214108 (2023).

[74] M. Penz, E. I. Tellgren, M. A. Csirik, M. Ruggenthaler, and A. Laestadius, "The Structure of Density-Potential Mapping. Part I: Standard Density-Functional Theory", ACS Physical Chemistry Au **3**, 334–347 (2023).

[75] M. Penz and R. van Leeuwen, "Geometry of Degeneracy in Potential and Density Space", Quantum **7**, 918 (2023).

[76] M. Penz and R. van Leeuwen, "Geometrical perspective on spin–lattice density-functional theory", The Journal of Chemical Physics **161**, 150901 (2024).

[77] Y.-K. Liu, M. Christandl, and F. Verstraete, "Quantum Computational Complexity of the N-Representability Problem: QMA Complete", Physical Review Letters **98**, 110503 (2007).

[78] R. Pino, O. Bokanowski, E. V. Ludeña, and R. L. Boada, "A re-statement of the Hohenberg–Kohn theorem and its extension to finite subspaces", Theoretical Chemistry Accounts **118**, 557–561 (2007).

[79] L. Garrigue, "Unique continuation for many-body Schrödinger operators and the Hohenberg-Kohn theorem", Mathematical Physics, Analysis and Geometry **21**, 27 (2018).

[80] J. Katriel, C. J. Appellof, and E. R. Davidson, "Mapping between local potentials and ground state densities", International Journal of Quantum Chemistry **19**, 293–300 (1981).

[81] W. Kohn, "V-Representability and Density Functional Theory", Physical Review Letters **51**, 1596–1598 (1983).

[82] C. A. Ullrich and W. Kohn, "Degeneracy in Density Functional Theory: Topology in the v and n Spaces", Physical Review Letters **89**, 156401 (2002).

[83] C. Schilling, "Communication: Relating the pure and ensemble density matrix functional", The Journal of Chemical Physics **149**, 231102 (2018).

[84] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis* (Birkhäuser, Boston, 2009).

[85] M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman, and E. A. Cornell, "Observation of Bose-Einstein Condensation in a Dilute Atomic Vapor", Science **269**, 198–201 (1995).

[86] C. C. Bradley, C. A. Sackett, J. J. Tollett, and R. G. Hulet, "Evidence of Bose-Einstein Condensation in an Atomic Gas with Attractive Interactions", Physical Review Letters **75**, 1687–1690 (1995).

[87] K. B. Davis, M. O. Mewes, M. R. Andrews, N. J. van Druten, D. S. Durfee, D. M. Kurn, and W. Ketterle, "Bose-Einstein Condensation in a Gas of Sodium Atoms", Physical Review Letters **75**, 3969–3973 (1995).

[88] I. Bloch, "Ultracold quantum gases in optical lattices", Nature Physics **1**, 23–30 (2005).

[89] I. Bloch, J. Dalibard, and W. Zwerger, "Many-body physics with ultracold gases", Reviews of Modern Physics **80**, 885–964 (2008).

[90] I. Bloch, J. Dalibard, and S. Nascimbène, "Quantum simulations with ultracold quantum gases", Nature Physics **8**, 267–276 (2012).

[91] C. Gross and I. Bloch, "Quantum simulations with ultracold atoms in optical lattices", Science **357**, 995–1001 (2017).

[92] F. Schäfer, T. Fukuhara, S. Sugawa, Y. Takasu, and Y. Takahashi, "Tools for quantum simulation with ultracold atoms in optical lattices", Nature Reviews Physics **2**, 411–425 (2020).

[93] D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, and P. Zoller, "Cold Bosonic Atoms in Optical Lattices", Physical Review Letters **81**, 3108–3111 (1998).

[94] D. Jaksch and P. Zoller, "The cold atom Hubbard toolbox", Annals of Physics **315**, 52–79 (2005).

[95] Y. Takahashi, "Quantum simulation of quantum many-body systems with ultracold two-electron atoms in an optical lattice", Proceedings of the Japan Academy, Series B **98**, 141–160 (2022).

[96] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, "Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms", Nature **415**, 39–44 (2002).

[97] I. Bloch and M. Greiner, "The superfluid-to-Mott insulator transition and the birth of experimental quantum simulation", Nature Reviews Physics **4**, 739–740 (2022).

[98] O. Penrose and L. Onsager, "Bose-Einstein Condensation and Liquid Helium", Physical Review **104**, 576–584 (1956).

[99] J. Schmidt, M. Fadel, and C. L. Benavides-Riveros, "Machine learning universal bosonic functionals", Physical Review Research **3**, L032063 (2021).

[100] C. Schilling, "Hubbard model: Pinning of occupation numbers and role of symmetries", Physical Review B **92**, 155149 (2015).

[101] J. Cioslowski, O. M. Brown, and T. Maciazek, "Natural orbitals and their occupation numbers for free anyons in the magnetic gauge", Physical Review A **109**, 023316 (2024).

[102] J. M. Danskin, "The Theory of Max-Min, with Applications", SIAM Journal on Applied Mathematics **14**, 641–664 (1966).

[103] M. F. Atiyah, "Convexity and Commuting Hamiltonians", Bulletin of the London Mathematical Society **14**, 1–15 (1982).

[104] V. Guillemin and S. Sternberg, "Convexity properties of the moment mapping", Inventiones Mathematicae **67**, 491–513 (1982).

[105] M. Walter, "Multipartite Quantum States and their Marginals", PhD thesis (2014), 1 Band.

[106] E. Michael, "Continuous Selections. I", Annals of Mathematics **63**, 361–382 (1956).

[107] C. Schilling, *The Quantum Marginal Problem*, Apr. 2014.

[108] C. Schilling, "Quantum Marginal Problem and its Physical Relevance", PhD thesis (2014), 1 Band.

[109] A. J. Coleman, "Structure of Fermion Density Matrices", Reviews of Modern Physics **35**, 668–686 (1963).

[110] R. E. Borland and K. Dennis, "The conditions on the one-matrix for three-body fermion wavefunctions with one-rank equal to six", Journal of Physics B: Atomic and Molecular Physics **5**, 7 (1972).

[111] A. Higuchi, A. Sudbery, and J. Szulc, "One-qubit reduced states of a pure many-qubit state: polygon inequalities", Physical Review Letters **90**, 107902 (2003).

[112] S. Bravyi, *Requirements for compatibility between local and multipartite quantum states*, Jan. 2003.

[113] A. Klyachko, *Quantum marginal problem and representations of the symmetric group*, Sept. 2004.

[114] A. Knutson, *The symplectic and algebraic geometry of Horn's problem*, May 2000.

[115] S. Daftuar and P. Hayden, "Quantum state transformations and the Schubert calculus", Annals of Physics **315**, 80–122 (2005).

[116] A. A. Klyachko, "Quantum marginal problem and N-representability", Journal of Physics: Conference Series **36**, 72 (2006).

[117] M. Christandl and G. Mitchison, "The Spectra of Density Operators and the Kronecker Coefficients of the Symmetric Group", Communications in Mathematical Physics **261**, 789–797 (2006).

[118] M. Altunbulak, "The Pauli Principle, Representation Theory, and Geometry of Flag Varieties", PhD thesis (Bilkent University, July 2008).

[119] M. Altunbulak and A. Klyachko, "The Pauli Principle Revisited", Communications in Mathematical Physics **282**, 287–322 (2008).

[120] A. Berenstein and R. Sjamaar, *Coadjoint orbits, moment polytopes, and the Hilbert-Mumford criterion*, Nov. 1999.

[121] G. J. Heckman, "Projections of orbits and asymptotic behavior of multiplicities for compact connected Lie groups", Inventiones mathematicae **67**, 333–356 (1982).

[122] V. Guillemin and S. Sternberg, "Convexity properties of the moment mapping. II", Inventiones mathematicae **77**, 533–546 (1984).

[123] M. F. Atiyah and R. Bott, "The moment map and equivariant cohomology", Topology **23**, 1–28 (1984).

[124] F. Kirwan, "Convexity properties of the moment mapping, III", Inventiones mathematicae **77**, 547–552 (1984).

[125] R. Sjamaar, "Convexity Properties of the Moment Mapping Re-examined", Advances in Mathematics **138**, 46–91 (1998).

[126] A. Sawicki, A. Huckleberry, and M. Kuś, "Symplectic geometry of entanglement", Communications in Mathematical Physics **305**, 441–468 (2011).

[127] M. Walter, B. Doran, D. Gross, and M. Christandl, "Entanglement Polytopes: Multiparticle Entanglement from Single-Particle Information", Science **340**, 1205–1208 (2013).

[128] A. Sawicki, M. Oszmaniec, and M. Kuś, "Convexity of momentum map, Morse index, and quantum entanglement", Reviews in Mathematical Physics **26**, 1450004 (2014).

[129] D. J. Simms and N. M. J. Woodhouse, *Lectures on Geometric Quantization*, Vol. 53, Lecture Notes in Physics (Springer, Berlin, Heidelberg, 1976).

[130] V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics* (Cambridge University Press, May 1990).

[131] V. Guillemin, *Convexity properties of Hamiltonian group actions*, CRM Monograph Series v. 26 (American Mathematical Society, Providence, R.I, 2005).

[132] M. Walter, D. Gross, and J. Eisert, *Multi-partite entanglement*, Feb. 2017.

[133] J. M. Lee, *Introduction to Smooth Manifolds*, Vol. 218, Graduate Texts in Mathematics (Springer, New York, NY, 2012).

[134] L. W. Tu, *Differential Geometry*, Vol. 275, Graduate Texts in Mathematics (Springer International Publishing, Cham, 2017).

[135] J. M. Lee, *Introduction to complex manifolds*, Graduate Studies in Mathematics volume 244 (American Mathematical Society, Providence, Rhode Island, 2024).

[136] J.-P. Ortega and T. S. Ratiu, *Momentum Maps and Hamiltonian Reduction* (Birkhäuser, Boston, MA, 2004).

[137] A. C. Da Silva, *Lectures on Symplectic Geometry*, Vol. 1764, Lecture Notes in Mathematics (Springer, Berlin, Heidelberg, 2008).

[138] J. E. Humphreys, *Introduction to Lie Algebras and Representation Theory*, Vol. 9, Graduate Texts in Mathematics (Springer, New York, NY, 1972).

[139] W. Fulton and J. Harris, *Representation Theory*, Vol. 129, Graduate Texts in Mathematics (Springer New York, New York, NY, 2004).

[140] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Vol. 222, Graduate Texts in Mathematics (Springer International Publishing, Cham, 2015).

[141] A. A. Klyachko, *The Pauli exclusion principle and beyond*, Apr. 2009.

[142] C. Schilling, C. L. Benavides-Riveros, and P. Vrana, "Reconstructing quantum states from single-party information", Physical Review A **96**, 052312 (2017).

Acknowledgements

This exciting project would not have been possible without the countless stimulating ideas from Christian, my thesis advisor, and fruitful discussions with him throughout the year. His everlasting enthusiasm for physics and sharp intuition have helped me overcome numerous scientific obstacles.

I would also like to express my gratitude to my amazing colleagues: Julia, Damiano, , Markus, Ludvík, Unik, and Paul. I feel lucky to be around such brilliant people.

The two-year academic journey in TMP would not have been complete without my friends – Omkar, Thomas, Miquel, , Afonso, András, Dmytro, Vittorio, Matteo, Martin, and many more. I thank them for their company and friendship.

Finally, I am deeply indebted to Laura for her constant and unfailing support and encouragement, and to my parents for their unconditional love and unwavering belief in my education.

Erklärung:

Hiermit erkläre ich, die vorliegende Arbeit selbständig verfasst zu haben und keine anderen als die in der Arbeit angegebenen Quellen und Hilfsmittel benutzt zu haben.

München, 5.10.2025

Chih-Chun Wang