Latin American Mathematics Series
UFSCar subseries

Jean-Bernard Bru
Walter Alberto de Siqueira Pedra

C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics

An Introduction

ufscar

Springer

Latin American Mathematics Series

Latin American Mathematics Series – UFSCar subseries

Series Editors

César R. de Oliveira, Center of Exact Sciences and Technology, Federal University of São Carlos, São Carlos, São Paulo, Brazil

Ruy Tojeiro, ICMC, Universidade de São Paulo, São Carlos, São Paulo, Brazil

Advisory Editors

Shiferaw Berhanu, Department of Mathematics, Temple University, Philadelphia, PA, USA

Ugo Bruzzo, SISSA - Scuola Internazionale Superiore di Studi Avanzati, Trieste, Italy

Irene Fonseca, Department of Math Sci, Carnegie Mellon University, Pittsburgh, PA, USA

Published under the Latin American Mathematics Series, which was created to showcase the new, vibrant mathematical output that is emerging from this region, this subseries aims to gather high-quality monographs, graduate textbooks, and contributing volumes based on mathematical research conducted at/with the Federal University of São Carlos, a technological pole located in the State of São Paulo, Brazil. Submissions are evaluated by an international editorial board and undergo a rigorous peer review before acceptance.

Jean-Bernard Bru •
Walter Alberto de Siqueira Pedra

C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics

An Introduction

Springer

Jean-Bernard Bru
Departamento de Matemáticas & EHU Quantum center, Facultad de Ciencia y Tecnología
Universidad del País Vasco / Euskal Herriko Unibertsitatea, UPV/EHU
Bilbao, Spain

BCAM - Basque Center for Applied Mathematics
Bilbao, Spain

IKERBASQUE, Basque Foundation for Science
Bilbao, Spain

Walter Alberto de Siqueira Pedra
Instituto de Ciências Matemáticas e da Computação – Departamento de Matemática
Universidade de São Carlos
São Carlos, Brazil

ISSN 2524-6755 ISSN 2524-6763 (electronic)
Latin American Mathematics Series
ISSN 2524-6755 ISSN 2524-6763 (electronic)
Latin American Mathematics Series – UFSCar subseries
ISBN 978-3-031-28948-4 ISBN 978-3-031-28949-1 (eBook)
https://doi.org/10.1007/978-3-031-28949-1

Mathematics Subject Classification: 82B10, 81V74, 46L05, 81R15, 46-01

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Dedicated to our families, teachers, and students

Preface

This book grew from lecture notes we have written for participants of lectures on applications of C^*-algebra theory to the foundations of quantum statistical mechanics, as well as a mini-course on thermodynamic equilibrium of quantum lattice systems with mean-field interactions, we held at the Institute of Physics of the University of São Paulo and at the Basque Center for Applied Mathematics (BCAM), in the last few years. In both cases, the audience was rather heterogeneous, composed by students at graduate and undergraduate level, from the Institute of Physics, the Institute of Mathematics and Statistics of University of São Paulo, and the BCAM. Most participants from the Institute of Physics had only very modest previous knowledge on fundamental mathematical disciplines like analysis, topology, and functional analysis. Thus, it was necessary to provide friendly and self-contained material, in order to allow them to follow the main ideas presented. In this sense, one important feature of our book is a good compromise between conceptual depth and technical simplicity. In fact, our book is mainly addressed to students stemming from physics departments, who are interested in mathematical foundations of physics and aim at studying physical theories in a mathematically rigorous way. From the point of view of a graduate student in mathematics, a considerable part of the material presented here is rather elementary, but, in contrast to many textbooks in mathematics, we systematically discuss the physical significance of each single (abstract) mathematical structure we use. Thus, mathematicians interested in quantum statistical physics can use our book as a quick introduction to the subject, written in a language that they easily understand, in particular specialists from the domains of C^*-algebras and convex analysis. Based on our experience as supervisors of master and PhD theses of students from physics departments, there is a lack of a textbook on the subject addressed to students that we refer to above. There are indeed excellent books available (for instance, by Bratteli and Robinson, Israel, Simon, and others), but they require close acquaintance with different mathematical disciplines, what is very frequently not the case for those students. We thus aim, among other things, at bringing physics students to a sufficient level to fruitfully read those (nowadays) classical books. In other words, the present work has a propaedeutic character. Notice, additionally, that these classical books are aged of about 40 years and do not cover (at least not in a systematic way) cases that are nowadays of great relevance in research, like interacting fermions and mean-field models. In fact, the "working case" of previous books are the quantum spin lattices. Our book will close this gap by presenting interacting fermions and mean-field models on the same theoretical ground as the quantum spin lattices of previous works. However, we do not discuss bosonic systems, for they are, from a technical point of view, quite peculiar and do not fit naturally in our setting. Another technical novelty of our exposition is that we present the C^*-algebras for fermions and quantum spins in the context of universal C^*-algebras of polynomial relations. This allows us, in particular, to construct all important algebra ($*$-)automorphisms, related to physical symmetries of the systems under consideration (like space translations, gauge and parity transformations, Bogoliubov automorphisms, etc.), in a technically simple and conceptually transparent way.

Bilbao, Spain — Jean-Bernard Bru
São Carlos, Brazil — Walter Alberto de Siqueira Pedra

Acknowledgments

W. de Siqueira Pedra is supported by CNPq (309723/2020-5). J.-B. Bru is supported by the Basque Government through the grant IT1615-22 and the BERC 2022–2025 program, by the COST Action CA18232 financed by the European Cooperation in Science and Technology (COST), by the BCAM Severo Ochoa accreditation SEV-2017-0718, and the grant PID2020-112948GB-I00 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF A way of making Europe."

About the Book

Our book is organized as follows: Chap. 1 is a simple and pragmatic review of the theory of ordered vector spaces. Among other things, we present the notion of positive linear transformations and prove various important properties of these transformations. In fact, C^*-algebras are examples of ordered vector spaces and various important results on these algebras, from both mathematical and physical point of view, are simple consequences of general results referring to abstract positive linear transformations. In our opinion, having this in mind from the beginning makes various aspects and results of the theory of C^*-algebras, as well as its applications to the mathematical foundations of quantum theory, conceptually much clearer. In particular, we discuss the crucial relationship between positivity and continuity of linear transformations on ordered vector spaces endowed with norms that are compatible with the order in a natural sense. In fact, the very mathematical notion of a physical state is derived from that of positivity of a linear functional. Chapter 1 is complemented by an appendix on Riesz spaces, which certainly represent the most well-known and studied class of ordered vector spaces. In fact, by a well-known theorem of Sherman's, the selfadjoint part of a given C^*-algebra is a Riesz-space if, and only if, the algebra is commutative. Observe that commutative C^*-algebras refer to classical physical systems, whereas the quantum systems are related to the non-commutative case. However, some important properties of Riesz spaces have analogies on the side of general C^*-algebras and shed some light on important aspects of the theory of C^*-algebras and its applications to quantum physics.

In Chap. 2 we show that the space of all bounded operators on a Hilbert space is naturally an ordered vector space. In fact, the scalar product of the Hilbert space naturally yields an order relation for bounded operators. This permits to define already at this point, for this particular example of C^*-algebra, the precise mathematical notion of state. Using the results from Chap. 1 and basic theory of Hilbert spaces, we prove, in this particular and technically very simple scope, various important properties of C^*-algebras and their states so that the reader can get some acquaintance to them. These results are then extended in Chap. 4 to general unital C^*-algebras. For completeness, we also discuss in Chap. 4 the main basic results concerning non-unital C^*-algebras. In fact, the non-unital case is also relevant in theoretical physics, but not in the examples of applications contained in this book. We provide a self-contained appendix on Hilbert space theory with proofs of all results needed in Chap. 2. At the end of Chap. 2 we show that the notion of density matrices (i.e., trace-one selfadjoint matrices, whose eigenvalues are non-negative), to which physics students are more familiar, is equivalent to the notion of states as positive linear functionals on the algebra of all complex matrices of a given finite dimension. In fact, we show that the usual trace of a matrix naturally defines a bijective and bipositive transformation from the selfadjoint matrices to the hermitian linear functionals on the matrices, which identifies the set of density matrices with the set of positive normalized functional on the matrices.

This fact serves as the motivation for the (C^*-)algebraic approach to quantum statistical physics: In Chap. 3 we prove that the Gibbs density matrix associated with any selfadjoint matrix (quantum Hamiltonian) is uniquely characterized by different static (free energy minimization, tangency to the pressure functions) and dynamic properties (KMS, passivity) of the corresponding state (positive functional). Later on, in Chap. 5, once more general C^*-algebras are studied in sufficiently many details, we will use the minimizing and tangencial properties of Gibbs density matrices to define thermodynamic equilibrium of infinitely extended quantum systems. In fact, one of the main objectives of the book is to show, in a mathematically rigorous way, that for very general models for fermions and quantum spins in crystals (discrete lattices with some notion of translation invariance), including models that may contain both short-range and mean-field interaction terms, this definition of equilibrium leads to a microscopic quantum theory of first-order phase transitions naturally capturing various important phenomenological aspects of the physical system. In particular, we show that for C^*-algebras related to infinite quantum systems, in contrast to the case matrix algebras (of finite systems), first-order phase transitions do occur, that is, thermodynamic equilibrium does not imply uniqueness of the system's state. Our proof of existence of phase transitions is an adaptation of the well-known one contributed by Robert Israel, based on a classical result of convex analysis, the Bishop-Phelps theorem. In this context we also discuss the spontaneous symmetry breaking and prove its existence by invoking another classical result of convex analysis, Choquet's theorem.

Chapter 4 is devoted to the abstract theory of C^*-algebra. As in the case of bounded operators in Hilbert spaces, we show that C^*-algebras are canonically ordered vector spaces, and that various important properties of these algebras are consequences of this fact. A small account on important results concerning the commutative case is also given. Recall that this case refers to classical physics. In fact, the C^*-algebraic approach allows us to make mathematically more transparent the essential differences between the classical and quantum realms, but also to construct fruitful analogies between both. After a complete exposition on the elementary theory of C^*-algebras, we study various important properties of the positive normalized linear functionals on these algebras (i.e., their states). In the last part of this chapter we give a small introduction to the so-called universal C^*-algebras of polynomial relations and define C^*-algebras that are important in quantum statistical physics, like those related to quantum spins and fermions on the lattice, in this setting. Among other things, the universal property of these algebras allows us to introduce in a very simple way various important physical symmetries (like translation and gauge invariance).

In Chap. 5, we show that, in the C^*-algebraic approach to quantum statistical mechanics, the variational principle for Gibbs states, discussed in Chap. 2 in the (very) simple case of matrix algebras (i.e., for finite quantum systems), has a natural extension for a large class of infinitely extended quantum systems that are homogeneous in space. The physical interactions considered in this chapter are short-range, in a sense. In particular, mean-field interaction terms cannot be handled at this point. However, as mentioned above, crucial phenomenological aspects of thermodynamic equilibrium of infinite systems, like existence of first-order phase transitions and spontaneous symmetry breaking, can be derived in this setting as simple corollaries of classical (deep) results of convex analysis. In this chapter we additionally contribute a brief introduction to the mathematical foundations of the Hartree-Fock theory (and its variants), which is very popular in theoretical and computational physics, as well as in quantum chemistry.

In Chap. 6 we show how the variational principle for equilibrium states of infinite space homogeneous quantum systems of the previous chapter can be extended to include mean-field models (like, for instance, the BCS model). In fact, the construction we present refers to the exactness of Bogoliubov's approximation, a classical problem in quantum statistical physics since more than half a century, which we prove for equilibrium states of lattice quantum systems. As important physical applications of our general approach, we give (i) a rigorous proof that the equilibrium states of the BCS model spontaneously break the gauge symmetry (in particular, the model presents a first-order phase transition) at low temperatures, and (ii) we show the exactness of Bogoliubov's approximation for the corresponding Hartree-Fock variational principle, for both pressure and equilibrium states. As far as we know, ours is the only textbook handling the theory of equilibrium states of quantum mean-field models in a complete and general manner.

For convenience and sake of completeness, we gather in the appendix various well-known mathematical results (basic and specific) that are important for an easy understanding of the main part of the text, or which are relevant complements to it.

Finally, all along the text we propose various exercises. They are of three types: (i) simple exercises that only aim at stressing important, albeit basic, definitions and primary results, (ii) parts of proofs of important results, and (iii) important corollaries of the results proven in the book. In fact, in our opinion, solving exercises of type (ii) and (iii) is very important for the beginner to properly understand the nature and scope of the different (important) results that are presented.

Notation

Metric spaces

- $M \equiv (M, d)$ denotes a generic metric space and d its metric. In the special case of a metric vector space we use the notation (V, d).
- K denotes a compact metric space, or a generic compact set in a metric space.
- Given a metric space M, $B_R(p)$ is the ball of radius $R > 0$ centered at $p \in M$. Its closure is denoted by $\overline{B}_R(p)$.
- For any subset Ω of a metric space M, Ω° denotes its interior, while $\overline{\Omega}$ its closure. See Definition 7.101.
- A general net (Definition 1.15) in an arbitrary set M is denoted by $(p_i)_{i \in I}$, whereas sequences are denoted by $(v_n)_{n \in \mathbb{N}}$. If M is a metric space and the net converges, we write $p_i \to p \doteq \lim_I p_i$, and $\lim_{n \to \infty} p_n$ for sequences.
- $d_{\|\cdot\|}$ is the canonical metric of a normed space $(X, \|\cdot\|)$. See Exercise 7.92.
- d_{H} stands for the Hausdorff distance. See Definition 7.94.

Vector spaces

- The notation $\mathbb{K} = \mathbb{R}, \mathbb{C}$ means that the considered field may be both the one of real numbers $\mathbb{R}$, or of complex numbers $\mathbb{C}$.
- Generic vector spaces are denoted by $V \equiv (V, +, \cdot)$ and their generic elements, by $v, v', w \in V$.
- $\mathrm{span}(M)$ denotes the linear span of a nonempty subset M in a vector space.
- $\mathrm{co}(M)$ denotes the convex hull of a nonempty subset M of a vector space. If the vector space is endowed with a metric, then $\overline{\mathrm{co}}(M)$ denotes the *closed* convex hull of M. See Definition 7.326.
- $\mathcal{E}(M) \subseteq M$ denotes the (possibly empty) set of all extreme points of a subset M of a vector space. See Definition 7.333.
- For vector spaces $V_1, \ldots, V_n$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $n \in \mathbb{N}$, $V_1 \otimes \cdots \otimes V_n$ denotes the corresponding algebraic tensor product. See Definition 7.9.
- The complex conjugation of a $*$-vector space V is denoted by $(\cdot)^*$.
- $\mathrm{Re}\{V\} \doteq \{v \in V : v^* = v\}$ is the real vector subspace of self-conjugate elements of a $*$-vector space V. See Definition 1.6.

- A generic preordered vector space is denoted by $V \equiv (V, \succeq)$, $\succeq$ being the preorder of V.
- V^+ stands for the positive cone of $(V, \succeq)$.
- $[v, v']$ is the "interval" between any v, v' in a preordered vector space V. See Definition 1.4.
- If the net $(v_i)_{i \in I}$ is order-convergent in a preordered vector space V, we use the notation $v_i \overset{o}{\to} v$. See Definition 7.284.
- If $(v_i)_{i \in I}$ is monotonically increasing and order-convergent to v in some preordered vector space V, we use the notation $v_i \uparrow v$. For order-converging decreasing nets $(v_i)_{i \in I}$, we write $v_i \downarrow v$.
- Ω^{d} denotes the order-disjoint complement of any nonempty subset Ω in a Riesz space. See Definition 7.278.
- A generic norm of a vector space V is denoted by $\|\cdot\|$. Sometimes we also use the notation $\|\cdot\|_V$ to be more explicit. Normed vector spaces are denoted by X and their elements, by $x, x', y \in X$.
- $\|\cdot\|_{\mathrm{e}}$ is the Euclidean norm in $\mathbb{R}^D$, for any $D \in \mathbb{N}$. See Definition 7.32.
- $\|\cdot\|_u$ denotes the seminorm associate with an order unit $u \in V$ of a real preordered vector space. See Definition 1.41.
- For any subset Ω of a normed space X, $\Omega^\perp$ denotes the largest subset of X that is orthogonal to Ω. See Sect. 7.3.1. In the special case of Hilbert spaces, it corresponds to the usual orthogonal complement, thanks to Lemma 7.206.
- Given Hilbert spaces $H_1, \ldots, H_n$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $n \in \mathbb{N}$, $H_1 \otimes_2 \cdots \otimes_2 H_n$ denotes the corresponding Hilbert-Schmidt tensor product. See Definition 7.257.
- Given a complex Hilbert space H, H_{sd} stands for the self-dual Hilbert space associated with H. See Definition 4.195.

Spaces of linear transformations

- $\mathcal{L}(V_1; V_2)$ denotes the vector space of all linear transformations $V_1 \to V_2$ between two vector spaces V_1 and V_2.
- $\mathcal{L}(V) \doteq \mathcal{L}(V; V)$ and the identity mapping is denoted by $\mathrm{id}_V \in \mathcal{L}(V)$, V being a vector space.
- $\mathcal{L}^+(V_1; V_2)$ is the convex cone of positive linear transformations $V_1 \to V_2$ between two preordered vector spaces V_1 and V_2. See Definition 1.7.
- $\mathcal{L}_{\mathrm{ob}}(V_1; V_2)$ is the family of all order-bounded linear transformations $V_1 \to V_2$ between two preordered vector spaces V_1 and V_2. See Definition 1.23.
- $\mathcal{L}_{\mathrm{n}}(V_1; V_2)$ is the set of all order-continuous linear transformation $V_1 \to V_2$ between two Riesz spaces V_1 and V_2. See Definition 7.290.
- $\mathcal{L}_{\mathrm{c}}(V_1; V_2)$ is the set of all σ-order-continuous linear transformation $V_1 \to V_2$ between two Riesz spaces V_1 and V_2. See Definition 7.290.
- Given two normed vector spaces X_1 and X_2,

$$\mathcal{B}(X_1; X_2) \doteq \left\{ \varphi \in \mathcal{L}(X_1; X_2) : \|\varphi\|_{\mathrm{op}} \doteq \sup_{x_1 \in X_1, \|x_1\|=1} \|\varphi(x_1)\| < \infty \right\} .$$

- Given a normed space X, $\mathcal{B}(X) \doteq \mathcal{B}(X; X)$ and the identity mapping is again denoted by $\mathrm{id}_X \in \mathcal{B}(V)$.
- $V'^{+} \doteq \mathcal{L}(V; \mathbb{K})^{+}$, V being a preordered vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.
- $V^{\mathrm{od}} \doteq \mathrm{span}(\mathcal{L}(V; \mathbb{K})^{+})$ is the order dual of a preordered vector space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. See Definition 1.25.
- $V^{\mathrm{td}} \doteq \mathcal{B}(V; \mathbb{K})$ is the topological dual of a preordered normed space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.
- $V^{\mathrm{n}} \doteq \mathcal{L}_{\mathrm{n}}(V; \mathbb{R})$ is the space of normal integrals on an Archimedean Riesz space V. See Definition 7.309.
- $V^{\mathrm{c}} \doteq \mathcal{L}_{\mathrm{c}}(V; \mathbb{R})$ are the space of integrals on an Archimedean Riesz space V. See Definition 7.309.
- $\ker(\Theta) \doteq \{v \in V_1 : \Theta(v) = 0\} \subseteq V_1$ denotes the kernel of a linear transformation $\Theta \in \mathcal{L}(V_1; V_2)$, V_1 and V_2 being two arbitrary vector spaces.
- $\mathrm{ran}(\Theta) \doteq \{\Theta(v) \in V_1\} \subseteq V_1$ denotes the range of a linear transformation $\Theta \in \mathcal{L}(V_1; V_2)$, V_1 and V_2 being two arbitrary vector spaces.

Function spaces

- $\mathcal{F}(M; V)$ is the space of V-valued functions f on a nonempty set M, V being a vector space.
- $\mathcal{F}^{+}(M; V)$ is the positive cone of the preordered vector space $\mathcal{F}(M; V)$, V being a preordered vector space.
- $\mathcal{F}_{\Omega}(M; V)$ is the space of V-valued functions on a nonempty set M that vanish on a subset $\Omega \subseteq M$, V being a vector space.
- $\mathcal{F}_{\mathrm{b}}(M; X)$ is the space of bounded functions $M \to X$ on a nonempty set M, X being a normed space.
- $C(M_1; M_2)$ is the set of all continuous functions $M_1 \to M_2$, M_1 and M_2 being two metric spaces. If M_1 is compact and M_2 a normed space, then $\|f\|_{\infty} \doteq \sup \|f(M)\|$ is the supremum norm of $f \in C(M_1; M_2)$.
- $C_0(M; X)$ is the space of functions $M \to V$ which are continuous and decay at infinity, M and X being a metric space and a normed space, respectively.
- $C_{\Omega}(M; X)$ is the space of continuous functions that vanish on a nonempty set $\Omega \subseteq M$, M and X being a metric space and a normed space, respectively.
- $C_{\mathrm{b}}(M; X)$ is the space of all continuous bounded functions $M \to X$, M and X being a metric space and a normed space, respectively.
- $\mathrm{M}_{\mathrm{b}}(M)$ is the space of bounded complex-valued (Borel-)measurable functions, M being a metric space.
- For any nonempty set M and subset $\Omega \subseteq M$, $\chi_{\Omega} \in \mathcal{F}(M; \mathbb{R})$ denotes the characteristic function of Ω, i.e., $\chi_{\Omega}(p) \doteq 1$ if $p \in \Omega$ and $\chi_{\Omega}(p) \doteq 0$, else.
- $\gamma(f)$ is the γ-regularization of a real-valued function f. See Definition 7.340.
- f^{*} is the Legendre-Fenchel transform of a function f. See Definition 7.348.
- f^{**} is the double Legendre-Fenchel transform of a function f. See Definition 7.351.
- $\mathrm{A}(M)$ is the space of all affine (i.e., both convex and concave) continuous real-valued functions on a convex subset M of a metric vector space.

- $\text{epi}(f)$ is the epigraph of a function f. See Definition 7.147.
- ∂f is the subdifferential of f. See Definition 7.355.

Algebras

- $\mathcal{A} \equiv (\mathcal{A}, +, \cdot, \circ)$ denotes a generic algebra and A, A', B, its elements.
- $\mathcal{A} \equiv (\mathcal{A}, +, \cdot, \circ, ^*)$ denotes a generic $*$-algebra and A, A', B, its elements.
- $\mathcal{A} \equiv (\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ denotes a generic normed algebra and A, A', B, its elements.
- $\mathcal{A} \equiv (\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ denotes a generic normed $*$-algebra and A, A', B, its elements.
- $\tilde{\mathcal{A}}$ denotes the unitization of an algebra $\mathcal{A}$.
- For any $A \in \mathcal{A}$, $|A|^2 \doteq A^* A$.
- M stands for a generic von Neumann algebra. (We use a notation similar to that for measurable functions, as these algebras are, in a sense, a generalization of the notion of measurable functions.)
- For any A, B in some algebra, $[A, B] \doteq AB - BA$ denotes the corresponding commutator.
- $\mathcal{Z}(\mathcal{A})$ is the center of an algebra $\mathcal{A}$. See Definition 7.23.
- Ideals of algebras are denoted by $\mathcal{I}$. See Definition 7.24.
- $\mathcal{P}$ usually denotes an arbitrary polynomial. In algebras, it refers to Definition 4.22.
- The $*$-subalgebra of all polynomials applied of some fixed element A of a $*$-algebra $\mathcal{A}$ is denoted by Pol_A. See Sect. 4.6.1.
- $R(A) \equiv R_{\mathcal{A}}(A)$ is the resolvent set of any element A of an algebra $\mathcal{A}$. See Definition 4.18.
- $\sigma(A) \equiv \sigma_{\mathcal{A}}(A) \doteq \mathbb{K}\backslash R(A)$ is the spectrum of any element A in an algebra $\mathcal{A}$. See Definition 4.18.
- For all A in a unital algebra and $\lambda \in R(A)$, $\mathcal{R}(A, \lambda)$ is the resolvent of A at λ. See Definition 4.18.
- $r(A) \equiv r_{\mathcal{A}}(A)$ is the spectral radius of any element A in an algebra $\mathcal{A}$. See Sect. 4.2.
- $r^{\text{Gelf}}(A) \equiv r^{\text{Gelf}}_{\mathcal{A}}(A)$ is Gelfand radius of any element A in a normed algebra $\mathcal{A}$. See Definition 4.28.
- $A^+_{\text{Re}}, A^-_{\text{Re}}, A^+_{\text{Im}}, A^-_{\text{Im}}$ refer to the positive decomposition of any element A of a C^*-algebra. See Propositions 2.6 and 2.15, 4.46 and Lemma 4.38.
- The function $\Xi(A) : E(\mathcal{A}) \to \mathbb{C}$ is the Gelfand transform of any element A in a C^*-algebra $\mathcal{A}$. See Definition 4.79.
- A generic representation of a $*$-algebra $\mathcal{A}$ is denoted by (H, π). If it is cyclic, then we use the notation (H, π, Ω), where Ω stands for a cyclic norm-one vector. See Definition 4.88.
- Given a nonempty (index) set I, a four-tuple $(\mathcal{P}, \bar{J}, J, \eta)$, where $\eta \in \mathbb{R}^+_0$, J, and $\bar{J}$ are, respectively, mappings $\{1, \ldots, n_J\} \to I$ and $\{1, \ldots, n_{\bar{J}}\} \to I$, $n_J, n_{\bar{J}} \in \mathbb{N}_0$, $n_J + n_{\bar{J}} \geq 1$, and $\mathcal{P}$ is a complex polynomial in $n_J + n_{\bar{J}}$ non-commuting variables, denotes a generic polynomial relation. See Definition 4.127.

- R stands for a generic family of polynomial relations.
- The pair $(\mathcal{A}, a)$, where $\mathcal{A}$ is a C^*-algebra and a a mapping $I \to \mathcal{A}$, denotes a generic C^*-representation of a family of polynomial relations R for the (index) set I. See Definition 4.127.
- $(C^*(I, \mathrm{R}), a)$ denotes a universal C^*-representation of a family R of polynomial relations for the index set I. Sometimes such a C^*-representation is denoted only by $C^*(I, \mathrm{R})$.
- $C^*(I, \mathrm{R})$ stands for the universal C^*-algebra of a family R of polynomial relations for the index set I.
- $C^*(\mathcal{G})$ denotes the C^*-subalgebra of a C^*-algebra $\mathcal{A}$, that is generated by a set $\mathcal{G}$ of elements of $\mathcal{A}$.
- $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega \equiv (\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega, a)$ denotes the universal tensor product of the family $\{\mathcal{A}_\omega\}_{\omega\in\Omega}$ of unital C^*-algebras. See Definition 4.147.
- For any subset Ω' of Ω, $\mathrm{i}_{\Omega\Omega'}$ denotes the natural $*$-homomorphism $\bigotimes_{\omega\in\Omega'} \mathcal{A}_\omega \to \bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$. See Definition 4.150.
- Given a nonempty family G of vectors in a pre-Hilbert space, $\mathrm{CAR}(G)$ denotes the corresponding CAR C^*-algebra. See Definition 4.163.
- For any subset G' of G, $\mathrm{i}_{GG'}$ denotes the canonical inclusion (i.e., faithful $*$-homomorphism) $\mathrm{CAR}(G') \to \mathrm{CAR}(G)$. See Definition 4.168.
- Given a nonempty family G of vectors in a pre-Hilbert space, $\mathrm{CAR}_\circ(G)$ is the corresponding gauge-invariant CAR C^*-algebra. See Definition 4.185 (i).
- Given a nonempty family G of vectors in a pre-Hilbert space, $\mathrm{CAR}_\mathrm{e}(G)$ is the corresponding even CAR C^*-subalgebra. See Definition 4.185 (ii).
- Given G_1 and G_2 nonempty families vectors in pre-Hilbert spaces, for any mapping $U : G_1 \to G_2$ preserving the scalar product, $\mathrm{Bog}(U)$ stands for the corresponding Bogoliubov $*$-homomorphism $\mathrm{CAR}(G_1) \to \mathrm{CAR}(G_2)$. See Definition 4.181.
- Given a self-dual Hilbert space H, $\mathrm{sCAR}(H)$ denotes the corresponding self-dual CAR C^*-algebra. See Definition 4.204.
- Given a self-dual Hilbert space H, $\mathrm{sCAR}_\mathrm{e}(H)$ stands for the corresponding even self-dual CAR C^*-algebra. See Definition 4.211.
- Given H_1 and H_2 self-dual Hilbert spaces, for any $*$-morphism $U : H_1 \to H_2$ preserving the scalar product, $\mathrm{Bog}(U)$ stands for the corresponding self-dual Bogoliubov $*$-homomorphism $\mathrm{sCAR}(H_1) \to \mathrm{sCAR}(H_2)$. See Definition 4.209.

Bounded operators on Hilbert spaces

- Generic Hilbert spaces are denoted by H. Note that the symbol H may also refer to Hamiltonians of finite quantum systems, with no risk of confusion.
- The scalar product of a generic Hilbert spaces H is denoted by $\langle\cdot,\cdot\rangle$, while the corresponding norm is $\|\varphi\| \equiv \|\varphi\|_{\langle\cdot,\cdot\rangle} \doteq \langle\varphi,\varphi\rangle^{1/2}$ for any $\varphi \in H$.
- $\Omega' \subseteq \mathcal{B}(H)$ denotes the commutant of a (nonempty) set $\Omega \subseteq \mathcal{B}(H)$, while $\Omega'' = (\Omega')'$. See Definition 7.253.
- $\mathcal{B}^+(H)$ stands for the positive cone of $\mathcal{B}(H)$. See Definition 2.1.
- $\mathcal{B}(H)^\mathbb{R}$ is a short notation for $\mathrm{Re}\{\mathcal{B}(H)\}$, H being a generic Hilbert space. In the same way, $\mathcal{L}(\mathbb{C}^n)^\mathbb{R}$ stands for $\mathrm{Re}\{\mathcal{L}(\mathbb{C}^n)\}$, $n \in \mathbb{N}$.

- $\mathrm{Tr}(\cdot)$ is the usual trace of matrix algebras. See Sect. 2.5.
- $\|\cdot\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm. See Definition 2.35.
- Given a subspace $G \subseteq H$ of a Hilbert space H, P_G denotes the orthogonal projector on H, whose range is G. See, for instance, Sect. 2.3. In the context of Riesz spaces, we use, mutatis mutandis, the same notation. See Lemma 7.303.
- U denotes a generic unitary operator on a Hilbert space.
- $N(A)$ denotes the numerical range of $A \in \mathcal{B}(H)$, H being a Hilbert space. See Definitions 2.1 and 7.241.

States

- $E(\mathcal{A})$ denotes the set of all states on a C^*-algebra $\mathcal{A}$. See Definition 4.61.
- $E(H) \equiv E(\mathcal{B}(H))$ is the set of all states on $\mathcal{B}(H)$, H being an Hilbert space. See Definition 2.25.
- For any norm-one vector x in a Hilbert space H, $\rho_x \in E(H)$ is the associated vector state of Definition 2.29.
- For any $n \in \mathbb{N}$, $\rho_D \in E(\mathbb{C}^n)$ is the state whose density matrix is $D \in \mathcal{L}(\mathbb{C}^n)^+$. See Sect. 3.1.
- For any $n \in \mathbb{N}$, every state $\rho \in E(\mathbb{C}^n)$ and unitary $U \in \mathcal{L}(\mathbb{C}^n)$, ρ_U is the state of Lemma 3.29.
- Given a compact metric space K and $p \in K$, $\rho_p \in E(C(K;\mathbb{C}))$ is the state of Example 4.65.
- μ denotes a generic probability measure in some metric space. See Definition 4.10.
- Given a compact metric space K, $\rho_\mu \in E(C(K;\mathbb{C}))$ denotes the restriction a probability measures μ to $C(K;\mathbb{C})$. See Example 4.67.
- Given a cyclic representation $R \doteq (H, \pi, \Omega)$ of a unital C^*-algebra $\mathcal{A}$ (Definition 4.88), $\rho_R \in E(\mathcal{A})$ denotes the state defined by Eq. (4.4).
- $\rho \mapsto \rho^{\mathcal{W}_1}$ is the mapping of Definition 5.14.
- $(H_\rho, \pi_\rho, \Omega_\rho)$ is a cyclic representation of a state $\rho \in E(\mathcal{A})$. See Theorem 4.113.
- Given a nonempty family G of vectors in a pre-Hilbert space, $E_\circ(\mathrm{CAR}(G)) \subseteq E(\mathrm{CAR}(G))$ is the set of all gauge-invariant states of $\mathrm{CAR}(G)$. See Definition 4.190.
- $E_{\mathrm{e}}(\mathrm{CAR}(G)) \subseteq E(\mathrm{CAR}(G))$ is the set of all even states of $\mathrm{CAR}(G)$. See Definition 4.190.
- $E_{\mathrm{e}}(\mathrm{sCAR}(H)) \subseteq E(\mathrm{CAR}(H))$ is the set of all even states of $\mathrm{sCAR}(H)$. See Definition 4.211.
- Given a collection of states $\rho_\omega \in E(\mathcal{A}_\omega)$, $\omega \in \Omega$, of unital C^*-algebras $\mathcal{A}_\omega$, $\omega \in \Omega$, $\bigotimes_{\omega\in\Omega} \rho_\omega \in E(\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega)$ denotes the corresponding product state. See Proposition 4.155.
- Given an orthonormal family G of vectors in a pre-Hilbert space, along with a partition P of this family, $\bigotimes_{\Lambda\in\mathrm{P}} \rho_\Lambda \in E_{\mathrm{e}}(\mathrm{CAR}(G))$ denotes the product of the even states $\rho_\Lambda \in E_{\mathrm{e}}(\mathrm{CAR}(\Lambda))$. See Proposition 4.193.
- Given a self-dual Hilbert space H, along with a family $\{G_i\}_{i\in I}$ of closed self-conjugate subspaces that are mutually orthogonal and span the whole space H,

for a family $\rho_i \in E_{\mathrm{e}}(\mathrm{sCAR}(G_i))$, $i \in I$, of even states, $\otimes_{i\in I}\rho_i \in E_{\mathrm{e}}(\mathrm{sCAR}(H))$ denotes the corresponding product state. See Corollary .
- Given a nonempty family G of vectors in a pre-Hilbert space, $\rho_{\mathrm{Fock}}^{G} \in E(\mathrm{CAR}(G))$ is the corresponding Fock state. See Proposition 4.176.
- Given a basis projection P on a self-dual Hilbert space H, $\rho_{\mathrm{Fock}}^{P} \in E(\mathrm{sCAR}(H))$ is the corresponding Fock state. See Definition 4.197 and Corollary 4.215.
- Given a self-dual Hilbert space H, $S_\rho \in \mathcal{B}(H)$ denotes the basis symbol of a state ρ of a sCAR C^*-algebra $\mathrm{sCAR}(H)$. See Definition 4.202 and Exercise 4.216.
- Given a self-dual Hilbert space H and a basis symbol $S \in \mathcal{B}(H)$, $\rho_S \in E_{\mathrm{e}}(\mathrm{sCAR}(H))$ is the unique quasi-free state whose basis symbol is S. See Definition 4.217 and Proposition 4.219.
- Given a complex Hilbert space H and a basis symbol $S \in \mathcal{B}(H_{\mathrm{sd}})$, $\rho_S \in E_{\mathrm{e}}(\mathrm{CAR}(H))$ is the unique quasi-free state whose basis symbol is S. See remarks after Proposition 4.219.

Finite quantum systems

- $S(\cdot)$ stands for the von Neumann entropy of states of the matrix algebra $\mathcal{L}(\mathbb{C}^n)$, $n \in \mathbb{N}$. See Definition 3.8.
- $S_{\mathcal{A}}(\cdot)$ denotes the von Neumann entropy of states of an algebra $\mathcal{A}$ that is $*$-isomorphic to the matrix algebra $\mathcal{L}(\mathbb{C}^n)$, $n \in \mathbb{N}$. See Definition 5.17.
- $F_{H,\beta}(\cdot)$ is the free energy functional associated with the Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$, at inverse temperature $\beta \in (0,\infty)$. See Definition 3.9.
- $P_{H,\beta}$ stands for the pressure associated with the Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ at inverse temperature $\beta \in (0,\infty)$. See Definition 3.16.
- $\omega_{H,\beta}$ is the Gibbs state associated with the Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ at $\beta \in (0,\infty)$. See Definition 3.7.
- $\{\tau_t^H\}_{t\in\mathbb{R}}$ is the group of $*$-automorphisms of $\mathcal{L}(\mathbb{C}^n)$ of Definition 3.26, associated with an Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$.

Algebraic setting of quantum lattice systems

- $\Gamma \doteq \mathbb{Z}^d$ for a fixed space dimension $d \in \mathbb{N}$.
- $\mathcal{P}_f \subsetneq 2^\Gamma$ denotes the set of all finite subsets of Γ.
- $d(\Lambda)$ is the diameter of a finite subset $\Lambda \in \mathcal{P}_f$. See Eq. (5.1).
- For all $\ell \in \mathbb{N}$, $\Lambda_\ell \doteq (\mathbb{Z} \cap [-\ell,\ell])^d \in \mathcal{P}_f$.
- Given $N, d \in \mathbb{N}$, $\mathrm{Spin}(N,\Lambda) \doteq \bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^N)$ is the spin algebra of the region $\Lambda \subseteq \Gamma$.
- $\mathrm{CAR}(\Omega,\Lambda) \doteq \mathrm{CAR}(\ell^2(\Omega\times\Lambda))$ is the CAR C^*-algebra of the region $\Lambda \subseteq \Gamma$, Ω being a nonempty finite (spin) set.
- $\mathcal{U}$ stands for either $\mathrm{Spin}(N,\Gamma)$ or $\mathrm{CAR}(\Omega,\Gamma)$, for a fixed $N \in \mathbb{N}$ and a nonempty finite set Ω, respectively.
- $\mathcal{U}^e$ denotes either $\mathrm{Spin}(N,\Gamma)$ or $\mathrm{CAR}_{\mathrm{e}}(\Omega,\Gamma)$.
- For every $\Lambda \in \mathcal{P}_f$, $\mathcal{U}_\Lambda \doteq \mathrm{Spin}(N,\Lambda)$ (quantum spin case) or $\mathcal{U}_\Lambda \doteq \mathrm{CAR}(\Omega,\Lambda)$ (fermion case).
- $\mathcal{U}_{\mathrm{loc}} \doteq \bigcup_{\Lambda\in\mathcal{P}_f} \mathcal{U}_\Lambda$.

- $\tau_x : \mathcal{U} \to \mathcal{U}$, $x \in \Gamma$, are the translation automorphisms of $\mathcal{U}$ defined by (5.4) in the quantum spin case or (5.9) in the fermion case.
- A_ℓ, $\ell \in \mathbb{N}$, are the space-averaged elements associated with $A \in \mathcal{U}$. See Corollary 5.34.
- $E_1(\mathrm{Spin}(N, \Gamma))$ denotes the set of all invariant states of $\mathrm{Spin}(N, \Gamma)$ and $\mathcal{E}_1(\mathrm{Spin}(N, \Gamma))$, the one of the corresponding ergodic states. See Definition 5.1.
- $E_1(\mathrm{CAR}(\Omega, \Gamma))$ denotes the set of all invariant states of $\mathrm{CAR}(\Omega, \Gamma)$ and $\mathcal{E}_1(\mathrm{CAR}(\Omega, \Gamma))$, the one of the corresponding ergodic states. See Definition 5.2.
- E_1 and $\mathcal{E}_1$, respectively, denote $E_1(\mathrm{Spin}(N, \Gamma))$ and $\mathcal{E}_1(\mathrm{Spin}(N, \Gamma))$ in the quantum spin case, while they denote $E_1(\mathrm{CAR}(\Omega, \Gamma))$ and $\mathcal{E}_1 = \mathcal{E}_1(\mathrm{CAR}(\Omega, \Gamma))$, in the fermion case.
- $\mathcal{V}$ denotes the space of (real) interactions and its generic elements are denoted by Φ, Φ', Ψ. See Definition 5.5.
- $\mathcal{V}^{\mathbb{C}}$ denotes de space of complex interactions and its generic elements are denoted by Φ, Φ', Ψ.
- $\mathcal{W}_1$ is the Banach space of (real) invariant interactions. See Definition 5.6.
- $\mathcal{W}_1^{\mathbb{C}}$ is the Banach space of complex invariant interactions. See Definition 6.2.
- $\mathcal{M}_1$ is the Banach space of mean-field models. See Definition 6.3. Its generic elements are denoted by $\mathfrak{m}$.
- For all $\Lambda \in \mathcal{P}_f$, H_Λ^Φ is the local energy observable of Definition 5.10 associated with an interaction Φ.
- For all $\Lambda \in \mathcal{P}_f$, $H_\Lambda^{\mathfrak{m}}$ is the local energy observable of Definition 6.4 associated with a mean-field model $\mathfrak{m} \in \mathcal{M}_1$.
- e_Φ is the energy density observable of Definition 5.10 and Eq. (6.6), associated with $\Phi \in \mathcal{W}_1^{\mathbb{C}}$.
- Δ_A is the space-averaging functional of Definition 6.9, associated with $A \in \mathcal{U}$.

Thermodynamic functions of quantum lattice systems

- For every nonempty $\Lambda \in \mathcal{P}_f$, $F_{H,\beta}(\cdot)$ is the free energy functional associated with the local Hamiltonian $H \in \mathrm{Re}\{\mathcal{U}_\Lambda\}$ at inverse temperature $\beta \in (0, \infty)$. See Definition 5.19.
- $P_{H,\beta}$ stands for the pressure associated with the local Hamiltonian $H \in \mathrm{Re}\{\mathcal{U}_\Lambda\}$ at inverse temperature $\beta \in (0, \infty)$. See Definition 5.19.
- e_Φ denotes the energy density functional associated with the complex invariant interaction $\Phi \in \mathcal{W}_1^{\mathbb{C}}$. See Definition 5.12 and Eq. (6.7). See also Definition 6.17 for the definition of the related object e_Ψ, $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$.
- Given $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$, Δ_Ψ stands for the corresponding long-range energy density functional. See Definition 6.12.
- s denotes the entropy density functional of Theorem 5.20.
- $\mathrm{f}_{\Phi,\beta}$ denotes the free energy density functional of Definition 5.22, associated with the invariant interactions $\Phi \in \mathcal{W}_1$ at inverse temperature $\beta \in (0, \infty)$.
- $\mathrm{f}_{\mathfrak{m},\beta}$ is the free energy density functional of Definition 6.12, associated with the mean-field model $\mathfrak{m} \in \mathcal{M}_1$ at inverse temperature $\beta \in (0, \infty)$.

- $\mathrm{g}_{\mathrm{m},\beta}$ stands for the nonlinear free energy density of Definition 6.17, associated with the mean-field model $\mathrm{m} \in \mathcal{M}_1$ at inverse temperature $\beta \in (0, \infty)$.
- $\mathrm{f}^{\flat}_{\mathrm{m},\beta}$ is the non-conventional free energy density functional of Eq. (6.30), associated with the mean-field model $\mathrm{m} \in \mathcal{M}_1$ at inverse temperature $\beta \in (0, \infty)$.
- $h_{\mathrm{m},\beta}$ denotes the approximating free energy density of Definition 6.28, associated with the mean-field model $\mathrm{m} \in \mathcal{M}_1$ at inverse temperature $\beta \in (0, \infty)$.
- p_{β} stands for the pressure function of Definitions 5.28 and 6.14 at inverse temperature $\beta \in (0, \infty)$.
- $h^{\sharp}_{\mathrm{m},\beta}$,$\mathrm{F}^{\flat}_{\mathrm{m},\beta}$ and $h^{\flat}_{\mathrm{m},\beta}$, $\mathrm{F}^{\sharp}_{\mathrm{m},\beta}$ the functions defined in Sect. 6.7 for any mean-field model $\mathrm{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$.
- $\mathcal{C}^{\flat}_{\mathrm{m},\beta}$ and $\mathcal{C}^{\sharp}_{\mathrm{m},\beta}$, defined by (6.28), are the sets of conservative strategies for the thermodynamic game associated with the mean-field model $\mathrm{m} \in \mathcal{M}_1$ at inverse temperature $\beta \in (0, \infty)$. We also use the notation $\mathcal{C}_{\mathrm{m},\beta}$ in some special cases, see (6.23).
- $P^{\sharp}_{\beta}(\eta_-, \eta_+)$ and $P^{\flat}_{\beta}(\eta_-, \eta_+)$ are, respectively, the conventional and non-conventional pressures (6.44) and (6.48) at inverse temperature $\beta \in (0, \infty)$ for a Kac-type model with parameters $\eta_-, \eta_+ \in \mathbb{R}^+_0$. In this context, see Eq. (6.49) for the definition of P_{β}.

Equilibrium states of quantum lattice systems

- $\omega_{H,\beta}$ is the Gibbs state associated with the Hamiltonian $H \in \mathrm{Re}\{\mathcal{U}_\Lambda\}$ at inverse temperature $\beta \in (0, \infty)$, i.e., the unique minimizer of $F_{H,\beta}(\cdot)$ in $E(\mathcal{U}_\Lambda)$. See Definition 5.19.
- $M_{\Phi,\beta}$ is the set of all minimizers of the free energy density $\mathrm{f}_{\Phi,\beta}$ for fixed $\Phi \in \mathcal{W}_1$ and $\beta \in (0, \infty)$. See Definition 5.22.
- $\Omega_{\mathrm{m},\beta}$ is the set of all approximated minimizers of the free energy density $\mathrm{f}_{\mathrm{m},\beta}$ for fixed $\mathrm{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$. See Definition 6.15.
- $\hat{M}_{\mathrm{m},\beta}$ is the set of all minimizers of $\mathrm{g}_{\mathrm{m},\beta}$ for fixed $\mathrm{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$. See Definition 6.21.
- $\Omega^{\flat}_{\mathrm{m},\beta}$ is the set of all minimizers of $\mathrm{f}^{\flat}_{\mathrm{m},\beta}$ for fixed $\mathrm{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$. See Sect. 6.8.
- Elements of $M_{\Phi,\beta}$, $\Omega_{\mathrm{m},\beta}$, $\hat{M}_{\mathrm{m},\beta}$, and $\Omega^{\flat}_{\mathrm{m},\beta}$ are denoted by ω.

Contents

Chapter 1
Ordered Vector Spaces and Positivity

1.1 Basic Notions

Let V be a vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We say that $P \subseteq V$ is a "cone" in V if, for all $\alpha \geq 0$ and $v \in P$, $\alpha v \in P$. The cone $P \subseteq V$ is a "convex cone" or "wedge" in V if, for all $v, v' \in P$, one has $v + v' \in P$. Note that any convex cone $P \subseteq V$ is a convex set in the usual sense: For all $\alpha \in [0, 1]$ and $v, v' \in P$, $\alpha v + (1 - \alpha)v' \in P$.

Any convex cone $P \subseteq V$ naturally defines a preorder relation in V, denoted here by $\succeq^P$:

$$v \succeq^P v' \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad v - v' \in P , \qquad v, v' \in V .$$

Recall that a relation $\succeq$ in V is, by definition, a preorder, if it is reflexive (i.e., $v \succeq v$ for all $v \in V$) and transitive (i.e., for all $v, v', v'' \in P$, one has that $v \succeq v'$ and $v' \succeq v''$ only if $v \succeq v''$). Observe that the above-defined preorder has additionally the following invariance properties:

(i) For all $v, v', w, w' \in V$ with $v \succeq^P v'$ and $w \succeq^P w'$, one has $v + w \succeq^P v' + w'$.
(ii) For all $\alpha \geq 0$ and $v, v' \in V$ with $v \succeq^P v'$, one has $\alpha v \succeq^P \alpha v'$.

Conversely, if $\succeq$ is a preorder in V having the two properties above, then it is the preorder associated with the convex cone

$$V^+ \doteq \{v \in V \ : \ v \succeq 0\} .$$

For this reason, we identify preorders in V, which satisfy the properties (i) and (ii), with their corresponding convex cones V^+.

Definition 1.1 (Preordered Vector Space) Let V be any vector space and $\succeq$ the preorder in V associated with a convex cone $V^+ \subseteq V$. The pair $(V, \succeq)$ is called

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_1

"preordered vector space," while V^+ is named the "positive cone" of this space, the elements of which are called "positive elements" of V. (Recall that $v \in V^+$ iff $v \succeq 0$.)

For simplicity, a generic preordered vector space $(V, \succeq)$, which is formally a pair, is sometimes denoted by the name of the vector space on which its preorder (or positive cone) is defined, i.e., V.

Definition 1.2 (Archimedean Property and Order Units) We say that the preordered vector space V is "Archimedean" if, for all $v \in V$ and $\tilde{v} \notin -V^+$ there is $\alpha \in \mathbb{R}_0^+$ such that $v \succeq \alpha\tilde{v}$ does *not* hold, i.e., if $0 \succeq \tilde{v}$ whenever $v \succeq \alpha\tilde{v}$ for all $\alpha \in \mathbb{R}_0^+$ and some $v \in V$. We say that the positive element $v \in V^+$ "dominates" a given subset $\Omega \subseteq V$ if, for any $v' \in \Omega$, there is $\alpha \in \mathbb{R}_0^+$ such that $\alpha v \succeq v'$. If V is a *real* vector space, we say that $u \in V$ is an "order unit" of V if it dominates the whole space V.

The simplest example of an Archimedean preordered vector space (over $\mathbb{R}$) is $(\mathbb{R}, \geq)$, where $\geq$ is the usual order for real numbers. Note that $\mathbb{R}_0^+ \doteq [0, \infty) \subseteq \mathbb{R}$ is the positive cone of $(\mathbb{R}, \geq)$ and any strictly positive real number is an order unit of $(\mathbb{R}, \geq)$. Remark that $\mathbb{C}$ endowed with the positive cone $\mathbb{R}_0^+ \subseteq \mathbb{C}$ is again an Archimedean ordered vector space, whereas it is non-Archimedean for the positive cone

$$\{0\} \cup \{z \in \mathbb{C} \ : \ \mathrm{Re}\{z\} > 0\} .$$

In the sequel, by default, the positive cone of both $\mathbb{R}$ and $\mathbb{C}$ will be $\mathbb{R}_0^+$.

Note that if V is a vector space, $P \subseteq V$ a convex cone and $W \subseteq V$ a vector subspace of V, then $P \cap W \subseteq W$ is clearly a convex cone in W. In fact, the preorder relation $\succeq^{P\cap W}$ in W is nothing else than the restriction to W of the preorder relation $\succeq^P$in V. Thus, we canonically see vector subspaces of preordered vector spaces again as preordered vector spaces. Remark that vector subspaces of an Archimedean preordered vector space are again Archimedean. This is so because, W being a vector subspace, one has the identity

$$-(P \cap W) = (-P) \cap W .$$

Another important example of an Archimedean preordered vector space is the following:

Example 1.3 Let $\mathcal{F}(\Omega; \mathbb{R})$ be the set of real-valued functions on a nonempty set Ω. Define the convex cone

$$\mathcal{F}^+(\Omega; \mathbb{R}) \doteq \{f \in \mathcal{F}(\Omega; \mathbb{R}) \ : \ f(p) \geq 0 \text{ for all } p \in \Omega\} \subseteq \mathcal{F}(\Omega; \mathbb{R}) .$$

$\mathcal{F}(\Omega; \mathbb{R})$ as preordered vector space, whose positive cone is $\mathcal{F}^+(\Omega; \mathbb{R})$, is Archimedean. More generally, if $(V, \succeq)$ is an arbitrary preordered vector space and $\mathcal{F}(\Omega; V)$ denotes the vector space of V-valued functions on Ω, we define the convex cone:

$$\mathcal{F}^+(\Omega; V) \doteq \{f \in \mathcal{F}(\Omega; V) \ : \ f(p) \in V^+ \text{ for all } p \in \Omega\} \subseteq \mathcal{F}(\Omega; V) \ .$$

$\mathcal{F}(\Omega; V)$ as preordered vector space, whose positive cone is $\mathcal{F}^+(\Omega; V)$, is Archimedean, whenever $(V, \succeq)$ is Archimedean.

By default, for any nonempty set Ω and preordered vector space V, $\mathcal{F}(\Omega; V)$ will be seen as the preordered vector space, whose positive cone is defined as in the last example. Note that, in contrast to $(\mathbb{R}, \geq)$, a non-zero positive element $f \in \mathcal{F}^+(\Omega; V)$ is not necessarily an order unit of $\mathcal{F}(\Omega; V)$, even if it is a *strictly* positive function (i.e., $f(p) \geq 0$ with $f(p) \neq 0$ for all $p \in \Omega$). If K is a compact metric space and $C(K; \mathbb{R})$ denotes the set of continuous real-valued functions on K, then any strictly positive function $f \in C(K; \mathbb{R})$ is an order unit of $C(K; \mathbb{R})$. However, if a metric space M is not compact, even functions that are strictly positive are again not necessarily order units of $C(M; \mathbb{R})$.

Definition 1.4 (Intervals, Order-Full Subsets, and Order Ideals) For any two vectors v, v' in the preordered vector space V, we define the "interval"

$$[v, v'] \doteq (v + V^+) \cap (v' - V^+) = \{v'' \in V \ : \ v' \succeq v'' \succeq v\} \subseteq V \ .$$

A subset $\Omega \subseteq V$ is called "(order-)full" if, for all $v, v' \in \Omega$, one has $[v, v'] \subseteq \Omega$. In particular, any interval in V is order-full. Order-full vector subspaces of V are called "(order) ideals" of V.

Notice that any interval in a preordered vector space is a convex subset, as the positive cone of the space is convex. The following examples of order ideals of preordered vector spaces are prototypical.

Exercise 1.5

(i) Let Ω be a nonempty set. For all $\widetilde{\Omega} \subseteq \Omega$, define the subspace:

$$\mathcal{F}_{\widetilde{\Omega}}(\Omega; \mathbb{R}) \doteq \{f \in \mathcal{F}(\Omega; \mathbb{R}) \ : \ f(\widetilde{\Omega}) = \{0\}\} \subseteq \mathcal{F}(\Omega; \mathbb{R}) \ .$$

Show that all such subspaces are order ideals of $\mathcal{F}(\Omega; \mathbb{R})$.

(ii) Let M be any metric space. Show that $C_0(M; \mathbb{R})$ (i.e., the space of continuous real-valued functions "decaying at infinity"; see Definition 7.166) is an order ideal of $C(M; \mathbb{R})$.

Observe that, in general, the subspace $C(M; \mathbb{R})$ of continuous real-valued functions on a metric space M is not an order ideal of $\mathcal{F}(M; \mathbb{R})$.

Any complex vector space is canonically seen a real vector space by restricting the scalar multiplication in the field $\mathbb{C}$ to $\mathbb{R}$. A $*$-vector space V is a complex vector space endowed with an antilinear[1] involution[2] $v \mapsto v^*$. In this case, the real part of a $*$-vector space V is the subset

$$\mathrm{Re}\{V\} \doteq \left\{v \in V : v^* = v\right\} \subseteq V ,$$

which is a real vector subspace of V. See the appendix for more details. If the vector space, on which a convex cone (i.e., a preorder) is given, is a $*$-vector space, one usually imposes the following compatibility condition between the preorder and the complex conjugation:

Definition 1.6 ($*$-Preordered Vector Space) Let V be a (complex) preordered vector space. We say that it is a "$*$-preordered vector space" if V is a $*$-vector space and $V^+ \subseteq \mathrm{Re}\{V\}$. In particular, in this case, $\mathrm{Re}\{V\}$ is a (real) order ideal of V, seen as a real vector space. We say that $u \in \mathrm{Re}\{V\}$ is an order unit of the complex vector space V if it is an order unit of the real vector space $\mathrm{Re}\{V\}$, in the previous sense.

Note that $\mathbb{C}$, the positive cone of which is $\mathbb{R}_0^+$, is a $*$-preordered vector space. By contrast, with the positive cones

$$\mathbb{R}_0^+ + i\mathbb{R} \qquad \text{and} \qquad \{0\} \cup \{z \in \mathbb{C} : \mathrm{Re}\{z\} > 0\} ,$$

it is not. Observe also that, for any nonempty set Ω and any $*$-preordered vector space V, $\mathcal{F}(\Omega; V)$ is again a $*$-preordered vector space. In this case, the antilinear involution $f \mapsto f^*$ on $\mathcal{F}(\Omega; V)$ is defined pointwise: $f^*(p) = f(p)^*$ for all $p \in \Omega$ and functions $f \in \mathcal{F}(\Omega; V)$. See also Example 7.61.

If V_1 and V_2 are two preordered vector spaces, then $\mathcal{L}(V_1; V_2)$ denotes the vector space of all linear transformations (or mappings) from V_1 to V_2, the vector space structure of which is defined by the pointwise addition and scalar multiplication. In this case, there is a natural notion of positivity for linear transformations:

Definition 1.7 (Cone of Positive Linear Transformations) Let V_1 and V_2 be any two preordered vector spaces. We define the following convex cone in $\mathcal{L}(V_1; V_2)$:

$$\mathcal{L}^+(V_1; V_2) \doteq \{\Theta \in \mathcal{L}(V_1; V_2) \ : \ \Theta(v_1) \in V_2^+ \text{ for all } v_1 \in V_1^+\} .$$

The elements of $\mathcal{L}^+(V_1; V_2)$ are called "positive," "positivity preserving" linear transformations, or "morphisms of preordered vector spaces." We say that a positive linear transformation $\Theta : V_1 \to V_2$ is "bipositive" if it reflects positivity, i.e., for all $v_1 \in V_1$, $\Theta(v_1) \in V_2^+$ only if $v_1 \in V_1^+$.

By default, for any two preordered vector spaces V_1 and V_2, $\mathcal{L}(V_1; V_2)$ is the preordered vector space whose positive cone is $\mathcal{L}^+(V_1; V_2)$, i.e., $\Theta \in \mathcal{L}(V_1; V_2)$ is a positive element of $\mathcal{L}(V_1; V_2)$ if it is positive valued on the positive cone V_1^+.

[1] For any $\lambda \in \mathbb{C}$ and $v, w \in V$, $(\lambda v)^* = \bar{\lambda} v^*$ and $(v + w)^* = v^* + w^*$.

[2] For any $v \in V$, $(v^*)^* = v$.

Observe that, in this case, for any two linear transformations $\Theta, \Theta' \in \mathcal{L}(V_1; V_2)$, $\Theta \succeq \Theta'$ iff $\Theta(v_1) \succeq \Theta'(v_1)$ for all $v_1 \in V_1^+$. In particular, dual (vector) spaces[3] of preordered vector spaces are again preordered vector spaces, their positive cones being the convex cones of positive functionals. The positive cone of the dual space of a given preordered vector space V is called "dual positive cone" of V.

In fact, positive linear transformations $\Theta \in \mathcal{L}^+(V_1; V_2)$ are special cases of order-preserving mappings $V_1 \to V_2$, i.e., $\Theta(v_1') \succeq \Theta(v_1)$ when $v_1' \succeq v_1$. Bipositive linear transformations are, additionally, "order-reflecting," that is, for all $v_1, v_1' \in V_1$, $\Theta(v_1) \succeq \Theta(v_1')$ only if $v_1 \succeq v_1'$.

At this point it is important to notice that $\mathcal{L}(V_1; V_2)$ is a vector subspace of the space $\mathcal{F}(V_1; V_2)$ of functions $V_1 \to V_2$ and, hence, also inherits the (canonical) preorder relation from $\mathcal{F}(V_1; V_2)$. See Example 1.3. However, in general,

$$\mathcal{L}^+(V_1; V_2) \neq \mathcal{L}(V_1; V_2) \cap \mathcal{F}^+(V_1; V_2)$$
$$= \{\Theta \in \mathcal{L}(V_1; V_2) \ : \ \Theta(v_1) \in V_2^+ \text{ for all } v_1 \in V_1\} \ .$$

In dealing with the vector space $\mathcal{L}(V_1; V_2)$ of *linear* transformations between preordered vector spaces V_1 and V_2, we will *not* consider the convex cones for subspaces of the (canonically) preordered space $\mathcal{F}(V_1; V_2)$.

Another important observation about the convex cone of positive linear transformations is that it does *not necessarily* preserve the $*$-preordered space structure: For arbitrary $*$-preordered spaces V_1 and V_2, the $*$-vector space $\mathcal{L}(V_1; V_2)$ is also a preordered vector space, the complex conjugation $\Theta \mapsto \Theta^*$ of which is defined by

$$\Theta^*(v_1) \doteq \Theta(v_1^*)^* \ , \qquad v_1 \in V_1 \ ,$$

for all $\Theta \in \mathcal{L}(V_1; V_2)$. See appendix for more details. It turns out that the $*$-vector space $\mathcal{L}(V_1; V_2)$ may fail to be a $*$-preordered vector space. In other words, it may happen that

$$\mathcal{L}^+(V_1; V_2) \nsubseteq \mathrm{Re}\{\mathcal{L}(V_1; V_2)\} \ .$$

This problem even occurs for linear functionals, i.e., for $V_2 = \mathbb{C}$. To see this, take, for instance, $V_1 = \mathbb{C}$ with $V_1^+ = \{0\}$. In this case $\mathcal{L}^+(V_1; V_2) = \mathcal{L}(V_1; V_2)$ but, of course, in general, $\mathrm{Re}\{\mathcal{L}(V_1; V_2)\} \neq \mathcal{L}(V_1; V_2)$. This "pathology" does not take place as soon as V_1^+ is generating for V_1, meaning that any element of V_1 is a linear combination of elements of V_1^+.

[3] That is, the vector space of all linear transformations from the vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ to the field itself $\mathbb{K}$.

Exercise 1.8 Let V_1 and V_2 be any two $*$-preordered vector spaces. Prove that

$$\mathcal{L}^+(V_1; V_2) \subseteq \mathrm{Re}\{\mathcal{L}(V_1; V_2)\}$$

when V_1^+ is generating for V_1.

We will see later on that this is the case for any C^*-algebra. In particular, positive functionals on such algebras are always Hermitian. This property is important for the (mathematical) theory of quantum states.

The following definition generalizes the usual notion of convexity for real-valued functions of one real variable to functions between two preordered vector spaces.

Definition 1.9 Let V_1 and V_2 be two real preordered vector spaces. We say that the function $f : V_1 \to V_2$ is "convex" if, for all $v, v' \in V_1$ and all $\alpha \in [0, 1]$, one has that

$$\alpha f(v) + (1 - \alpha) f(v') \succeq f(\alpha v + (1 - \alpha)v') .$$

Clearly, as in the case of functions $\mathbb{R} \to \mathbb{R}$, if $f : V_1 \to V_2$ is convex in the above sense, then for all $v_1, \ldots, v_n \in V_1$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}_0^+$, $n \in \mathbb{N}$, such that $\alpha_1 + \cdots + \alpha_n = 1$,

$$\alpha_1 f(v_1) + \cdots + \alpha_n f(v_n) \succeq f(\alpha_1 v_1 + \cdots + \alpha_n v_n) .$$

In particular, any $\Theta \in \mathcal{L}(V_1, V_2)$ is a convex function.

We now introduce the notion of isomorphisms of preordered vector spaces and equivalent preordered vector spaces:

Definition 1.10 (Isomorphisms of Preordered Vector Spaces) Let V_1 and V_2 be two preordered vector spaces. We say that the isomorphism of vector spaces $\Theta : V_1 \to V_2$ (i.e., Θ is a linear one-to-one correspondence) is an "isomorphism of preordered vector spaces," when Θ and Θ^{-1} are positive. If there exists an isomorphism of preordered vector spaces $V_1 \to V_2$, we say that V_1 and V_2 are "equivalent preordered vector spaces."

Exercise 1.11 Let V_1 and V_2 be two preordered vector spaces. Show that an isomorphism of vector spaces $\Theta : V_1 \to V_2$ is an isomorphism of preordered vector spaces iff it is order-preserving and order-reflecting.

Note that, for any two fixed preordered vector spaces V_1 and V_2, if there is an injective bipositive linear transformation $\Theta : V_1 \to V_2$, then V_1 is equivalent (as a preordered vector space) to a subspace of V_2, that is, there is a "copy" of V_1 within V_2. In many important cases, such a Θ exists for $V_2 = \mathcal{F}(\Omega; \mathbb{K})$, where $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and Ω is some nonempty set. We will see later on that this is always the case if V_1 is a C^*-algebra and $\mathbb{K} = \mathbb{C}$. In fact, any *separable* C^*-algebra is equivalent, as a preordered vector space, to a subspace of the space of continuous functions on a locally compact metric space. (This space is even compact if the given C^*-algebra is unital.)

Definition 1.12 (Regularly Preordered Spaces) A preordered vector space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ is said to be "order-semisimple" if, for all $v \in V$, $v = 0$ whenever $\varphi(v) = 0$ for all $\varphi \in V'^+ \doteq \mathcal{L}^+(V; \mathbb{K})$. V is called "regularly preordered" if it is Archimedean and order-semisimple.

The order semisimplicity defined above means that positive linear functionals separate points of the given preordered vector space. We will see later on that any C^*-algebra is regularly ordered. In particular they are order-semisimple. From the side of quantum physics, this corresponds to the fact that any two different observables of a given physical system can be distinguished by some possible state of the system.

One important property of any order-semisimple preordered vector space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ is the existence of a canonical *injective* morphism of preordered vector spaces $\Pi : V \to \mathcal{F}(V'^+; \mathbb{K})$, defined as follows:

$$[\Pi(v)](\varphi) \doteq \varphi(v) , \qquad \varphi \in V'^+ , \ v \in V .$$

It is important to notice, however, that Π is not necessarily bipositive. Recall that V'^+ is the dual positive cone of V. In some cases, Π is also bipositive, like, for instance, if V is any C^*-algebra.

Definition 1.13 (Directed Sets) Let $(I, \succeq)$ be any preordered set. We say that I is an "upward directed set" if any two elements of I have an upper bound in I, i.e., for all $i, i' \in I$, there is $i'' \in I$ such that $i'' \succeq i$ and $i'' \succeq i'$. Similarly, if any two elements of I have a lower bound, then I is said to be a "downward directed set." As the first case is the most relevant here, for simplicity, we will refer to upward directed set as "directed sets." Finally, we say that a preordered vector space is "directed," or a "directed vector space," whenever it is upward directed as a preordered set.

One simple, though important, example of a directed set is given by $(\mathbb{N}, \geq)$, the natural numbers with their usual order relation. Another example is given by the positive cone of any preordered vector space.

Exercise 1.14 Let V be any preordered *real* vector space. Show that V^+ is generating for V iff V is directed and that V is directed if it has an order unit. If V is a $*$-preordered vector space, show that V^+ is generating for V iff $\mathrm{Re}\{V\}$ is directed.

Recall that sequences in a given nonempty set Ω are, by definition, arbitrary functions $\mathbb{N} \to \Omega$. We introduce below a notion which generalizes the one of sequences, by replacing $\mathbb{N}$ with any directed set.

Definition 1.15 (Nets and Monotone Nets) Let Ω be any nonempty set. A "net" or "Moore-Smith sequence" in Ω is, by definition, a mapping $I \to \Omega$, whose domain I is a directed set. For any given directed set I, $(p_i)_{i \in I}$ denotes the mapping (i.e., the net in Ω) $i \mapsto p_i$, from I to Ω. Let $(v_i)_{i\in I}$ be a net in a preordered vector space $(V, \succeq)$. We say that it is "monotonically increasing" if $v_i \succeq v_{i'}$ for all $i, i' \in I$ so that $i \succeq i'$. Similarly, it is "monotonically decreasing" if $v_{i'} \succeq v_i$ for all $i, i' \in I$ so that $i \succeq i'$.

In fact, one may see monotonically increasing nets in a preordered vector space as (upward) directed subsets in this space and, similarly, monotonically decreasing nets as downward directed subsets.

Definition 1.16 (Order Boundedness, Infimum, and Supremum) If Ω is a subset of a preordered vector space $(V, \succeq)$ and $v \in V$ satisfies $v \succeq v'$ for all $v' \in \Omega$, we say that v is an "upper bound" of Ω. In this case we say that Ω is "(order-)bounded from above." The upper bound v is a "supremum" of Ω if, for any second upper bound $\tilde{v}$ of Ω, one has $\tilde{v} \succeq v$, that is, if v is a least upper bound for Ω. Similarly, if $v' \succeq v$ for all $v' \in \Omega$, v is a "lower bound" of Ω and, in this case, Ω is "(order-)bounded from below." The lower bound v is called an "infimum" of Ω if, for any second lower bound $\tilde{v}$ of Ω, one has $v \succeq \tilde{v}$, that is, if v is a largest lower bound for Ω. We say that Ω is "order-bounded," whenever it is order-bounded simultaneously from above and below, i.e., Ω lies in some interval.

Observe that a given order-bounded subset of a preordered vector space V does not necessarily has an infimum or a supremum. Moreover, infima and suprema in preordered vector spaces (like in any preordered set) are not necessarily unique, when they exist.

Definition 1.17 (Monotone Order Limits) If $(v_i)_{i\in I}$ is a monotonically increasing net in a preordered vector space V and $v \in V$ is a supremum for (the upward directed subset) $\{v_i : i \in I\} \subseteq V$, we write $v_i \uparrow v$. Similarly, we write $v_i \downarrow v$ if $(v_i)_{i\in I}$ is a monotonically decreasing net in V and v is an infimum for (the downward directed subset) $\{v_i : i \in I\}$. The vector v is called a "(monotone) order limit" of the monotone net $(v_i)_{i\in I}$, in these two cases.

Observe that if $(v_i)_{i\in I}$ is a monotonically increasing (decreasing) net in V with $v_i \uparrow v$ ($v_i \downarrow v$) for some $v \in V$, then $(-v_i)_{i\in I}$ is a monotonically decreasing (increasing) net and one has $(-v_i) \downarrow (-v)$ ($(-v_i) \uparrow (-v)$).

Exercise 1.18 Let V be any preordered vector space and let $(v_i)_{i\in I}$ and $(v_i')_{i\in I}$ be any two monotonically increasing (decreasing) nets in V with $v_i \uparrow v$, $v_i' \uparrow v'$ ($v_i \downarrow v$, $v_i' \downarrow v'$) for some $v, v' \in V$. Show that, for any constants $\alpha, \alpha' \geq 0$, $(\alpha v_i + \alpha' v_i') \uparrow (\alpha v + \alpha' v')$ ($(\alpha v_i + \alpha' v_i') \downarrow (\alpha v + \alpha' v')$).

Definition 1.19 (Order-Dense Subspaces) Let V be any preordered vector space. We say that the vector subspace $W \subseteq V$ is "order-dense" ("σ-order-dense") if every element of V is a monotone limit of some net (sequence) of elements of W.

Definition 1.20 (Order Completeness and Closedness) We say that the preordered vector space V is "order-complete" or "Dedekind complete" ("σ-order-complete") if every subset (countable subset) of V which is bounded from above has a supremum. (Equivalently, every subset (countable subset) of V which is bounded from below has infimum.) Finally, we say that a vector subspace $W \subseteq V$ is "order-closed" ("σ-order-closed") if the suprema in V of subsets (countable subset) of W are again elements of W. (Equivalently, infima in V of subsets (countable subset) of W are elements of W.) In particular, such suprema (infima) in V are also suprema (infima) in W.

Notice that an order-closed (σ-order-closed) subspace of an order-complete (σ-order-complete) preordered space is again order-complete (σ-order-complete). However, an order-complete (σ-order-complete) subspace of a preordered vector space may fail to be order-closed (σ-order-closed) in this space. By the definition of order-closed (σ-order-closed) subspaces, observe also that a subset (countable subset) $\Omega \subseteq W$ of an *order-closed* (*σ-order-closed*) subspace $W \subseteq V$ having a supremum in V is necessarily bounded within W (and not only for the whole space V). In particular, if V is order-complete (σ-order-complete), then the subset (countable subset) $\Omega \subseteq W$ is bounded from above (below) within W iff it is bounded from above (below) within V. This is generally not true for subspaces that are not order-closed (σ-order-closed), i.e., it may happen for general subspaces $W \subseteq V$ that some subset (countable subset) $\Omega \subseteq W$ has indeed a supremum (infimum) in V, but is not even bounded from above (below) within W.

The preordered vector space $\mathcal{F}(\Omega; V)$ is order-complete (σ-order-complete) iff V is order-complete (σ-order-complete). In fact, in this case, for any subset (countable subset) $S \subseteq \mathcal{F}(\Omega; V)$ that is bounded from above, one has that the mapping $\Omega \to V$:

$$p \mapsto \sup\{f(p) \ : \ f \in S\} ,$$

where, for any $p \in \Omega$, $\sup\{f(p) : f \in S\}$ stands for any supremum in V of the subset $\{f(p) : f \in S\} \subseteq V$, is a supremum in $\mathcal{F}(\Omega; V)$ for S.

Exercise 1.21 Prove that any order ideal of an order-complete (σ-order-complete) preordered vector space is also order-complete (σ-order-complete). Prove the same for order-closed (σ-order-closed) subspaces.

Observe that an arbitrary order ideal of a preordered vector space is not necessarily order-closed (σ-order-closed). The next exercise gives some intuition on the nature of such order ideals.

Exercise 1.22 Let M be any metric space. For any nonempty subset $\Omega \subseteq M$, define the subspace $C_\Omega(M; \mathbb{R})$ of $C(M; \mathbb{R})$ by

$$C_\Omega(M; \mathbb{R}) \doteq C(M; \mathbb{R}) \cap \mathcal{F}_\Omega(M; \mathbb{R}) = \{f \in C(M; \mathbb{R}) \ : \ f(\Omega) = \{0\}\} .$$

Show that such subspaces are order ideals of $C(M; \mathbb{R})$ but generally not σ-order-closed. Prove additionally that $C_\Omega(M; \mathbb{R})$ is order-closed whenever Ω is an *open* subset.

The notion of order boundedness of subsets of preordered vector spaces naturally induces a notion of order boundedness for linear transformations between preordered vector spaces:

Definition 1.23 (Order-Bounded Linear Transformations) Let V_1 and V_2 be any two preordered vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We say that the linear transformation $\Theta \in \mathcal{L}(V_1; V_2)$ is "order-bounded" if it maps any order bounded subset of V_1 into an order-bounded subset of V_2. The family of all order-bounded linear transformations $V_1 \to V_2$ is denoted by $\mathcal{L}_{\text{ob}}(V_1; V_2)$.

Remark that $\mathcal{L}^+(V_1; V_2) \subseteq \mathcal{L}_{\text{ob}}(V_1; V_2)$, i.e., any positive linear transformation $V_1 \to V_2$ is order-bounded, because such mappings preserve order.

Exercise 1.24 Let V_1 and V_2 be any two preordered vector spaces. Show that $\mathcal{L}_{\text{ob}}(V_1; V_2)$ is a *real* order ideal of $\mathcal{L}(V_1; V_2)$. In the case of V_1 and V_2 being $*$-preordered vector spaces, show additionally that $\mathcal{L}_{\text{ob}}(V_1; V_2)$ is a self-conjugate subspace of $\mathcal{L}(V_1; V_2)$, i.e.,

$$\mathcal{L}_{\text{ob}}(V_1; V_2)^* \doteq \{\Theta^* : \Theta \in \mathcal{L}_{\text{ob}}(V_1; V_2)\} = \mathcal{L}_{\text{ob}}(V_1; V_2) .$$

Definition 1.25 (Order Dual) Let V be any preordered vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

$$V^{\text{od}} \doteq \text{span}(\mathcal{L}^+(V; \mathbb{K}))$$

is called the "order dual" of V.

If $\mathbb{K} = \mathbb{R}$, by the last exercise, one has $V^{\text{od}} \subseteq \mathcal{L}_{\text{ob}}(V; \mathbb{K})$ but, in general, $V^{\text{od}} \neq \mathcal{L}_{\text{ob}}(V; \mathbb{K})$. The space of order-bounded linear functionals $\mathcal{L}_{\text{ob}}(V; \mathbb{K})$ is sometimes called "order-bounded dual" of V.

Note that the order dual of any (not necessarily real) preordered vector space V is an order ideal of the dual space V'.

Exercise 1.26 Let V be any $*$-preordered vector space. Show that the order dual of V is a self-conjugate subspace of its dual space, V', and that the restriction to $\text{Re}\{V\}$ of mappings $V \to \mathbb{C}$ establishes an one-to-one correspondence from $\text{Re}\{V^{\text{od}}\}$ to $\text{Re}\{V\}^{\text{od}}$.

In the case of normed spaces, there is a tight relation between positivity and order boundedness of linear transformation and their continuity. This will be discussed in some detail later on, in the current chapter. This fact is of particular importance for the theory of states of C^*-algebras, which are the mathematical objects representing physical states of general quantum systems.

Let V be an arbitrary vector space. We say that the cone $P \subseteq V$ is "pointed," whenever

$$(-P) \cap P = [0, 0] = \{0\}$$

(i.e., the only element $v \in V$ such that $-v, v \in P$ is $v = 0$). It is easy to verify that the preorder $\succeq^P$ is a partial order in V, i.e., it is antisymmetric (meaning that, for all $v, v' \in V$, $v \succeq^P v'$ and $v' \succeq^P v$ iff $v = v'$), iff P is pointed. $\mathbb{R}_0^+$ and $C(M; \mathbb{K})^+$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and M being a metric space, are important examples of a pointed cone.

Definition 1.27 (Ordered Vector Spaces) Let V be a preordered vector space. We say that this space is an "ordered" vector space if its positive cone is pointed (and, hence, $\succeq$ is a partial order in V). Note that in ordered vector spaces (like in any partially ordered set), infima and suprema are unique, when they exist. If V is an ordered vector space and $\Omega \subseteq V$ has an infimum, it is denoted by $\inf \Omega$. Analogously, $\sup \Omega$ denotes the supremum of Ω, when it exists. We say that the ordered vector space V is a "$*$-ordered vector space" if it is a $*$-preordered vector space.

$\mathbb{K}$ and $C(M; \mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$, M being any metric space, are important examples of an ordered vector space. $\mathbb{C}$ and $C(M; \mathbb{C})$ are additionally $*$-ordered vector spaces. We will show later on that any unital C^*-algebra is a $*$-ordered vector space. The algebra unit of such an algebra is additionally an order unit. In particular, like $\mathrm{Re}\{C(M; \mathbb{C})\}$, the real subspace of any unital C^*-algebra is directed (because any preordered vector space having order units is automatically directed). Recall that, in mathematical physics, the set of all observables of a given physical system, quantum or classical, is seen as the space of selfadjoint elements of some unital C^*-algebra, which is commutative in the classical and noncommutative in the quantum case.

Exercise 1.28 Let V_1 be any preordered vector space and V_2 any ordered vector space. Show that, for any $\Theta \in \mathcal{L}^+(V_1; V_2)$,

$$\ker(\Theta) \doteq \{v \in V_1 : \Theta(v) = 0\} \subseteq V_1$$

is an order ideal.

Notice that the assertion of the last exercise is generally false if both V_1 and V_2 are only preordered.

Exercise 1.29 Let V_1 be a vector space whose positive cone is generating and V_2 any ordered vector space. Show that $\mathcal{L}(V_1; V_2)$ is again an ordered vector space.

In particular, by the last exercise, the dual space and, hence, the order dual of any directed vector space are always an ordered vector space.

Exercise 1.30 Let V_1 and V_2 be any two ordered vector spaces. Show that a linear transformation $\Theta : V_1 \to V_2$ is an isomorphism of (pre)ordered vector spaces iff it is bipositive and surjective. (In contrast to the (more general) preordered case, we do not have to assume that Θ is bijective.)

In ordered vector spaces, we have the following equivalent formulation for the Archimedean property.

Exercise 1.31 Let V be any ordered vector space. Show that V is Archimedean iff, for all $v \in V^+$, one has

$$\inf\{\alpha v \ : \alpha > 0\} = 0 \ .$$

Deduce from this that any σ-order-complete ordered vector space is Archimedean.

1.1.1 Ordered Normed Spaces

If a given vector space is to be simultaneously equipped with a preorder (positive cone) and a norm[4] (or, more generally, with some vector space topology), one usually requires some compatibility between the norm (or topology) and the preorder relation. We give below the most commonly used of such conditions and briefly discuss some of their important consequences. In particular, we will discuss the important role played by the Archimedean property, the order fullness, and order units.

Definition 1.32 (Preordered Normed Spaces) Let X be a preordered vector space, which is equipped with a norm, i.e., X is also a normed space. We say that X is a “preordered normed space” if its positive cone is closed with respect to the given norm. An ordered vector space which is also a preordered normed space is called here an “ordered normed space.” A (pre)ordered normed space, which is complete with respect to its norm (i.e., it is a Banach space) is called here a “(pre)ordered Banach space.”

For any compact metric space K, it is easy to see that $C(K;\mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$, with the supremum norm:

$$\|f\|_\infty \doteq \sup f(K) \doteq \sup\{f(p) : p \in K\} \ , \qquad f \in C(K;\mathbb{K}) \ ,$$

is an ordered Banach space. This is an important example of such a space. We will proof later on that any C^*-algebra is also an ordered Banach space.

Exercise 1.33 Show that any preordered normed space is Archimedean.

Beyond the Archimedean property of preordered normed spaces, the (norm) closedness of the positive cone yields that positive linear transformations on preordered normed spaces map norm-bounded subsets of the positive cone to order-bounded sets.

---

[4] A norm $v \mapsto \|v\|$ on a vector space V over a field $\mathbb{K}$ is a real-valued function satisfying $\|v + w\| \leq \|v\| + \|w\|$ (the triangle inequality) and $\|\lambda v\| = |\lambda| \, \|v\|$ (homogeneity) for any $v, w \in V$ and $\lambda \in \mathbb{K}$, as well as $\|v\| = 0 \Rightarrow v = 0$ (positive definiteness). See the appendix for more details.

Lemma 1.34 *Let X be any preordered Banach space and V any preordered vector space having an order unit. For any $\Theta \in \mathcal{L}^{+}(X; V)$ and any* norm-*bounded subset $\Omega \subseteq X^{+}$, the image $\Theta(\Omega) \subseteq V^{+}$ is* order-*bounded.*

Proof Take any positive linear transformation $\Theta \in \mathcal{L}^{+}(X; V)$. By linearity of Θ, it suffices to show that, for some $w_{\Theta} \in V^{+}$, one has

$$\Theta(X^{+} \cap B_1(0)) \subseteq [0, w_{\Theta}] ,$$

where

$$B_1(0) \doteq \{x \in X : \|x\| < 1\}$$

is the unit ball (the ball of radius 1 centered at zero) of the normed space X. Fix any order unit u of V, and assume, by contradiction, that there is no such $w_{\Theta} \in V^{+}$. In particular, there is a sequence $x_n \in X^{+} \cap B_1(0)$ such that, for all $n \in \mathbb{N}$, $n^3 u \not\succeq \Theta(x_n)$ (i.e., $n^3 u \succeq \Theta(x_n)$ does not hold true). For X is a Banach space and X^{+} is closed:

$$x \doteq \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n^2} x_n \in X^{+}$$

and $x \succeq \frac{1}{n^2} x_n \succeq 0$ for all $n \in \mathbb{N}$. In particular, as $\Theta \in \mathcal{L}^{+}(X; V)$,

$$n^2 \Theta(x) \succeq n^2 \Theta\left(\frac{1}{n^2} x_n\right) = \Theta(x_n) \in V^{+}$$

for all $n \in \mathbb{N}$. As $n^3 u \not\succeq \Theta(x_n)$, it follows that, for all $n \in \mathbb{N}$,

$$nu \not\succeq \Theta(x) \in V^{+} .$$

But this contradicts the assumption that $u \in V^{+}$ is an order unit of V and the assertion follows. $\square$

This property is behind the fact that, under fairly general conditions, positive linear transformations between preordered normed spaces are automatically continuous. In particular, we have the following corollary.

Corollary 1.35 (Positive Functionals Are Continuous—I) *Let X be any preordered Banach space over $\mathbb{R}$ in which any vector is the difference of two positive vectors whose length is at most the length of the original vector. (This situation indeed occurs in many important cases.) Any positive linear functional is continuous. In particular,*

$$X^{\mathrm{od}} \doteq \mathrm{span}(\mathcal{L}^{+}(X; \mathbb{R})) \subseteq X^{\mathrm{td}} \doteq \mathcal{B}(X; \mathbb{R}) \doteq \left\{\varphi \in \mathcal{L}(X; \mathbb{R}) : \|\varphi\|_{\mathrm{op}} < \infty\right\} .$$

Here,

$$\|\varphi\|_{\text{op}} \doteq \sup_{x \in X,\ \|x\|=1} |\varphi(x)| \in [0, \infty]$$

is the "operator norm" of the linear functional $\varphi \in \mathcal{L}(X; \mathbb{R})$. *See the appendix for more details.*

In fact, the assumption of the last corollary, on the decomposition of vectors, can be weakened, and the following stronger result also holds true:

Theorem 1.36 (Positive Functionals Are Continuous—II) *Let X be any directed real Banach space (i.e., X is a preordered real Banach space in which any vector is the difference of two positive vectors). Then,* $X^{\text{od}} \subseteq X^{\text{td}}$.

In the following we discuss the so-called order fullness of norms in a preordered vector space. This property has very important consequences, as it will be shown.

Definition 1.37 (Normal Positive Cones) Let X be a *real* preordered vector space, which is equipped with a norm $\|\cdot\|$. We say that the positive cone X^+ is "normal" with respect to this norm if, for any $R > 0$, there is $r > 0$ and an *order full* subset $\Omega \subseteq X$ such that

$$B_r(0) \doteq \{x \in X : \|x\| < r\} \subseteq \Omega \subseteq B_R(0) \doteq \{x \in X : \|x\| < R\} \ .$$

Normed spaces that are preordered vector spaces and whose positive cones are normal are called "locally full." Consistently, if $\|\cdot\|$ is a norm in X, with respect to which X^+ is normal, then the norm is called "locally full." As is usual, we extend these definitions to the case of $*$-preordered vector spaces by referring the above properties to the real subspace $\text{Re}\{X\} \doteq \{x \in X : x^* = x\}$.

Note the canonical positive cone $\mathbb{R}_0^+$ of $\mathbb{R}$ is normal with respect to the absolute value. The same is true for $\mathbb{R}_0^+$ being the positive cone of $\mathbb{C}$ as a $*$-preordered vector space.

One important example of a preordered vector space having a normal positive cone is $C(K; \mathbb{R})$, K being a compact metric space, endowed with its supremum norm $\|\cdot\|_\infty$: For any fixed $R > 0$,

$$B_{R/2}(0) \subseteq [-R/2, R/2] \subseteq B_R(0) \ ,$$

where $[-R/2, R/2]$ is the interval in $C(K; \mathbb{R})$, from the constant function $-R/2$ to the constant function $R/2$. Another example of this situation is the (real) vector space of all selfadjoint elements of arbitrary unital C^*-algebras, as we will prove at the appropriate point. In particular, any such an algebra is locally full.

A crucial property of a preordered vector space, whose positive cone is normal with respect to a norm, is the following.

Exercise 1.38 Let X be a real preordered vector space, which is equipped with a norm. Show that if X^+ is normal with respect to this norm, then any order-bounded subset of X is also norm-bounded.

We will see below that, for ordered vector spaces with order units, one can always choose locally full norms in a way that order boundedness and norm boundedness of subsets are equivalent properties. It will be shown later on that the norm of any unital-C^* algebra is of this kind.

In the following we discuss one important class of locally full norms.

Definition 1.39 (Monotone Norms) Let X be any preordered vector space and let $\|\cdot\|$ be a (semi)norm[5] in X. We say that $\|\cdot\|$ is a "monotone (semi)norm" in the preordered space if, for all $x, x' \in X^+$ so that $x \succeq x'$, one has $\|x\| \geq \|x'\|$. The (semi)norm is said to be "fully monotone" if, for all $x, x', x'' \in X$ satisfying $x \succeq x'' \succeq x'$, one has

$$\|x''\| \leq \max\{\|x\|, \|x'\|\},$$

i.e., in any interval $[x', x] \subseteq X$, the (semi)norm is maximized at the extreme points x', x of the interval.

Observe that any fully monotone norm is monotone, but the converse is false.

Exercise 1.40 Let X be any preordered vector space and let $\|\cdot\|$ be a norm in X.

(i) Show that the norm is fully monotone iff the corresponding closed (open) unit ball

$$\overline{B}_1(0) \doteq \{x \in X : \|x\| \leq 1\} \subseteq X$$

($B_1(0) \subseteq X$) is order-full.

(ii) Show that the norm is locally full if it is equivalent[6] to a fully monotone one. In particular, any fully monotone norm is locally full.

In fact, it can be shown that the second part of the last exercise refers to an equivalence: A norm in X is locally full iff it is equivalent to a fully monotone one. In particular, regarding topological properties (like continuity), it suffices to consider fully monotone norms, instead of all locally full ones.

[5] A seminorm $v \mapsto \|v\|$ on a vector space V is a real-valued function satisfying all properties of a norm except the positive definiteness, i.e., $\|v\| = 0$ does not yield $v = 0$. See the appendix for more details.

[6] $\|\cdot\|^{(1)}$ and $\|\cdot\|^{(2)}$ are "equivalent norms" if, for some $C \in [1, \infty)$. and all $v \in V$, $C^{-1}\|v\|^{(2)} \leq \|v\|^{(1)} \leq C\|v\|^{(2)}$. See the appendix for more details.

Definition 1.41 (Seminorm of an Order Unit) Let V be any real preordered vector space and let $u \in V$ be an order unit. We define the mapping $\|\cdot\|_u : V \to \mathbb{R}_0^+$ by

$$\|v\|_u \doteq \inf\{\alpha > 0 \ : \ v \in [-\alpha u, \alpha u]\} , \qquad v \in V .$$

Observe that, by the definition of order units, the interval $[-u, u]$ is absorbing.[7] Moreover, it is balanced and convex. In fact, $\|\cdot\|_u$ is nothing else than the so-called Minkowski functional associated with $[-u, u] \subseteq V$.

Exercise 1.42 Prove that such a mapping $\|\cdot\|_u : V \to \mathbb{R}_0^+$ is a fully monotone seminorm.

Proposition 1.43 *Let V be any* Archimedean *real ordered vector space and let $u \in V$ be an order unit. $\|\cdot\|_u$ is a fully monotone norm in V and $V^+ \subseteq V$ is closed with respect to this norm. In particular, $(V, \|\cdot\|_u)$ is a preordered normed space. For all $v \in V$,*

$$v \in [-\|v\|_u \, u, \|v\|_u \, u] .$$

The interval $[-u, u] \subseteq V$ is the closed unit ball $\overline{B}_1(0)$ of the normed space $(V, \|\cdot\|_u)$. In particular, any subset $\Omega \subseteq V$ is norm-bounded (with respect to $\|\cdot\|_u$) iff it is order-bounded.

Proof

1. By the last exercise, $\|\cdot\|_u$ is a seminorm. Take any $v \in V$ and assume that $\|v\|_u = 0$. Then, by the definition of the seminorm $\|\cdot\|_u$, for all $\alpha > 0$,

   $$v \in [-\alpha^{-1}u, \alpha^{-1}u] .$$

   That is, for all $\alpha \geq 0$,

   $$u \succeq \alpha(-v) \qquad \text{and} \qquad u \succeq \alpha v ,$$

   where we used for $\alpha = 0$ that $u \in V^+$, by definition. Thus, from the Archimedean property, $0 \succeq -v$ and $0 \succeq v$. As V is an ordered vector space, it follow that $v = 0$. That is, the seminorm μ_u is actually a norm. Again by the exercise, it is fully monotone.
2. Take any $v \in V$. Then, similarly as above, by the definition of $\|\cdot\|_u$, for all $\alpha > 0$,

   $$v \in [-(\|v\|_u + \alpha^{-1})u, (\|v\|_u + \alpha^{-1})u] .$$

[7] A absorbing subset of a vector space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ is a set $W \subseteq V$ such that for any $v \in V$, there is $r > 0$ so that $v \in \lambda W$ for any $\lambda \in \mathbb{K}$ with $|\lambda| > r$.

Hence, for all $\alpha \geq 0$,

$$u \succeq \alpha(-v - \|v\|_u u) \qquad \text{and} \qquad u \succeq \alpha(v - \|v\|_u u) ,$$

keeping in mind that $u \in V^+$, by definition. Again by the Archimedean property, we arrive at

$$0 \succeq -v - \|v\|_u u \qquad \text{and} \qquad 0 \succeq v - \|v\|_u u .$$

Hence, $v \succeq - \|v\|_u u$ and $\|v\|_u u \succeq v$. That is, for all $v \in V$,

$$v \in [- \|v\|_u u, \|v\|_u u] .$$

3. By the definition of the norm $\|\cdot\|_u$, one clearly has that $[-u, u] \subseteq \overline{B}_1(0)$, where $\overline{B}_1(0)$ is the unit closed ball centered at 0 (for the norm $\|\cdot\|_u$). By 2.,

$$\overline{B}_1(0) \subseteq \bigcup_{v \in V \ , \ \|v\|_u \leq 1} [- \|v\|_u u, \|v\|_u u] = [-u, u] .$$

Note here that $[-\alpha u, \alpha u] \subseteq [-u, u]$ for every $\alpha \in [0, 1]$.
4. Recall that any fully monotone norm is locally full and that order-bounded subsets are also norm-bounded with respect to such norms. Thus, from 3., we conclude that any $\Omega \subseteq V$ is norm-bounded with respect to $\|\cdot\|_u$ iff it is order-bounded.
5. Finally, to prove the closedness of the positive cone $V^+ \subseteq V$ with respect to the norm $\|\cdot\|_u$, let $(v_n)_{n\in\mathbb{N}}$ be any sequence of *positive* vectors (i.e., $v_n \in V^+$) converging (in V) to some $v \in V$. As $[-u, u] = \overline{B}_1(0)$, for any $k \in \mathbb{N}$, we can choose $n_k \in \mathbb{N}$ such that

$$\frac{1}{k} u \succeq v - v_{n_k} \succeq -\frac{1}{k} u .$$

In particular, one has

$$v = (v - v_{n_k}) + v_{n_k} \succeq v - v_{n_k} \succeq -\frac{1}{k} u .$$

Hence, for all $k \in \mathbb{N}$, $u \succeq k(-v)$. By the Archimedean property, $0 \succeq -v$, that is, $v \in V^+$.

□

Note that 1 is an order unit for $(\mathbb{R}, \geq)$ with $\|\cdot\|_1$ being nothing else than the absolute value $|\cdot|$. Similarly, for any compact metric space K, the constant function 1 is an order unit for $C(K; \mathbb{R})$ and $\|\cdot\|_1$ is exactly the supremum norm $\|\cdot\|_\infty$. We will prove later on that, more generally, for any unital C^*-algebra $\|\cdot\|_{\mathbb{1}}$, where $\mathbb{1}$

is the (algebraic) unit of the algebra (which is, simultaneously, also an order unit), coincides on selfadjoint elements with the original norm of the algebra.

Exercise 1.44 Let V be any *Archimedean* real ordered vector space and let $u, u' \in V$ be any two order units. Prove that $\|\cdot\|_u$ and $\|\cdot\|_{u'}$ are equivalent norms.

From the last exercise, Archimedean real ordered vector spaces having order units have a canonical norm topology, the one associated with any (order) unit of the space. By default, this will be the (norm) topology associated with such a space.

We summarize the relations between different types of norms for preordered vector spaces, presented above:

(i) Any fully monotone norm is locally full.
(ii) Locally full norms are exactly the norms that are equivalent to some fully monotone one.
(iii) For each real ordered space that is Archimedean and has order units, there are canonical fully monotone norms associated with each order unit of the space. All such norms are equivalent to each other. In particular, these spaces have a canonical norm topology. Moreover, in this topology the positive cone is closed, that is, such a space together with a norm associated with an arbitrary order unit is automatically an ordered normed space.

Given two preordered vector spaces V_1 and V_2 over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, recall that $\mathcal{L}_{\text{ob}}(V_1; V_2)$ denotes the family of all order-bounded linear transformations $V_1 \to V_2$.

Corollary 1.45 *Let V be any* Archimedean *real ordered vector space having an order unit and X any real preordered vector space equipped with a locally full norm. Then,*

$$\mathcal{L}_{\text{ob}}(V; X) \subseteq \mathcal{B}(V; X) \doteq \left\{\Theta \in \mathcal{L}(V; X) : \|\Theta\|_{\text{op}} < \infty\right\} ,$$

i.e., any order-bounded linear transformation $V \to X$ is continuous. Here,

$$\|\Theta\|_{\text{op}} \doteq \sup_{v \in V,\ \|v\|_u = 1} \|\Theta(v)\| \in [0, \infty]$$

is the "operator norm" of the linear mapping $\Theta \in \mathcal{L}(V; X)$, where u is any order unit of V. See the appendix for more details on the operator norms. In particular, morphisms $V \to X$ of such preordered vector spaces (i.e., positive linear transformations; see Definition 1.7) are automatically continuous. Thus, $\mathcal{B}(V; X)$ is an order ideal of $\mathcal{L}(V; X)$, in this case. (Recall that $\mathcal{L}^+(V; X) \subseteq \mathcal{L}_{\text{ob}}(V; X)$.) If X is an Archimedean *ordered vector space and its norm is the one associated with some order unit, then*

$$\mathcal{L}_{\text{ob}}(V; X) = \mathcal{B}(V; X) .$$

Proof Exercise. □

Another very important property of Archimedean real ordered vector space having an order unit is the fact that the order and topological duals of such spaces coincide. We will demonstrate this by using the following well-known result of the theory of topological vector spaces.

Theorem 1.46 (Grosberg-Krein) *Let X be a preordered vector space over $\mathbb{R}$, which is endowed with a norm. Assume that, for all $R > 0$, there is an interval $[x, x'] \subseteq X$ and some $r > 0$ such that*

$$B_r(0) \subseteq [x, x'] \subseteq B_R(0) \ .$$

Then, for all $\varphi \in X^{\mathrm{td}} \doteq \mathcal{B}(X; \mathbb{R})$, there are

$$\varphi^-, \varphi^+ \in \mathcal{B}(X; \mathbb{R})^+ \doteq \{\varphi \in \mathcal{B}(X; \mathbb{R}) : \varphi \geq 0\}$$

such that

$$\varphi = \varphi^+ - \varphi^- \ .$$

Observe that the assumptions of the above theorem imply that X is locally full. The proof of this theorem, in its general form, uses the Hahn-Banach separation theorem ([18, 3.4 Theorem], or Theorem 7.331 for a simpler version of it) and will not be presented here. For a reference, see [101].

Proposition 1.47 (Positive Functionals Are Continuous—III) *If V is an Archimedean real ordered space having an order unit, then*

$$V^{\mathrm{td}} \doteq \mathcal{B}(V; \mathbb{R}) = V^{\mathrm{od}} \doteq \mathrm{span}(\mathcal{L}^+(V; \mathbb{R})) = \mathcal{L}_{\mathrm{ob}}(V; \mathbb{R}) \ .$$

In other words, the topological dual, the order dual, as well as the order- bounded dual of V are all the same vector space. If V is an Archimedean $*$-ordered space having an order unit, then

$$V^{\mathrm{td}} \doteq \mathcal{B}(V; \mathbb{C}) = V^{\mathrm{od}} \doteq \mathrm{span}(\mathcal{L}^+(V; \mathbb{C})) \ .$$

Proof Let first V be a real vector space. Recall that the inclusion $V^{\mathrm{od}} \subseteq \mathcal{L}_{\mathrm{ob}}(V; \mathbb{R})$ always hold, in this case. For $\mathbb{R}$ is an Archimedean ordered space having an order unit, by the last corollary we also have the identity $\mathcal{L}_{\mathrm{ob}}(V; \mathbb{R}) = V^{\mathrm{td}}$. Finally, by the Grosberg-Krein theorem combined with the last proposition, from the assumptions, it follows that $V^{\mathrm{td}} \subseteq V^{\mathrm{od}}$. The case of $*$-preordered vector spaces follows from the real one, by observing that V^{td} and V^{od} are self-conjugate subspaces of the dual space of V and decomposing elements of these spaces in their real and imaginary parts. See also Exercises 1.26 and 7.77. □

To conclude the paragraph, we give two direct, albeit nontrivial, applications of the last proposition.

Corollary 1.48 *Any Archimedean real ordered space V having an order unit is order-semisimple, i.e., for all $v \in V$, $v = 0$ whenever $\varphi(v) = 0$ for all $\varphi \in V'^+ \doteq \mathcal{L}^+(V;\mathbb{R})$.*

Proof Exercise. *Hint*: Use Corollary 7.41. □

Corollary 1.49 (Riesz-Markov-Kakutani for Compact Spaces) *Let K be any compact metric space. A linear functional $\varphi : C(K;\mathbb{R}) \to \mathbb{R}$ is continuous with respect to the supremum norm iff it is the difference of two (finite positive) Borel measures[8] on K.*

Proof Clearly, the difference of two (finite positive) Borel measures always defines a bounded (and hence continuous) linear functional $C(K;\mathbb{R}) \to \mathbb{R}$. Recall that the supremum norm in $C(K;\mathbb{R})$ is a unit norm and that $C(K;\mathbb{R})$ is Archimedean. By the last proposition, for any continuous linear functional $\varphi : C(K;\mathbb{R}) \to \mathbb{R}$, there are $\varphi^+, \varphi^- \in C(K;\mathbb{R})'^+$ such that $\varphi = \varphi^+ - \varphi^-$. By the Riesz-Markov theorem (Theorem 4.68), any positive linear functional $C(K;\mathbb{R}) \to \mathbb{R}$ corresponds to a (unique) Borel measure. □

In fact, notice that any (finite positive) Borel measure on K, an arbitrary compact metric space, is uniquely identified with a linear functional on the bounded Borel-measurable functions of K, which is an "integral" in the sense of the theory of Riesz spaces, as defined in Sect. 7.4.5. See Definition 4.10 and remarks following it.

[8] See Definition 4.10 and remarks following it.

Chapter 2
The Space of Bounded Operators on a Hilbert Space as Ordered Vector Space

2.1 The Positive Cone of Bounded Operators on a Hilbert Space

In the present section, we define the cone of positive bounded operators on a Hilbert space and study its main properties, as well as those of the corresponding dual positive cone, i.e., the cone of positive functionals acting on the space of bounded operators on the Hilbert space.

In fact, the scalar product $\langle \cdot, \cdot \rangle$ of any Hilbert space H naturally induces a preorder in the space:

$$\mathcal{B}(H) \doteq \left\{ A \in \mathcal{L}(H) : \ \|A\|_{\mathrm{op}} \doteq \sup \{\|A\varphi\| : \varphi \in H, \ \|\varphi\| = 1\} < \infty \right\}$$

of bounded (linear) operators on H, $\|\psi\| \doteq \langle \psi, \psi \rangle^{1/2}$, $\psi \in H$, being the norm of H.

Definition 2.1 (**$\mathcal{B}(H)$ as a Preordered Vector Space**) Let H be any (real or complex) Hilbert space and define the convex cone:

$$\begin{aligned} \mathcal{B}^+(H) &\doteq \{A \in \mathcal{B}(H) \ : \ \langle x, A(x) \rangle \geq 0 \text{ for all } x \in H\} \\ &= \{A \in \mathcal{B}(H) \ : \ N(A) \subseteq \mathbb{R}_0^+\} , \end{aligned}$$

where

$$N(A) \doteq \{\langle x, A(x) \rangle : x \in H, \|x\| = 1\}$$

denotes the "numerical range" of A. $\mathcal{B}(H)$ is canonically seen as the preordered vector space whose positive cone is $\mathcal{B}^+(H)$. The corresponding (pre)order relation is denoted here by $\geq$.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_2

Recall that the "spectrum" $\sigma(A)$ of $A \in \mathcal{B}(H)$ is the complement in the field $\mathbb{K} = \mathbb{R}, \mathbb{C}$ of its resolvent set $R(A)$, i.e., $\sigma(A) \doteq \mathbb{K} \backslash R(A)$ with

$$R(A) \doteq \{z \in \mathbb{K} \ : \ (A - z\mathrm{id}_H) \text{ has an inverse in } \mathcal{B}(H)\} \subseteq \mathbb{K} ,$$

$\mathrm{id}_H \in \mathcal{B}(H)$ being the identity operator. For more details, see the appendix. By Proposition 7.242, $\sigma(A) \subseteq \overline{N(A)}$ for all $A \in \mathcal{B}(H)$ and, hence, any positive $A \in \mathcal{B}(H)$ has necessarily a positive spectrum, i.e., $\sigma(A) \subseteq \mathbb{R}_0^+$. Note, however, that the converse of this property is not true, i.e., there may be an element $A \in \mathcal{B}(H)$ not lying in the positive cone $\mathcal{B}(H)^+$, whose spectrum is indeed positive. In fact, this can only occur if A is not selfadjoint, and we have the following corollary of Proposition 7.244, regarding the positivity of selfadjoint operators.

Corollary 2.2 *Let H be any (real or complex) Hilbert space. A* selfadjoint *operator $A \in \mathcal{B}(H)$ is positive iff $\sigma(A) \subseteq \mathbb{R}_0^+$.*

Observe that the above notion of positivity of elements of $\mathcal{B}(H)$ has no relation at all to the positivity-preserving property of Definition 1.7, for H is not assumed to be a preordered vector space. In specific cases where the Hilbert space H is itself preordered and both notions of positivity are available, in order to avoid confusion, we will use the term "positivity preserving" to refer to the positivity of a linear transformation in the sense of Definition 1.7.

The identity operator $\mathrm{id}_H \in \mathcal{B}(H)$ is trivially positive, by the positivity of scalar products (i.e., by $\langle x, x\rangle \geq 0$). Similarly, for any fixed vector $x \in H$, $A \mapsto \langle x, A(x)\rangle$ is (trivially) a positive linear functional on $\mathcal{B}(H)$. For all $A \in \mathcal{B}(H)$, the operator $A^*A \in \mathcal{B}(H)$ is positive, for

$$\left\langle x, A^*A(x)\right\rangle = \langle A(x), A(x)\rangle \geq 0 , \qquad x \in H .$$

In fact, we will see later on that any positive element of $\mathcal{B}(H)$ is of this form. Recall that the orthogonal projectors on a given Hilbert space H are exactly the selfadjoint projectors in $\mathcal{B}(H)$, i.e., those $P \in \mathcal{B}(H)$ for which $P^2 = P = P^*$. Hence, $P = P^*P$ and any such a projector is positive.

Exercise 2.3 Let H be any Hilbert space. Show that the positive cone of $\mathcal{B}(H)$ is closed with respect to the operator norm:

$$\|A\|_{\mathrm{op}} \doteq \sup\{\|A\varphi\| : \varphi \in H,\ \|\varphi\| = 1\} , \qquad A \in \mathcal{B}(H) .$$

In particular, since any preordered normed space (Definition 1.32) is Archimedean (Exercise 1.33), $\mathcal{B}(H)$ is an Archimedean preordered vector space. It turns out that, in the complex case, $\mathcal{B}(H)$ is a $*$-ordered (and not only ($*$-)preordered) vector space.

Exercise 2.4 Let H be any *complex* Hilbert space. Show that

$$\mathcal{B}(H)^+ \subseteq \mathrm{Re}\{\mathcal{B}(H)\} \doteq \{A \in \mathcal{B}(H) \ : \ A = A^*\} \doteq \mathcal{B}(H)^{\mathbb{R}} ,$$

i.e., $\mathcal{B}(H)$ is a $*$-preordered vector space, and $\mathcal{B}(H)^+$ is pointed, i.e., $\mathcal{B}(H)$ is an ordered vector space. Show additionally that $\mathrm{id}_H \in \mathcal{B}(H)^+$ is an order unit of $\mathcal{B}(H)$ and its associated norm in $\mathrm{Re}\{\mathcal{B}(H)\}$ is nothing else than the operator norm.

Hint: Use Proposition 7.233 and Corollary 7.234.

Combining the first part of the last exercise with Corollary 2.2, for any complex Hilbert space H, we arrive at the following characterization of the positive cone $\mathcal{B}(H)^+$:

Corollary 2.5 *Let H be any* complex *Hilbert space. $A \in \mathcal{B}(H)$ is positive iff it is selfadjoint and its spectrum is positive.*

Observe that the above corollary can be used, in the complex case, to equivalently define the positive cone $\mathcal{B}(H)^+$ only by means of the algebraic structure of $\mathcal{B}(H)$, as the selfadjointness and the spectrum of an operator are purely algebraic notions. For this reason, the positivity of the spectrum and selfadjointness are used to define positivity of elements of general $*$-algebras. In C^*-algebras, which are normed algebras, the positivity of elements can, additionally, be characterized via a condition relative to the norm. This will also be shown below for $\mathcal{B}(H)$.

By the existence of an order unit, we conclude that $\mathrm{Re}\{\mathcal{B}(H)\} \doteq \mathcal{B}(H)^{\mathbb{R}}$ is directed and, hence, as $\mathcal{B}(H)$ is a $*$-preordered vector, $\mathcal{B}(H)^+$ is generating for $\mathcal{B}(H)$ (see Exercise 1.14). In fact, we have the following property regarding the decomposition of arbitrary elements of $\mathcal{B}(H)$ as linear combinations of positive ones.

Proposition 2.6 (Positive Decomposition of Operators) *Let H be any complex Hilbert space. For all $A \in \mathcal{B}(H)$, there are $A_{\mathrm{Re}}^+, A_{\mathrm{Re}}^-, A_{\mathrm{Im}}^+, A_{\mathrm{Im}}^- \in \mathcal{B}(H)^+$ such that*

$$\|A_{\mathrm{Re}}^+\|_{\mathrm{op}}, \|A_{\mathrm{Re}}^-\|_{\mathrm{op}}, \|A_{\mathrm{Im}}^+\|_{\mathrm{op}}, \|A_{\mathrm{Im}}^-\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}} \quad \textit{and}$$
$$A = A_{\mathrm{Re}}^+ - A_{\mathrm{Re}}^- + i(A_{\mathrm{Im}}^+ - A_{\mathrm{Im}}^-) .$$

If $A \in \mathcal{B}(H)^{\mathbb{R}}$,i.e., if $A = A^*$, then one can take $A_{\mathrm{Im}}^+ = A_{\mathrm{Im}}^- = 0$.

Proof If $A = 0$, the proposition trivially holds true with

$$A_{\mathrm{Re}}^+ = A_{\mathrm{Re}}^- = A_{\mathrm{Im}}^+ = A_{\mathrm{Im}}^- = 0 .$$

Thus, assume that $A \neq 0$. By the polarization identity for operators,

$$\begin{aligned} A &= \|A\|_{\mathrm{op}} (\|A\|_{\mathrm{op}}^{-1} A^*)^* \mathrm{id}_H \\ &= \sum_{n=1}^{4} (-i)^n \frac{\|A\|_{\mathrm{op}}}{4} (\|A\|_{\mathrm{op}}^{-1} A^* + i^n \mathrm{id}_H)^* (\|A\|_{\mathrm{op}}^{-1} A^* + i^n \mathrm{id}_H) . \end{aligned}$$

Observe that, for any $n \in \{1, 2, 3, 4\}$,

$$\frac{\|A\|_{\text{op}}}{4}(\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H)^*(\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H) \geq 0 \,.$$

If $A = A^*$, then

$$\sum_{n=1,3} (-i)^n(\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H)^*(\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H) = 0$$

and A_{Im}^+, A_{Im}^- are both chosen equal to zero, in this case. Finally, note that

$$\begin{aligned}\left\|(\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H)^*(\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H)\right\|_{\text{op}} &= \left\|\|A\|_{\text{op}}^{-1} A^* + i^n \text{id}_H\right\|_{\text{op}}^2 \\ &\leq (\|A\|_{\text{op}}^{-1} \left\|A^*\right\|_{\text{op}} + 1)^2 = 4 \,.\end{aligned}$$

□

At this point, note from Exercise 1.8 that the fact that $\mathcal{B}(H)^+$ is generating for $\mathcal{B}(H)$, where H is any complex Hilbert space, implies that any *positive* linear functional $\varphi \in \mathcal{B}(H)'^+ \doteq \mathcal{L}^+(\mathcal{B}(H); \mathbb{C})$ is real, i.e., it is self-conjugate in $\mathcal{B}(H)' \doteq \mathcal{L}(\mathcal{B}(H); \mathbb{C})$. This property allows a (canonical) construction of scalar semiproducts on $\mathcal{B}(H)$ for any $\varphi \in \mathcal{B}(H)'^+$, as proven in the next proposition. This fact is very important in the theory of operator algebras and C^*-algebras.

Proposition 2.7 (GNS Scalar Semiproduct) *Let H be a Hilbert space over $\mathbb{C}$. For any linear functional $\varphi \in \mathcal{B}(H)' \doteq \mathcal{L}(\mathcal{B}(H); \mathbb{C})$, define the sesquilinear form $\langle \cdot, \cdot \rangle_\varphi : \mathcal{B}(H) \times \mathcal{B}(H) \to \mathbb{C}$ (Definition 7.200) by*

$$\langle A, B \rangle_\varphi \doteq \varphi(A^* B) \,, \qquad A, B \in \mathcal{B}(H) \,.$$

$\langle \cdot, \cdot \rangle_\varphi$ is Hermitian[1] if φ is self-conjugate[2] in $\mathcal{L}(\mathcal{B}(H); \mathbb{C})$. It is a scalar semiproduct[3] whenever $\varphi \in \mathcal{B}(H)'^+ \doteq \mathcal{L}^+(\mathcal{B}(H); \mathbb{C})$ (i.e., φ is a positive linear functional on $\mathcal{B}(H)$).

Proof

1. The linearity of the mapping $B \mapsto \langle A, B \rangle_\varphi$ is obvious, and the antilinearity of $A \mapsto \langle A, B \rangle_\varphi$ is a consequence of the antilinearity of the involution $A \mapsto A^*$.

[1] For all $A, B \in \mathcal{B}(H)$, $\langle A, B \rangle_\varphi = \overline{\langle B, A \rangle_\varphi}$.

[2] For any $A \in \mathcal{B}(H)$, $\varphi(A^*) = \overline{\varphi(A^*)}$.

[3] A scalar semiproduct $\langle \cdot, \cdot \rangle$ is the same notion as a scalar product, except the positive definiteness does not necessarily hold true, i.e., $\langle v, v \rangle = 0$ does not yield $v = 0$.

2. From the polarization identity for operators, for $A, B \in \mathcal{B}(H)$,

$$\varphi(A^*B) = \frac{1}{4}\sum_{n=1}^{4}(-i)^n \varphi((A + i^n B)^*(A + i^n B)) .$$

3. If φ is self-conjugate in $\mathcal{L}(\mathcal{B}(H); \mathbb{C})$ then, for any $n \in \{1, 2, 3, 4\}$, $\varphi((A + i^n B)^*(A + i^n B)) \in \mathbb{R}$, because $(A + i^n B)^*(A + i^n B)$ is selfadjoint. We obtain in this case that

$$\begin{aligned}\overline{\varphi(A^*B)} &= \frac{1}{4}\sum_{n=1}^{4} i^n \varphi((A + i^n B)^*(A + i^n B)) \\ &= \frac{1}{4}\sum_{n=1}^{4} i^n \varphi((B + (-i)^n A)^*(B + (-i)^n A)) \\ &= \frac{1}{4}\sum_{n=1}^{4} (-i)^n \varphi((B + i^n A)^*(B + i^n A)) \\ &= \varphi(B^*A) .\end{aligned}$$

   In other words, for all $A, B \in \mathcal{B}(H)$, $\langle A, B\rangle_\varphi = \overline{\langle B, A\rangle_\varphi}$.
4. If φ is positive, then

$$\langle A, A\rangle_\varphi = \varphi(A^*A) \geq 0$$

   for all $A \in \mathcal{B}(H)$, because $A^*A \in \mathcal{B}(H)^+$. Recall, finally, that any $\varphi \in \mathcal{L}^+(\mathcal{B}(H); \mathbb{C})$ is self-conjugate in $\mathcal{L}(\mathcal{B}(H); \mathbb{C})$.

□

Given a complex Hilbert space H, recall that $\mathcal{B}(H)$ is an Archimedean $*$-ordered space having an order unit, by Exercises 2.3 and 2.4. Therefore, we can apply Proposition 1.47 to $V = \mathcal{B}(H)$ and arrive at the following important fact.

Corollary 2.8 *For any complex Hilbert space H*

$$\mathcal{B}(H)^{\mathrm{td}} \doteq \mathcal{B}(\mathcal{B}(H); \mathbb{C}) = \mathcal{B}(H)^{\mathrm{od}} \doteq \mathrm{span}(\mathcal{L}^+(\mathcal{B}(H); \mathbb{C})) .$$

In other words, the topological and the order dual spaces of $\mathcal{B}(H)$ coincide.

Proof By Exercises 1.33, 1.40, and 1.42 together with Exercises 2.3 and 2.4, any real space $\mathcal{B}(H)^{\mathbb{R}} \doteq \mathrm{Re}\{\mathcal{B}(H)\}$ of selfadjoint operators acting on a Hilbert space H is an Archimedean real ordered vector space with an order unit and a locally full norm. Therefore, it suffices to apply Corollary 1.45 on $V_1 = \mathcal{B}(H_1)^{\mathbb{R}}$ and $V_2 = \mathcal{B}(H_2)^{\mathbb{R}}$ to get the assertion. □

In fact, we have the following important identity for the operator norm of *positive* linear functional on $\mathcal{B}(H)$.

Lemma 2.9 *Let H be any complex Hilbert space. For all $\varphi \in \mathcal{B}(H)'^+$ (i.e., for any positive linear functional on $\mathcal{B}(H)$), one has*

$$\|\varphi\|_{op} \doteq \sup\left\{|\varphi(A)| : A \in \mathcal{B}(H),\ \|A\|_{op} = 1\right\} = \varphi(\mathrm{id}_H) ,$$

where we recall that id_H *is the identity operator on H.*

Proof Recall that $\mathcal{B}(H)^{\mathrm{od}} = \mathcal{B}(H)^{\mathrm{td}}$ and thus, for all $\varphi \in \mathcal{B}(H)'^+ \subseteq \mathcal{B}(H)^{\mathrm{od}}$, $\|\varphi\|_{op} < \infty$. Take any $\varphi \in \mathcal{B}(H)'^+$. Clearly, $\|\varphi\|_{op} \geq \varphi(\mathrm{id}_H)$, because $\|\mathrm{id}_H\|_{op} = 1$. By the Cauchy-Schwarz inequality for the scalar semiproduct $\langle\cdot,\cdot\rangle_\varphi$, for all $A, A' \in \mathcal{B}(H)$, one has

$$\left|\varphi(A^* A')\right| \leq \sqrt{\varphi(A^* A)}\sqrt{\varphi(A'^* A')} .$$

By choosing $A = \mathrm{id}_H = \mathrm{id}_H^*$, we arrive at

$$\left|\varphi(A')\right|^2 \leq \varphi(\mathrm{id}_H)\varphi(A'^* A') \leq \varphi(\mathrm{id}_H)\,\|\varphi\|_{op}\,\left\|A'\right\|_{op}^2$$

for all $A' \in \mathcal{B}(H)$. Now, taking the supremum over $A' \in \mathcal{B}(H)$, $\left\|A'\right\|_{op} = 1$, we conclude that $\|\varphi\|_{op}^2 \leq \varphi(\mathrm{id}_H)\,\|\varphi\|_{op}$ from which the inequality $\|\varphi\|_{op} \leq \varphi(\mathrm{id}_H)$ follows. □

We will prove later on that the identity $\|\varphi\|_{op} = \varphi(\mathrm{id}_H)$ is not only a necessary condition for the positivity of a linear functional $\varphi \in \mathcal{B}(H)'$, but it is also sufficient.

Observe from Exercise 7.77 (ii) that, given any two complex Hilbert spaces H_1, H_2, the space $\mathrm{Re}\{\mathcal{B}(\mathcal{B}(H_1); \mathcal{B}(H_2))\}$ can be canonically identified (by restriction) with $\mathcal{B}(\mathcal{B}(H_1)^{\mathbb{R}}; \mathcal{B}(H_2)^{\mathbb{R}})$, recalling that $\mathcal{B}(H_j)^{\mathbb{R}}$ is a short notation for $\mathrm{Re}\{\mathcal{B}(H_j\}$, $j = 1, 2$. Thus, by Exercises 2.3 and 2.4, we apply Corollary 1.45 on $V_1 = \mathcal{B}(H_1)^{\mathbb{R}}$ and $V_2 = \mathcal{B}(H_2)^{\mathbb{R}}$ to get the following assertion.

Corollary 2.10 *Let H_1, H_2 be any two complex Hilbert spaces. Every element of $\mathcal{L}_{\mathrm{ob}}(\mathcal{B}(H_1)^{\mathbb{R}}; \mathcal{B}(H_2)^{\mathbb{R}})$, i.e., any order-bounded*[4] *(real) linear transformation $\mathcal{B}(H_1)^{\mathbb{R}} \to \mathcal{B}(H_2)^{\mathbb{R}}$, is the restriction of a unique continuous linear transformation $\mathcal{B}(H_1) \to \mathcal{B}(H_2)$. In particular*

$$\mathcal{L}^+(\mathcal{B}(H_1); \mathcal{B}(H_2)) \subseteq \mathcal{B}(\mathcal{B}(H_1); \mathcal{B}(H_2)) ,$$

i.e., every positive linear transformation $\mathcal{B}(H_1) \to \mathcal{B}(H_2)$ is continuous. Conversely, the restriction to $\mathcal{B}(H_1)^{\mathbb{R}}$ of any self-conjugate *continuous linear trans-*

[4] It maps any order-bounded subset of $\mathrm{Re}\{\mathcal{B}(H_1)\}$ to an order-bounded subset of $\mathrm{Re}\{\mathcal{B}(H_2)\}$. See Definitions 1.16 and 1.23.

formation $\mathcal{B}(H_1) \to \mathcal{B}(H_2)$ defines an order-bounded (real) linear transformation $\mathcal{B}(H_1)^{\mathbb{R}} \to \mathcal{B}(H_2)^{\mathbb{R}}$.

Proof By Exercises 1.33, 1.40, and 1.42 together with Exercises 2.3 and 2.4, any real space $\mathcal{B}(H)^{\mathbb{R}} \doteq \mathrm{Re}\{\mathcal{B}(H)\}$ of selfadjoint operators acting on a Hilbert space H is an Archimedean real ordered vector space with an order unit and a locally full norm. Therefore, it suffices to apply Corollary 1.45 on $V_1 = \mathcal{B}(H_1)^{\mathbb{R}}$ and $V_2 = \mathcal{B}(H_2)^{\mathbb{R}}$ to get the assertion. □

Once again from the last two exercises combined now with Corollary 1.48, we conclude that, for any complex Hilbert space H, $\mathcal{B}(H)$ is a semisimple preordered vector space, i.e., for any $A, A' \in \mathcal{B}(H)$, $A \neq A'$ only if, for some *positive* $\varphi \in \mathcal{B}(H)'$, one has $\varphi(A) \neq \varphi(A')$. In other words, there is a canonical *injective* morphism of preordered vector spaces $\Pi : \mathcal{B}(H) \to \mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{C})$, defined by

$$[\Pi(A)](\varphi) \doteq \varphi(A) , \qquad \varphi \in \mathcal{B}(H)'^{+} , \ A \in \mathcal{B}(H) .$$

We prove below that Π is *bipositive* and self-conjugate in $\mathcal{L}(\mathcal{B}(H); \mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{C}))$, that is, for any complex Hilbert space H, $\mathcal{B}(H)$ is equivalent to a self-conjugate subspace of $\mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{C})$, as a preordered $*$-vector space.

Lemma 2.11 *For any complex Hilbert space H, $\Pi : \mathcal{B}(H) \to \mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{C})$, as defined above, is a bipositive self-conjugate linear transformation.*

Proof

1. By Exercise 1.8, any positive linear functional $\varphi : \mathcal{B}(H) \to \mathbb{C}$ is Hermitian, i.e., it is self-conjugate in $\mathcal{L}(\mathcal{B}(H); \mathbb{C})$. Thus, for all $A \in \mathcal{B}(H)$ and all $\varphi \in \mathcal{B}(H)'^{+}$,

$$[\Pi(A^*)](\varphi) = \varphi(A^*) = \overline{\varphi(A)} = [\Pi(A)]^*(\varphi) .$$

   In other words, $\Pi \in \mathrm{Re}\{\mathcal{L}(\mathcal{B}(H); \mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{C}))\}$.
2. We already know that Π is positive. Thus, it only remains to prove that it reflects positivity. Take any $A \in \mathcal{B}(H)$ and assume that, for all $\varphi \in \mathcal{B}(H)'^{+}$, $[\Pi(A)](\varphi) \geq 0$, that is, $\varphi(A) \geq 0$. In this case, for all $x \in H$, $\langle x, A(x)\rangle \geq 0$ (recall that $\langle x, (\cdot)(x)\rangle \in \mathcal{B}(H)'^{+}$), i.e., A is positive.

□

By the last lemma, for any complex Hilbert space H, the real ordered vector space $\mathcal{B}(H)^{\mathbb{R}}$ is equivalent to a vector subspace of $\mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{R})$. However, this subspace is generally not a function space (in the sense of Definition 7.270), i.e., a Riesz subspace of the Riesz space $\mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{R})$, for $\mathcal{B}(H)^{\mathbb{R}}$ is (generally) not a Riesz space, $\mathcal{B}(H)$ being (generally) noncommutative. (See Sherman's theorem, referring here to Theorem 7.319.) For more details on Riesz spaces, see the appendix. Note finally that, for all $A \in \mathcal{B}(H)$, the mapping $\mathcal{B}(H)'^{+} \to \mathbb{C}$, $\varphi \mapsto \varphi(A)$, is trivially continuous with respect to the metric defined by the operator norm on $\mathcal{B}(H)'^{+} \subseteq \mathcal{B}(\mathcal{B}(H); \mathbb{C})$. Hence, $\mathcal{B}(H)$ is (canonically) equivalent, as a $*$-preordered vector space, to a (self-conjugate) vector subspace of the space of complex-valued continuous functions on some metric space.

Proposition 2.12 *Let H be any complex Hilbert space. Then, $A \in \mathcal{B}(H)^{\mathbb{R}}$, $A \neq 0$, is positive iff*

$$\left\| \mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A \right\|_{\mathrm{op}} \leq 1 .$$

Proof

1. Take any $A \in \mathcal{B}(H)^{+} \subseteq \mathcal{B}(H)^{\mathbb{R}}$, $A \neq 0$. Note that $\mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A \in \mathcal{B}(H)^{\mathbb{R}}$. By Proposition 7.233

$$\begin{aligned}
\left\| \mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A \right\|_{\mathrm{op}} &= \sup_{x \in H,\ \|x\|=1} \left| \left\langle x, (\mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A)(x) \right\rangle \right| \\
&= \sup_{x \in H,\ \|x\|=1} \left| 1 - \|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \right| \\
&= \sup_{x \in H,\ \|x\|=1} \left(1 - \|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \right) \\
&= 1 - \|A\|_{\mathrm{op}}^{-1} \inf_{x \in H,\ \|x\|=1} \langle x, A(x) \rangle \leq 1 .
\end{aligned}$$

Observe that we used that $\|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \in [0, 1]$ for $x \in H$, $\|x\| = 1$, which is a consequence of the positivity of A and the Cauchy-Schwarz inequality.

2. Now, assume, for $A \in \mathcal{B}(H)^{\mathbb{R}}$, that

$$\left\| \mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A \right\|_{\mathrm{op}} \leq 1 .$$

Again by Proposition 7.233

$$\begin{aligned}
1 &\geq \sup_{x \in H,\ \|x\|=1} \left| \left\langle x, (\mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A)(x) \right\rangle \right| \\
&= \sup_{x \in H,\ \|x\|=1} \left| 1 - \|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \right| .
\end{aligned}$$

As $A \in \mathcal{B}(H)^{\mathbb{R}}$, $\langle x, A(x) \rangle \in \mathbb{R}$ and, by the Cauchy-Schwarz inequality, $\|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \leq 1$ for all $x \in H$, $\|x\| = 1$. Hence,

$$\begin{aligned}
1 &\geq \sup_{x \in H,\ \|x\|=1} \left| 1 - \|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \right| \\
&= \sup_{x \in H,\ \|x\|=1} \left(1 - \|A\|_{\mathrm{op}}^{-1} \langle x, A(x) \rangle \right) \\
&= 1 - \|A\|_{\mathrm{op}}^{-1} \inf_{x \in H,\ \|x\|=1} \langle x, A(x) \rangle .
\end{aligned}$$

That is,

$$\inf_{x \in H,\ \|x\|=1} \langle x, A(x) \rangle \geq 0$$

and, thus, $A \in \mathcal{B}(H)^+$.

□

Observe that the above characterization of positivity in $\mathcal{B}(H)$ does not make any direct reference to the scalar product of the Hilbert space H, but only to the operator norm of $\mathcal{B}(H)$, as well as selfadjointness. For this reason, it may be used as the *definition* of positivity in general (abstract) unital C^*-algebras. This characterization of the positive cone $\mathcal{B}(H)^+$ leads to the following important property of positive functionals on $\mathcal{B}(H)$:

Corollary 2.13 *Let H be a complex Hilbert space. $\varphi \in \mathrm{Re}\{\mathcal{B}(H)'\}$ (i.e., φ is Hermitian) is positive iff $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$.*

Proof Recall that we already proved above that $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$ for all $\varphi \in \mathcal{B}(H)'^{+} \subseteq \mathrm{Re}\{\mathcal{B}(H)^{\mathrm{td}}\}$. So, for a given $\varphi \in \mathrm{Re}\{\mathcal{B}(H)'\}$, assume now that $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$. For any $A \in \mathcal{B}(H)^+ \subseteq \mathcal{B}(H)^{\mathbb{R}}$, $A \neq 0$, one has

$$\varphi\left(\mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A\right) \in \mathbb{R} ,$$

because $\mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A \in \mathcal{B}(H)^{\mathbb{R}}$ and φ is a Hermitian functional. As A is positive, by the last proposition,

$$\left\| \mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A \right\|_{\mathrm{op}} \leq 1 .$$

Hence, as $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$,

$$\varphi\left(\mathrm{id}_H - \|A\|_{\mathrm{op}}^{-1} A\right) = \varphi\left(\mathrm{id}_H\right) - \|A\|_{\mathrm{op}}^{-1} \varphi\left(A\right) \leq \varphi(\mathrm{id}_H)$$

and we conclude that $0 \leq \varphi\left(A\right)$. □

In the last corollary, we assumed that the linear functional φ is Hermitian. It turns out that this condition can be removed.

Corollary 2.14 *Let H be a complex Hilbert space. $\varphi \in \mathcal{B}(H)'$ (i.e., φ is not assumed to be Hermitian) is positive iff $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$.*

Proof

1. Recall that positive functionals are automatically Hermitian. Thus, from the last corollary, if $\varphi \in \mathcal{B}(H)'$ is positive, then $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$. Assume thus that $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H) \neq 0$ (the special case $\varphi = 0$ is trivial) holds for a given $\varphi \in$

$\mathcal{B}(H)'$. We will show that this implies that φ must be Hermitian and the positivity of φ then follows from the above corollary.

2. Take any $A \in \mathcal{B}(H)^{\mathbb{R}}$. Then, for all $\alpha \in \mathbb{R}$

$$\varphi(A + i\alpha \mathrm{id}_H) = \mathrm{Re}\{\varphi(A)\} + i(\alpha\varphi(\mathrm{id}_H) + \mathrm{Im}\{\varphi(A)\}) .$$

Note that we used the equality $\varphi(\mathrm{id}_H) = \|\varphi\|_{\mathrm{op}} \in \mathbb{R}$. In particular

$$|\alpha\varphi(\mathrm{id}_H) + \mathrm{Im}\{\varphi(A)\}| \le |\varphi(A + i\alpha \mathrm{id}_H)| \le \varphi(\mathrm{id}_H) \|A + i\alpha \mathrm{id}_H\|_{\mathrm{op}} .$$

Hence, for all $\alpha \in \mathbb{R}$

$$(\alpha + \varphi(\mathrm{id}_H)^{-1} \mathrm{Im}\{\varphi(A)\})^2 \le \|A + i\alpha \mathrm{id}_H\|^2_{\mathrm{op}} .$$

3. But by Proposition 7.233

$$\begin{aligned} \|A + i\alpha \mathrm{id}_H\|^2_{\mathrm{op}} &= \sup_{x\in H,\ \|x\|=1} \left\langle x, (A + i\alpha \mathrm{id}_H)^* (A + i\alpha \mathrm{id}_H)(x)\right\rangle \\ &= \sup_{x\in H,\ \|x\|=1} \left\langle x, (A^2 + \alpha^2 \mathrm{id}_H)(x)\right\rangle \\ &= \alpha^2 + \sup_{x\in H,\ \|x\|=1} \left\langle x, A^2(x)\right\rangle = \|A\|^2_{\mathrm{op}} + \alpha^2 . \end{aligned}$$

Hence, for all $\alpha \in \mathbb{R}$,

$$2\alpha\varphi(\mathrm{id}_H)^{-1} \mathrm{Im}\{\varphi(A)\} + \varphi(\mathrm{id}_H)^{-2} \mathrm{Im}\{\varphi(A)\}^2 \le \|A\|^2_{\mathrm{op}} .$$

This condition can only be satisfied for all $\alpha \in \mathbb{R}$ when $\mathrm{Im}\{\varphi(A)\} = 0$.

4. Recall finally that $\varphi(A) \in \mathbb{R}$ for all $A \in \mathcal{B}(H)^{\mathbb{R}}$ is equivalent to φ being a Hermitian functional.

□

We summarize in the next two propositions the most important properties of the positive cones $\mathcal{B}(H)^+$ and $\mathcal{B}(H)'^+$, proven above.

Proposition 2.15 (Properties of $\mathcal{B}(H)^+$) *Let H be any complex Hilbert space. The convex cone $\mathcal{B}(H)^+$ of positive bounded operators on H has the following properties:*

(i) *It is a norm-closed pointed cone. In particular, $\mathcal{B}(H)$ is an Archimedean ordered vector space.*
(ii) *It is a subset of $\mathcal{B}(H)^{\mathbb{R}}$, i.e., a positive operator on H is necessarily selfadjoint. Thus, $\mathcal{B}(H)$ is a $*$-(pre)ordered vector space.*
(iii) *For all $A \in \mathcal{B}(H)$, $A^*A \in \mathcal{B}(H)^+$, the positivity of orthogonal projectors being a special case of this property.*

(iv) *For any* $A \in \mathcal{B}(H)$*, there are* $A^+_{\mathrm{Re}}, A^-_{\mathrm{Re}}, A^+_{\mathrm{Im}}, A^-_{\mathrm{Im}} \in \mathcal{B}(H)^+$ *such that*

$$\|A^+_{\mathrm{Re}}\|_{\mathrm{op}}, \|A^-_{\mathrm{Re}}\|_{\mathrm{op}}, \|A^+_{\mathrm{Im}}\|_{\mathrm{op}}, \|A^-_{\mathrm{Im}}\|_{\mathrm{op}} \leq \|A\|_{\mathrm{op}} \quad \textit{and}$$
$$A = A^+_{\mathrm{Re}} - A^-_{\mathrm{Re}} + i(A^+_{\mathrm{Im}} - A^-_{\mathrm{Im}}) .$$

In particular, $\mathcal{B}(H)^+$ *is generating for* $\mathcal{B}(H)$*.*

(v) *Take any* $A \in \mathcal{B}(H)^{\mathbb{R}}$*,* $A \neq 0$*. One has* $A \in \mathcal{B}(H)^+$ *iff*

$$\left\| \mathrm{id}_H - \|A\|^{-1}_{\mathrm{op}} A \right\|_{\mathrm{op}} \leq 1 .$$

Proposition 2.16 (Properties of Dual Positive Cone of $\mathcal{B}(H)$**)** *Let* H *be any complex Hilbert space. The convex cone* $\mathcal{B}(H)'^+$ *of positive linear functionals on* $\mathcal{B}(H)$ *has the following properties:*

(i) *It is a pointed cone whose elements are all bounded linear functionals, that is,* $\mathcal{B}(H)'^+ \subseteq \mathcal{B}(H)^{\mathrm{td}}$*. In particular, the dual space* $\mathcal{B}(H)'$ *is an ordered vector space. In fact, note that* $\mathcal{B}(H)'^+$ *is pointed, because* $\mathcal{B}(H)^+$ *is generating for* $\mathcal{B}(H)$*. See Exercise 1.29.*

(ii) *It is a subset of* $\mathrm{Re}\{\mathcal{B}(H)'\}$*, i.e., a positive linear functional on* $\mathcal{B}(H)$ *is necessarily Hermitian. Thus,* $\mathcal{B}(H)'$ *is a* $*$*-(pre)ordered vector space.*

(iii) $\mathcal{B}(H)^{\mathrm{td}} = \mathcal{B}(H)^{\mathrm{od}}$*. In particular,* $\mathcal{B}(H)'^+ \subseteq \mathcal{B}(H)^{\mathrm{td}}$ *is generating for* $\mathcal{B}(H)^{\mathrm{td}}$*.*

(iv) *Any* $\varphi \in \mathcal{B}(H)'$ *is a positive linear functional iff* $\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H)$*.*

(v) $\mathcal{B}(H)'^+$ *separates the elements of* $\mathcal{B}(H)$*, that is, for any* $A, B \in \mathcal{B}(H)$*,* $A \neq B$*, there is* $\varphi \in \mathcal{B}(H)'^+$ *such that* $\varphi(A) \neq \varphi(B)$*. (Recall that this property refers to the semisimplicity of* $\mathcal{B}(H)$ *as a preordered vector space, deduced from Corollary 1.48.)*

Observe that, by simple adaptations of proofs, all properties given above for the (C^*-)algebra of (all) bounded operators on any complex Hilbert spaces and its dual space can also be derived for its unital C^*-subalgebras (i.e., for any concrete C^*-algebra) containing the identity. In fact, we will see later on that these properties (up to simple modifications) hold true for arbitrary (abstract, not necessarily unital) C^*-algebras.

2.2 Monotone Order Completeness and von Neumann Algebras

By Sherman's theorem (Theorem 7.319), for a general complex Hilbert space H, $\mathrm{Re}\{\mathcal{B}(H)\}$ is not a Riesz space (because $\mathcal{B}(H)$ is generally noncommutative). Thus, $\mathrm{Re}\{\mathcal{B}(H)\}$ and, consequently, $\mathcal{B}(H)$ are not order-complete, in general. However, as we demonstrate below, $\mathrm{Re}\{\mathcal{B}(H)\}$ still satisfies a consequence of order completeness, the existence of an order limit for any bounded *monotone* net. See, in this context, Corollary 7.276. As we will also discuss in this paragraph, the (related) property of closedness with respect to order limits of bounded monotone nets characterizes von Neumann algebras and can be used as an abstract definition of these algebras.

Proposition 2.17 *Let H be any (real or complex) Hilbert space. Any increasing (decreasing) monotone net $(A_i)_{i\in I}$ of selfadjoint elements in $\mathcal{B}(H)$ that is bounded from above (below) has a supremum (infimum) in $\mathcal{B}(H)$. The supremum (infimum) is itself also selfadjoint and is the weak operator*[5] *limit of the net.*

Proof We only consider the complex case, the real one being even simpler.

1. Let $(A_j)_{j\in J}$ be an increasing net in $\mathcal{B}(H)^{\mathbb{R}}$ with $A_j \leq A$ for some $A \in \mathcal{B}(H)^{\mathbb{R}}$ and all $j \in J$. From (a version of) the polarization identity for (the sesquilinear form) $\langle \cdot, A_j(\cdot)\rangle$, for all $x, x' \in H$ and $j \in J$,

$$\langle x, A_j(x')\rangle = \frac{1}{4}\sum_{n=1}^{4}(-i)^n \langle x + i^n x', A_j(x + i^n x')\rangle \ .$$

   As, for fixed $x, x' \in H$, $n \in \{1, 2, 3, 4\}$ and all $j_1, j_2 \in J$, $j_1 \preceq j_2$, one has

$$\begin{aligned}\langle x + i^n x', A_{j_1}(x + i^n x')\rangle &\leq \langle x + i^n x', A_{j_2}(x + i^n x')\rangle \\ &\leq \langle x + i^n x', A(x + i^n x')\rangle \ ,\end{aligned}$$

   the net $(\langle x, A_j(x')\rangle)_{j\in J}$ in $\mathbb{C}$ converges to

$$[x, x'] \doteq \frac{1}{4}\sum_{n=1}^{4}(-i)^n \sup_{j\in J}\langle x + i^n x', A_j(x + i^n x')\rangle \ .$$

   Note that $(x, x') \mapsto [x, x']$ defines a sesquilinear form on H (see Definition 7.200).
2. Additionally, from the bound

$$\begin{aligned}\langle x + i^n x', A_{j_0}(x + i^n x')\rangle &\leq \sup_{j\in J}\langle x + i^n x', A_j(x + i^n x')\rangle \\ &\leq \langle x + i^n x', A(x + i^n x')\rangle \ ,\end{aligned}$$

   for any fixed $j_0 \in J$ and $n \in \{1, 2, 3, 4\}$, we arrive at the uniform estimate:

[5] A net $(A_i)_{i\in I}$ of operators in $\mathcal{B}(H)$ converges in the weak (operator) topology to $A \in \mathcal{B}(H)$ when $\langle x, A_i x'\rangle$ converges to $\langle x, Ax'\rangle$ for any fixed $x, x' \in H$. For more details on weak operator convergence and general net convergence, see the appendix.

$$|[x, x']| \leq 4 \max\{\|A_{j_0}\|_{\text{op}}, \|A\|_{\text{op}}\}$$

for all $x, x' \in H$, $\|x\| = \|x'\| = 1$. By the sesquilinearity of $[\cdot, \cdot]$, it follows that

$$|[x, x']| \leq 4 \max\{\|A_{j_0}\|_{\text{op}}, \|A\|_{\text{op}}\} \|x\| \|x'\| , \qquad x, x' \in H .$$

In other words, $[\cdot, \cdot]$ is a bounded sesquilinear form (see Definition 7.200).

3. Hence, by the Riesz representation theorem for bounded sesquilinear forms (Corollary 7.216), there is $A_\infty \in \mathcal{B}(H)$ such that

$$[x, x'] = \langle x, A_\infty(x') \rangle , \qquad x, x' \in H .$$

By construction, A_∞ is the limit of the net A_j in the weak operator topology. Moreover, for all $x \in H$

$$\langle x, A_\infty(x) \rangle = \sup_{j \in J} \langle x, A_j(x) \rangle \in \mathbb{R} .$$

Thus, by Corollary 7.234, A_∞ is selfadjoint. From the above equality, we also conclude that, for all $j \in J$, $A_\infty \geq A_j$. That is, A_∞ is an upper bound for the net $(A_j)_{j \in J}$.

4. Let $A \in \mathcal{B}(H)$ be any upper bound for the net $(A_j)_{j \in J}$. Then, for all $x \in H$

$$\langle x, A(x) \rangle \geq \sup_{j \in J} \langle x, A_j(x) \rangle = \langle x, A_\infty(x) \rangle .$$

Hence, A_∞ is the supremum in $\mathcal{B}(H)$ of the net $(A_j)_{j \in J}$. The existence of the infimum for a net $(A_j)_{j \in J}$ which is bounded from below follows from the existence of the supremum for the net $(-A_j)_{j \in J}$, which is, in this case, bounded from above.

□

Recalling that (concrete) von Neumann algebras are algebras of bounded operators on a Hilbert space that are closed in the weak operator topology (see Appendix for more details), it follows that these algebras are closed with respect to monotone order convergence.

Corollary 2.18 *Let H be any complex Hilbert space and $\mathfrak{M} \subseteq \mathcal{B}(H)$ a von Neumann algebra. For any increasing (decreasing) monotone net $(A_i)_{i \in I}$ in* $\text{Re}\{\mathfrak{M}\}$ *that is bounded from above (below), one has that $A_i \uparrow A$ ($A_i \downarrow A$) for some $A \in$* $\text{Re}\{\mathfrak{M}\}$.

In fact, from a well-known result proven by Kadison, it follows that if $\mathfrak{M} \subseteq \mathcal{B}(H)$ is a (concrete) C^*-algebra, then $\mathfrak{M}$ is a von Neumann algebra (i.e., it is closed in the weak operator topology) if every increasing net in $\mathfrak{M}^+$ ($\subseteq \text{Re}\{\mathfrak{M}\}$) that is bounded in $\mathfrak{M}$ has its monotone order limit in $\mathfrak{M}^+$. See, for instance, [51, Section 2.4.3, Theorem (Kadison)]. We will discuss more about this result below. See discussions prior to Corollary 2.31.

In view of the last corollary and Exercise 7.291, we introduce the following notion, which is a sort of "noncommutative generalization" of normal positive integrals (see Sect. 7.4.5) on Riesz spaces.

Definition 2.19 (Normal Positive Functionals on von Neumann Algebras) Let H be any complex Hilbert space and $\mathfrak{M} \subseteq \mathcal{B}(H)$ a von Neumann algebra. We say that the positive functional $\varphi \in \mathfrak{M}'^{+}$ is "normal" if, for any bounded monotone increasing net $(A_i)_{i\in I}$ of positive elements of $\mathfrak{M}$, the (bounded monotone increasing) net $(\varphi(A_i))_{i\in I}$ converges in $\mathbb{R}_0^+$ to $\varphi(\sup_{i\in I} A_i)$.

Normal positive functionals are very important in the theory of general von Neumann algebras. In the special case of commutative von Neumann algebras, it turns out that these algebras are order-complete Riesz spaces and the corresponding normal positive functionals on them are exactly the normal integrals of Definition 7.309.

2.3 The Lattice of Projectors

For complex Hilbert spaces, we have the following characterization of orthogonal projectors in relation to the preordered vector space structure of $\mathcal{B}(H)$.

Exercise 2.20 Let H be any complex Hilbert space. Prove that $P = P_G$ for some closed vector subspace $G \subseteq H$ iff $P \in \mathcal{L}(H)$ is a projector (i.e., $P^2 = P$) satisfying $\mathrm{id}_H \geq P \geq 0$.

Compare the last exercise with Lemma 7.303, the second part of which is a very similar property of projectors on Riesz spaces.

As already discussed, for any Hilbert space, the mapping $G \mapsto P_G$, where $P_G \in \mathcal{B}(H)^+$ is the orthogonal projector associated with the closed vector subspace $G \subseteq H$, defines a one-to-one correspondence between the family of all closed vector subspaces of H and the family of all selfadjoint projectors in $\mathcal{B}(H)$. It turns out that this mapping preserves and reflects order, the closed vector subspaces of H being partially ordered by the inclusion relation.

Exercise 2.21 Let H be any (real or complex) Hilbert space. Prove that, for any closed vector subspaces $G, G' \subseteq H$, $P_G \geq P_{G'}$ iff $G' \subseteq G$.

Hint: In order to prove that $P_G \geq P_{G'}$ show that $P_G - P_{G'}$ is an orthogonal projector.

For a given complex Hilbert space H, let $\mathcal{G}(H)$ denote the collection of all its *closed* subspaces, which is canonically a partially ordered space by the inclusion relation. Note that any nonempty subset $\tilde{\mathcal{G}} \subseteq \mathcal{G}(H)$ has an infimum:

$$\inf \tilde{\mathcal{G}} = \cap \tilde{\mathcal{G}} \ .$$

In fact, this follows from the fact that the intersection of any family of closed vector subspaces of a normed space is again a closed vector subspace.

Let $\mathcal{P}(H) \subseteq \mathcal{B}(H)^+$ be the family of all orthogonal projectors on H. By the last exercise and remark, any nonempty subset $\tilde{\mathcal{P}} \subseteq \mathcal{P}(H)$ and has infimum in $\mathcal{P}(H)$, where $\mathcal{P}(H)$ is a partially ordered set by the restriction of the (canonical) order relation of $\mathcal{B}(H)$. In fact, observe that $\tilde{\mathcal{P}}$ also has a supremum:

$$\sup \tilde{\mathcal{P}} = \mathrm{id}_H - \inf\{\mathrm{id}_H - P \ : \ P \in \tilde{\mathcal{P}}\} .$$

In other words, $\mathcal{P}(H)$ is a complete lattice.[6]

We will show in the following that the family orthogonal projectors $\mathcal{P}(\mathfrak{M}) \doteq \mathfrak{M} \cap \mathcal{P}(H)$ of any von Neumann algebra $\mathfrak{M} \subseteq \mathcal{B}(H)$ are order-closed in $\mathcal{P}(H)$. In particular, $\mathcal{P}(\mathfrak{M})$ is again a complete lattice. Note that this property is *not* satisfied by general C^*-subalgebras of $\mathcal{B}(H)$.

Lemma 2.22 *Let H be any complex Hilbert space and $\Omega \subseteq \mathcal{B}(H)$ any* self-conjugate[7] *nonempty subset. An orthogonal projector $P \in \mathcal{B}(H)$ belongs to the commutant:*

$$\Omega' \doteq \left\{A \in \mathcal{B}(H) : \forall B \in \Omega,\ [A, B] \doteq AB - BA = 0\right\} \subseteq \mathcal{B}(H)$$

(Definition 7.253) iff, for every $A \in \Omega$, one has that $AP(H) \subseteq P(H)$.

Proof Let $P \in \mathcal{B}(H)$ be any orthogonal projector. If $P \in \Omega'$, then, for all $A \in \Omega$, $AP = PA$ and the inclusion $AP(H) \subseteq P(H)$ trivially follows. Assume now that this inclusion holds true for all $A \in \Omega$. Then, for all $A \in \Omega$, $PAP = AP$ and, hence, by taking adjoints, $PA^*P = PA^*$. As Ω is self-conjugate, for all $A \in \Omega$, $A^* \in \Omega$, and we infer from the last observations that $AP = PAP = PA$ and conclude that

$$[A, P] = AP - PA = PAP - PAP = 0$$

for all $A \in \Omega$. That is, $P \in \Omega'$. □

Corollary 2.23 *Let H be any complex Hilbert space and $\mathfrak{M} \subseteq \mathcal{B}(H)$ any von Neumann algebra. The collection $\mathcal{P}(\mathfrak{M}) \subseteq \mathfrak{M}^+ \subseteq \mathcal{P}(H)$ of all orthogonal projectors in $\mathfrak{M}$ is order-closed in $\mathcal{P}(H)$, i.e., for any nonempty family $\tilde{\mathcal{P}} \subseteq \mathcal{P}(\mathfrak{M})$, the infimum and supremum of $\tilde{\mathcal{P}}$ in the set $\mathcal{P}(H)$ of all orthogonal projectors on H are elements of $\mathcal{P}(\mathfrak{M})$.*

---

[6] A partially ordered set $(P, \succeq)$ is called a "lattice" if any two elements $p, p' \in P$ have an infimum and a supremum. It is a complete lattice if all subsets of P have both a supremum and an infimum. See the appendix for more details.

[7] That is, for any $A \in \Omega$, $A^* \in \Omega$.

Proof Take any nonempty family $\tilde{\mathcal{P}} \subseteq \mathcal{P}(\mathfrak{M})$. Recall that, as $\mathfrak{M}$ is a von Neumann algebra, the bicommutant theorem [51, Theorem 2.4.11] tells us that $\mathfrak{M} = \mathfrak{M}'' \doteq (\mathfrak{M}')'$. In particular, $\mathfrak{M}$ is a commutant (the commutant of $\mathfrak{M}'$). By the last lemma, for all $P \in \tilde{\mathcal{P}} \subseteq \mathfrak{M}''$ and $A \in \mathfrak{M}'$

$$AP(H) \subseteq P(H) .$$

By Exercise 2.21 and the remark after it

$$[\inf \tilde{\mathcal{P}}](H) = \cap\{P(H) \ : \ P \in \tilde{\mathcal{P}}\} \subseteq H .$$

Thus, for all $A \in \mathfrak{M}'$

$$A[\inf \tilde{\mathcal{P}}](H) \subseteq [\inf \tilde{\mathcal{P}}](H) .$$

Again by the lemma, it follows that $\inf \tilde{\mathcal{P}} \in (\mathfrak{M}')' = \mathfrak{M}$. Finally, as $\mathrm{id}_H \in \mathfrak{M}$

$$\sup \tilde{\mathcal{P}} = \mathrm{id}_H - \inf\{\mathrm{id}_H - P : P \in \tilde{\mathcal{P}}\} \in \mathfrak{M} .$$

This last equality follows from the previous arguments applied to the family:

$$\{\mathrm{id}_H - P : P \in \tilde{\mathcal{P}}\} \subseteq \mathcal{P}(\mathfrak{M})$$

of orthogonal projectors in $\mathfrak{M}$. □

In order to complement the above corollary, we prove next that $\mathcal{P}(\mathfrak{M})$ is closed in $\mathfrak{M}$ with respect to the monotone order convergence.

Lemma 2.24 *Let H be any complex Hilbert space and $\mathfrak{M} \subseteq \mathcal{B}(H)$ any von Neumann algebra. The order limit of any monotonically increasing or decreasing net $(P_i)_{i\in I}$ of orthogonal projectors in $\mathfrak{M}$ (i.e., a net in $\mathcal{P}(\mathfrak{M})$) belongs to $\mathcal{P}(\mathfrak{M})$ (i.e., it is itself an orthogonal projector in $\mathfrak{M}$).*

Proof

1. Take any increasing net $(P_i)_{i\in I}$ of orthogonal projectors in $\mathfrak{M}$. Recall that any orthogonal projector is selfadjoint. The net is bounded from above (by the identity operator $\mathrm{id}_H \in \mathfrak{M}$) and thus has a supremum $P_\infty \in \mathfrak{M}^{\mathbb{R}}$, which is the limit of the net in the weak operator topology, by Proposition 2.17.
2. Note that, for all $j \in J$, $P_j P_\infty = P_j$. In fact, for all $x, x' \in H$ and all $j \in J$

$$\begin{aligned}\langle x, P_j P_\infty(x')\rangle &= \langle P_j(x), P_\infty(x')\rangle \\ &= \lim_{j'\in J} \langle P_j(x), P_{j'}(x')\rangle = \lim_{j'\in J} \langle x, P_j P_{j'}(x')\rangle \\ &= \lim_{j'\in J} \langle x, P_j(x')\rangle = \langle x, P_j(x')\rangle .\end{aligned}$$

Observe that $P_j P_{j'} = P_j$, because $P_j \leq P_{j'}$ whenever $j' \succeq j$, the net being monotone. See Exercises 7.239 (i) and 2.21.

3. Hence, for all $x, x' \in H$, one has

$$\left\langle x, P_\infty^2(x') \right\rangle = \lim_{j\in J} \left\langle x, P_j(P_\infty(x')) \right\rangle = \lim_{j\in J} \left\langle x, P_j(x') \right\rangle = \left\langle x, P_\infty(x') \right\rangle \ .$$

In other words, $P_\infty^2 = P_\infty$ and, hence, $P_\infty \in \mathfrak{M}^{\mathbb{R}}$ is an orthogonal projector of $\mathfrak{M}$.

4. If $(P_i)_{i\in I}$ is a decreasing net of orthogonal projectors, $(\mathrm{id}_H - P_j) \in \mathfrak{M}^{\mathbb{R}}$, $j \in J$, is a increasing net of orthogonal projectors and has thus a supremum. The element

$$\mathrm{id}_H - \sup_{j\in J}(\mathrm{id}_H - P_j)$$

is thus the infimum of the net.

□

2.4 General States of $\mathcal{B}(H)$

We discuss now the general notion of state on the algebra $\mathcal{B}(H)$ of bounded operators on a complex Hilbert space H. This is a key concept in the algebraic approach to quantum mechanics and refers to the following definition.

Definition 2.25 (States on $\mathcal{B}(H)$) Let H be any Hilbert space over $\mathbb{C}$:

(i) A linear functional $\rho \in \mathcal{B}(H)'$ is a "state" on $\mathcal{B}(H)$ if it is positive and "normalized," i.e., $\rho(\mathrm{id}_H) = 1$. The set of all states on $\mathcal{B}(H)$ is denoted by $E(H) \subseteq \mathcal{B}(H)'^+$.

(ii) We say that $\rho \in E(H)$ is a "normal state" if it is a normal positive functional on (the von Neumann algebra) $\mathcal{B}(H)$. $E^{\mathrm{n}}(H) \subseteq E(H)$ denotes the set of all normal states on $\mathcal{B}(H)$.

Note that, like the whole positive cone $\mathcal{B}(H)'^+$, the set $E(H)$ of all states on $\mathcal{B}(H)$ is a convex set. Observe also that $E(H)$ is a norm-closed subset of $\mathcal{B}(H)^{\mathrm{td}}$.

Exercise 2.26 Let H be any complex Hilbert space. Show that the set $E^{\mathrm{n}}(H) \subseteq \mathcal{B}(H)^{\mathrm{td}}$ of all normal states on $\mathcal{B}(H)$ is convex and norm-closed.

Definition 2.27 (Extreme States on $\mathcal{B}(H)$) Let H be any Hilbert space over $\mathbb{C}$. We say that $\rho \in E(H)$ is a "extreme state" on $\mathcal{B}(H)$, whenever it is an extreme point of the convex set $E(H)$, i.e., if

$$\rho = \lambda\rho' + (1-\lambda)\rho'' \ , \qquad \lambda \in (0,1) \ , \qquad \rho', \rho'' \in E(H) \ ,$$

implies that $\rho' = \rho'' = \rho$.

The following fact is well-known and results from the Krein-Milman theorem [18, Theorem 3.23] (see also Proposition 7.334) for locally convex spaces.

Theorem 2.28 *For every complex Hilbert space H, $E(H)$ contains extreme states.*

In fact, the above theorem holds true for all general C^*-algebras (thanks to the Krein-Milman theorem).

In the algebraic approach to quantum theory, the states in the above (mathematical) sense are identified with the (physical) states of a physical system whose observables are represented by the selfadjoint elements of $\mathcal{B}(H)$. In this context, for any observable $A \in \mathcal{B}(H)$ and state $\rho \in E(H)$, the number $\rho(A) \in \mathbb{R}$ is interpreted as being the expected value (in statistical sense) for the quantity represented by A when the system is found in the state ρ.

One simple, though very important, example of state on $\mathcal{B}(H)$ is the following:

Definition 2.29 (Vector State) Let H be any complex Hilbert space. For all $x \in H$, $\|x\| = 1$, we define the state $\rho_x \in E(H)$ by

$$\rho_x(A) \doteq \langle x, A(x) \rangle \ , \qquad A \in \mathcal{B}(H) \ .$$

A state $\rho \in E(H)$ is called a "vector state" if it has this form for some $x \in H$ satisfying $\|x\| = 1$.

If $\dim(H) > 1$, then not every state on $\mathcal{B}(H)$ is a vector state. In fact, vector states are examples of extreme states of $\mathcal{B}(H)$.

Exercise 2.30 Let H be any complex Hilbert space. Show that, for all $x \in H$, $\|x\| = 1$, one has $\rho_x \in E^{\mathrm{n}}(H)$. In other words, any vector state is normal.

By combining the last exercise with Corollary 7.234, we see that normal states separate the elements of $\mathcal{B}(H)$, that is, for any $A, A' \in \mathcal{B}(H)$, $A \neq A'$ only if $\rho(A) \neq \rho(A')$ for some state $\rho \in E^{\mathrm{n}}(H)$. (In particular, general states separate the elements of $\mathcal{B}(H)$.) Kadison used this fact, combined with the monotone order completeness (as discussed after Corollary 2.18), to give an abstract definition (i.e., with no reference to Hilbert spaces) of von Neumann algebras. See [51, Section 2.4.3, Theorem (Kadison)].

From Proposition 7.242 we arrive at the following property of the spectrum of selfadjoint operators:

Corollary 2.31 *Let H be a complex Hilbert space. For any $A \in \mathcal{B}(H)$*

$$\sigma(A) \subseteq \overline{[\Pi(A)](E(H))} \ ,$$

where we recall that $\sigma(A)$ denotes the "spectrum" of $A \in \mathcal{B}(H)$, i.e., the complement in $\mathbb{C}$ of its resolvent set $R(A)$, while

$$[\Pi(A)](E(H)) = \{\rho(A) \ : \ \rho \in E(H)\} \ .$$

Physically speaking, the above corollary says that, for any possible value a of the physical quantity represented by $A \in \mathcal{B}(H)$, there is some state $\rho \in E(H)$ of the corresponding physical system such that the expected value $\rho(A)$ of this quantity is as near as desired to the given value a. In fact, we will prove later on, for general unital C^*-algebras, that there is always $\rho \in E(H)$ such that $\rho(A) = a$, that is, one actually has

$$\sigma(A) \subseteq [\Pi(A)](E(H)) .$$

(By contrast,

$$\sigma(A) \nsubseteq N(A) \doteq \{\langle x, A(x)\rangle : x \in H, \|x\| = 1\} ,$$

in general.) In fact, such a state ρ reproducing any given spectral point of $A \in \mathcal{B}(H)$ can even be requested to be extreme. Note also that, in general, the expected values $\rho(A)$, $\rho \in E(H)$, are not necessarily in the spectrum of A, even if the state ρ is extreme. This is so, because of the statistical nature of physical states. In the special case of commutative C^*-algebras, i.e., for classical physical systems, $\rho(A)$ is a spectral point of A, whenever the state ρ is extreme.

From Corollary 2.14, we arrive at the following important characterization of the set of states on $\mathcal{B}(H)$:

Proposition 2.32 *Let H be any complex Hilbert space. A linear functional $\varphi \in \mathcal{B}(H)'$ is a state iff*

$$\|\varphi\|_{\mathrm{op}} = \varphi(\mathrm{id}_H) = 1 .$$

Any positive functional $\varphi \in \mathcal{B}(H)'^{+}$ is a positive multiple of a state.

Definition 2.33 (Gelfand Transform in $\mathcal{B}(H)$) Let H be any complex Hilbert space. We define a linear transformation $\Xi : \mathcal{B}(H) \to \mathcal{F}(E(H); \mathbb{C})$ by

$$\Xi(A) \doteq \Pi(A)|_{E(H)} , \qquad A \in \mathcal{B}(H) .$$

The function $\Xi(A) : E(H) \to \mathbb{C}$ is called here the "Gelfand transform" of $A \in \mathcal{B}(H)$.

Observe that, strictly speaking, what is most commonly called "Gelfand transform" refers rather to the case of *commutative* C^*-algebras and the domain of such transforms is not the set of all states on the given algebra, as above, but only the set of *extreme* states. This special and very important case will be discussed in more details later on.

Recall that $\Pi : \mathcal{B}(H) \to \mathcal{F}(\mathcal{B}(H)'^{+}; \mathbb{C})$ is an injective bipositive linear transformation. As the positive functionals on $\mathcal{B}(H)$ are exactly the positive multiples of the states on $\mathcal{B}(H)$, it follows that also $\Xi : \mathcal{B}(H) \to \mathcal{F}(E(H); \mathbb{C})$ is injective and bipositive. Moreover, like Π, Ξ is self-conjugate and its image is a vector subspace of $C(E(H); \mathbb{C})$, where $E(H)$ is endowed with the metric associated with the operator norm for linear functionals. Thus, $\mathcal{B}(H)$ is equivalent, as a $*$-ordered vector space, to the subspace of $C(E(H); \mathbb{C})$ formed by the Gelfand transforms of all elements of $\mathcal{B}(H)$. In fact, it is possible to choose a topology on $E(H)$ such that $E(H)$ is a compact Hausdorff[8] space, and the Gelfand transforms of elements of $\mathcal{B}(H)$ are still continuous. This holds for all general unital C^*-algebras. We will discuss this fact in more details later on, for the case of separable C^*-algebras (in which the referred topology is a metric one).

By combining the last proposition with Proposition 7.233, we obtain the following alternative definition for the operator norm in $\mathcal{B}(H)$.

Corollary 2.34 *Let H be any complex Hilbert space. For every selfadjoint $A \in \mathcal{B}(H)^{\mathbb{R}}$,*

$$\|A\|_{\mathrm{op}} = \sup_{\rho \in E(H)} |\rho(A)| \ .$$

In particular, for all (not necessarily selfadjoint) $A \in \mathcal{B}(H)$,

$$\|A\|_{\mathrm{op}} = \sup_{\rho \in E(H)} \sqrt{\rho(A^*A)} \ .$$

2.5 Finite-Dimensional Case: States as Density Matrices

In all the present subsection, H will denote any complex Hilbert space of *finite* dimension $N \in \mathbb{N}$ and $\{\mathrm{e}_n\}_{n=1}^N$ some fixed Hilbert basis (or orthonormal basis) of this space, that is,

$$\langle \mathrm{e}_n, \mathrm{e}_{n'} \rangle = \delta_{n,n'} \ , \qquad n, n' \in \{1, 2, \ldots, N\} \ ,$$

where $\delta_{\cdot,\cdot}$ is the "Kronecker delta." It is easy to check that, for all $A \in \mathcal{B}(H) = \mathcal{L}(H)$, the quantity

$$\mathrm{Tr}(A) \doteq \sum_{n=1}^{N} \langle \mathrm{e}_n, A(\mathrm{e}_n) \rangle \in \mathbb{C} \ ,$$

known as the "trace" of A, does not depend on the particular choice of the Hilbert basis $\{\mathrm{e}_n\}_{n=1}^N$. Another important property of traces are their "cyclicity" :

$$\mathrm{Tr}(AA') = \mathrm{Tr}(A'A)$$

for all $A, A' \in \mathcal{B}(H)$. For more details, see for instance [52, Chapter VI, Section 6].

[8] Distinct points are separated by disjoint neighborhoods.

For every $A \in \mathcal{B}(H)$, we define the linear functional $\varphi_A \in \mathcal{B}(H)'$ by

$$\varphi_A(A') \doteq \mathrm{Tr}(A^* A') = \mathrm{Tr}(A' A^*) , \qquad A' \in \mathcal{B}(H) .$$

We will show below that the mapping $A \mapsto \varphi_A$ defines an equivalence of ordered vector spaces (i.e., a bipositive bijective linear transformation) $\mathrm{Re}\{\mathcal{B}(H)\} \rightarrow \mathrm{Re}\{\mathcal{B}(H)'\}$. In other words, in the finite-dimensional case, Hermitian linear functions on $\mathcal{B}(H)$ and selfadjoint elements of $\mathcal{B}(H)$ can be canonically identified with each other, as elements of ordered vector spaces. (The equivalence of $\mathrm{Re}\{\mathcal{B}(H)\}$ and $\mathrm{Re}\{\mathcal{B}(H)'\}$ as vector spaces is of course trivial, for they has the same finite dimension.) In the case of states, this identification leads to the notion of "density matrices," which are sometimes formally taken in theoretical physics for the states themselves.

Definition 2.35 (Hilbert-Schmidt Norm) Let $\{\mathrm{e}_n\}_{n=1}^N$ be any Hilbert basis of H. For all $A \in \mathcal{B}(H)$, define

$$\|A\|_{\mathrm{HS}} \doteq \sqrt{\sum_{n=1}^{N} \|A(\mathrm{e}_n)\|^2} = \sqrt{\mathrm{Tr}(A^* A)} \in \mathbb{R}_0^+ .$$

$\|A\|_{\mathrm{HS}}$ is known as the "Hilbert-Schmidt norm" of A.

Observe that the norm $\|\cdot\|_{\mathrm{HS}}$ does not depend on the particular choice of the basis $\{\mathrm{e}_n\}_{n=1}^N$ in its definition. It results from a scalar product.

Lemma 2.36 *The unique scalar product $\langle \cdot, \cdot \rangle_{\mathrm{HS}}$ associated with the Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}}$ is*

$$\left\langle A, A' \right\rangle_{\mathrm{HS}} = \mathrm{Tr}(A^* A') , \qquad A, A' \in \mathcal{B}(H) .$$

In other words, $\langle \cdot, \cdot \rangle_{\mathrm{HS}}$ is the GNS scalar product associated with the positive functional $\mathrm{Tr}(\cdot) \in \mathcal{B}(H)'^{+}$ *(see Proposition 2.7).*

Proof The linear functional $\mathrm{Tr}(\cdot) \in \mathcal{B}(H)'$ is clearly positive, being a sum of positive multiples of vector states. From the identity

$$\|A\|_{\mathrm{HS}}^2 = \sum_{n=1}^{N} \|A(\mathrm{e}_n)\|^2 = \sum_{n=1}^{N} \langle A(\mathrm{e}_n), A(\mathrm{e}_n) \rangle = \sum_{n=1}^{N} \left\langle \mathrm{e}_n, A^* A(\mathrm{e}_n) \right\rangle = \mathrm{Tr}(A^* A)$$

and the polarization identity one arrives at the assertion. □

Exercise 2.37 Show that $(\mathcal{B}(H), \|\cdot\|_{\mathrm{HS}})$ is a (complex) Hilbert space.

Corollary 2.38 *For every $\varphi \in \mathcal{B}(H)'$, there is a unique $D_\varphi \in \mathcal{B}(H)$ such that $\varphi = \varphi_{D_\varphi}$. The linear functional φ is Hermitian iff D_φ is selfadjoint, in which case one has $\mathrm{Tr}(D_\varphi) = \varphi(\mathrm{id}_H)$. In particular, the mapping $A \mapsto \varphi_A$ from $\mathrm{Re}\{\mathcal{B}(H)\}$ to $\mathrm{Re}\{\mathcal{B}(H)'\}$ is (real) linear and bijective, i.e., it is a equivalence of (real) vector spaces.*

Proof

1. Take any linear functional $\varphi \in \mathcal{B}(H)'$. Observing from the finite dimensionality of $\mathcal{B}(H)$ that

$$\mathcal{B}(H)' = \mathcal{B}((\mathcal{B}(H), \|\cdot\|_{\mathrm{HS}}); \mathbb{C}) ,$$

   the existence and uniqueness of $D_\varphi \in \mathcal{B}(H)$ such that $\varphi = \varphi_{D_\varphi}$ is a direct consequence of the Riesz-Fréchet theorem (Theorem 7.214) applied to the Hilbert space $(\mathcal{B}(H), \|\cdot\|_{\mathrm{HS}})$.
2. If $\varphi \in \mathcal{B}(H)'$ is Hermitian, i.e., $\varphi(A) = \overline{\varphi(A^*)}$, then, for all $A \in \mathcal{B}(H)$

$$\mathrm{Tr}(D_\varphi^* A) = \varphi(A) = \overline{\varphi(A^*)} = \overline{\mathrm{Tr}(D_\varphi^* A^*)} = \overline{\overline{\mathrm{Tr}(A^{**} D_\varphi^{**})}} = \mathrm{Tr}((D_\varphi^*)^* A) .$$

   Note that we used the Hermiticity (which follows from the positivity) and cyclicity of the trace $\mathrm{Tr}(\cdot)$. By uniqueness of D_φ, one has in this case that $D_\varphi = D_\varphi^*$.
3. It is easy to see, again from the Hermiticity and cyclicity of the trace, that any $A \in \mathcal{B}(H)^{\mathbb{R}}$ (i.e., any selfadjoint $A \in \mathcal{B}(H)$) defines a Hermitian linear functional ρ_A. Hence, $\varphi \in \mathrm{Re}\{\mathcal{B}(H)'\}$ iff $D_\varphi \in \mathcal{B}(H)^{\mathbb{R}}$. Additionally, in this case one has

$$\mathrm{Tr}(D_\varphi) = \mathrm{Tr}(D_\varphi^* \mathrm{id}_H) = \varphi(\mathrm{id}_H) .$$

□

In the following we will prove that the mapping $A \mapsto \varphi_A$ from $\mathrm{Re}\{\mathcal{B}(H)\}$ to $\mathrm{Re}\{\mathcal{B}(H)'\}$ preserves and reflects positivity.

Exercise 2.39 Show that, for any subspace $G \subseteq H$, $\mathrm{Tr}(P_G) = \dim(G)$.

Lemma 2.40 *For every $A \in \mathcal{B}(H)$, $\varphi_A \in \mathcal{B}(H)'$ is a positive linear functional iff A is positive.*

Proof

1. Take any positive $A \in \mathcal{B}(H)^+$. Then, the spectrum $\sigma(A)$ is positive. In particular, all eigenvalues of A are nonnegative real numbers. Recall that any selfadjoint operator on a finite-dimensional complex Hilbert H space has an eigenbasis, i.e., a Hilbert basis of eigenvectors of H. Thus, let $\{\mathrm{e}_n\}_{n=1}^N$ be any eigenbasis of A.

Then, for all $A' \in \mathcal{B}(H)^+$, one has

$$\varphi_A(A') = \mathrm{Tr}(A'A) = \sum_{n=1}^{N} \left\langle \mathrm{e}_n, A'(A(\mathrm{e}_n)) \right\rangle = \sum_{n=1}^{N} \lambda_n \left\langle \mathrm{e}_n, A'(\mathrm{e}_n) \right\rangle \geq 0 ,$$

where $\lambda_n \geq 0$ is the eigenvalue of A associated with the eigenvector e_n. Hence, $\varphi_A \in \mathcal{B}(H)'$ is a positive linear functional, whenever $A \in \mathcal{B}(H)$ is positive.

2. Take now any $A \in \mathcal{B}(H)$ and assume that φ_A is a positive linear functional. Then, by the last corollary, A is selfadjoint. In particular, the spectrum $\sigma(A)$ is real. Recall that, as H has finite dimension, every point in the spectrum $\sigma(A)$ is an eigenvalue of A.
3. Let $\lambda \in \mathbb{R}$ be any eigenvalue of A and $x_\lambda \in H \backslash \{0\}$ any corresponding eigenvector. Then,

$$AP_{\mathrm{span}(\{x_\lambda\})} = \lambda P_{\mathrm{span}(\{x_\lambda\})} .$$

Hence, by the last exercise, as orthogonal projectors are positive, one has

$$0 \leq \varphi_A(P_{\mathrm{span}(\{x_\lambda\})}) = \mathrm{Tr}(AP_{\mathrm{span}(\{x_\lambda\})}) = \lambda \mathrm{Tr}(P_{\mathrm{span}(\{x_\lambda\})}) = \lambda .$$

As A is selfadjoint, this implies that A is positive.

□

These last results motivate the following definition of so-called density matrices.

Definition 2.41 (Density Matrices) We say that $D \in \mathcal{B}(H)$ is a "density matrix" if $\mathrm{Tr}(D) = 1$ and $D \geq 0$. $\mathcal{D}(H) \subseteq \mathcal{B}(H)^+$ denotes the set of all density matrices in $\mathcal{B}(H)$.

Observe from Lemma 2.40 that $D \in \mathcal{B}(H)$ is a density matrix iff φ_D is a positive linear functional and $\varphi_D(\mathrm{id}_H) = 1$, that is, iff φ_D is a state. In other words, the following assertion holds true:

Corollary 2.42 *The mapping $D \mapsto \varphi_D$ defines a one-to-one correspondence $\mathcal{D}(H) \to E(H)$, which preserves and reflects order.*

Exercise 2.43 Show that $\rho \in E(H)$ is an extreme state on $\mathcal{B}(H)$ iff it is a vector state.

Hint: (i) Note that ρ is a vector state iff its density matrix is an orthogonal projector whose range is one-dimensional. (ii) Observe that, for all $\lambda \in [0, 1]$ and any density matrices D, D', one has

$$\varphi_{\lambda D + (1-\lambda) D'} = \lambda \varphi_D + (1 - \lambda) \varphi_{D'} .$$

Chapter 3
Thermodynamic Equilibrium of Finite Quantum Systems

3.1 Gibbs States

In the present section we discuss the theory of equilibrium states for quantum systems whose associated C^*-algebra is some finite matrix algebra $\mathcal{L}(\mathbb{C}^n) = \mathcal{B}(\mathbb{C}^n)$, $n \in \mathbb{N}$. Here, $\mathbb{C}^n$ is (canonically) the n-dimensional complex Hilbert space, the norm of which is the Euclidean one.

Definition 3.1 (Functional Calculus for Selfadjoint Matrices) Fix $n \in \mathbb{N}$ and $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. The eigenvalues[1] of A are denoted by $\sigma_1, \ldots, \sigma_m \in \mathbb{R}$, $m \leq n$, and, for all $k = 1, \ldots, m$, the eigenspace associated with the eigenvalue σ_k is denoted by

$$E_k \doteq \{x \ : \ A(x) = \sigma_k x\} \subseteq \mathbb{C}^n \ .$$

Then, for any function $f : \sigma(A) \to \mathbb{C}$, we define

$$f(A) \doteq f(\sigma_1) P_{E_1} + \cdots + f(\sigma_m) P_{E_m} \in \mathcal{B}(\mathbb{C}^n) = \mathcal{L}(\mathbb{C}^n) \ .$$

The mapping $f \mapsto f(A)$ from $\mathcal{F}(\sigma(A); \mathbb{C})$ to $\mathcal{L}(\mathbb{C}^n)$ is called the "functional calculus" for A.

Recall that $\mathcal{F}(\sigma(A); \mathbb{C})$ stands for the space of all functions $\sigma(A) \to \mathbb{C}$. By construction, note that if $x \in \mathbb{C}^n$ is an eigenvector of $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ for the eigenvalue $\sigma_1 \in \sigma(A)$, then, for all $f \in \mathcal{F}(\sigma(A); \mathbb{C})$, $f(\sigma_1) \in \mathbb{C}$ is an eigenvalue of $f(A)$ and

[1] In particular, $\sigma(A) = \{\sigma_1, \ldots, \sigma_m\}$.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_3

x is an eigenvector corresponding to this eigenvalue. In fact, it is easy to see that

$$\sigma(f(A)) = \{f(\sigma_1) \ : \ \sigma_1 \in \sigma(A)\} .$$

We gather other here relevant properties of the functional calculus for selfadjoint matrices in the following exercises.

Exercise 3.2 Prove that, for any $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, the functional calculus $f \mapsto f(A)$ is a self-conjugate algebra homomorphism (i.e., a $*$-homomorphism) from $\mathcal{F}(\sigma(A);\mathbb{C})$ to $\mathcal{L}(\mathbb{C}^n)$, for which $1(A) = \mathrm{id}_{\mathbb{C}^n}$ and $\mathrm{id}_{\sigma(A)}(A) = A$, where $\mathrm{id}_{\sigma(A)} \in \mathcal{F}(\sigma(A);\mathbb{C})$ is the identity mapping on $\sigma(A) \subseteq \mathbb{C}$. In particular, for any polynomial $f(s) = a_0 + a_1 s + \cdots + a_m s^m$, with $m \in \mathbb{N}$ and $a_0, \ldots, a_m \in \mathbb{C}$,

$$f(A) = a_0 \mathrm{id}_{\mathbb{C}^n} + a_1 A + \cdots + a_m A^m ,$$

where $A^k \doteq A A^{k-1}$ for $k \in \{2, \ldots, m\}$.

Exercise 3.3 Prove that, for any $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, any unitary $U \in \mathcal{L}(\mathbb{C}^n)$ and any function $f : \sigma(A) = \sigma(UAU^*) \to \mathbb{C}$, one has

$$Uf(A)U^* = f(UAU^*) .$$

Exercise 3.4 Prove that, for any $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $z \in \mathbb{C}$

$$\exp(zA) = \mathrm{id}_{\mathbb{C}^n} + \sum_{k=1}^{\infty} \frac{z^k}{k!} A^k .$$

Show, moreover, that $\exp(zA)$ is unitary when $z \in i\mathbb{R}$. (As $\mathcal{L}(\mathbb{C}^n)$ has finite dimension, the convergence of the above series refers to any norm in $\mathcal{L}(\mathbb{C}^n)$).

Exercise 3.5 Prove, by using the functional calculus, that $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ is positive iff there is $B \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ such that $A = B^2$.

We now introduce energy observables of finite quantum systems, which are usually called Hamiltonians, as well their associated equilibrium states.

Definition 3.6 (Hamiltonian) Fix $n \in \mathbb{N}$. The selfadjoint element $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ associated with the total energy of the physical system is called "Hamiltonian" of the system.

In principle, any selfadjoint element of $\mathcal{L}(\mathbb{C}^n)$ could be the Hamiltonian of some physical system.

Recall from Corollary 2.42 that, for any $n \in \mathbb{N}$ and every state $\rho \in E(\mathbb{C}^n)$, there is a unique density matrix $D_\rho \in \mathcal{L}(\mathbb{C}^n)^+$ such that

$$\rho(A) = \mathrm{Tr}(D_\rho A)$$

for all $A \in \mathcal{L}(\mathbb{C}^n)$. Conversely, for any $n \in \mathbb{N}$ and all $D \in \mathcal{L}(\mathbb{C}^n)^+$ such that $\mathrm{Tr}(D_\rho) = 1$, i.e., for any density matrix, there is a state $\rho_D \in E(\mathbb{C}^n)$ such that

$$\rho_D(A) = \mathrm{Tr}(DA)$$

for all $A \in \mathcal{L}(\mathbb{C}^n)$.

For any given Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$, we define its associated Gibbs state as follows.

Definition 3.7 (Gibbs State) Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$. For all $\beta \in (0, \infty)$, we define the density matrix:

$$D_{H,\beta} \doteq Z_{H,\beta}^{-1} \exp(-\beta H) ,$$

where

$$Z_{H,\beta} \doteq \mathrm{Tr} \exp(-\beta H) > 0 .$$

The state $\omega_{H,\beta} \in E(\mathbb{C}^n)$ whose density matrix is $D_{H,\beta}$ is called the "Gibbs state" associated with the Hamiltonian H, at temperature $T = \beta^{-1}$.

$\exp(-\beta H)$ is a non-zero positive element of $\mathcal{L}(\mathbb{C}^n)$, and thus $D_{H,\beta}$ is a density matrix, because the image of the exponential function $\exp(\cdot) : \mathbb{R} \to \mathbb{R}$ is strictly positive. In fact, note that $\exp(-\beta H) \in \mathcal{L}(\mathbb{C}^n)$ is selfadjoint, because the exponential function $\exp(\cdot)$ is real-valued and the functional calculus is a $*$-homomorphism from $\mathcal{F}(\sigma(A); \mathbb{C})$ to $\mathcal{L}(\mathbb{C}^n)$. Thus, $\exp(-\beta H)$ is positive, its spectrum being positive.

In quantum statistical physics, the Gibbs state $\omega_{H,\beta}$ is seen, for various reasons, as the thermal equilibrium state of the quantum system at temperature $T = \beta^{-1}$, the Hamiltonian of which is H. In the following we show that the Gibbs states are completely determined by some of their properties. In other words, we will *equivalently define* the thermal equilibrium states of *finite* quantum systems via different particular properties of Gibbs states. In fact, in the case of infinitely extended systems, the Gibbs states, as define above, have, in many important cases, no clear mathematical meaning, whereas various of their properties do have. This is how reasonable mathematical *definitions* for the thermodynamical equilibrium of infinitely extended quantum systems were found. In this section we discuss the most important properties of Gibbs states that turned out to be relevant for the mathematical foundations of quantum statistical mechanics. We divide them into two types: statical and dynamical properties. For technical simplicity, later on we will restrict ourselves to the *statical* properties only, to go deeper into the mathematical foundations of thermodynamic equilibrium of infinite systems, in which first-order phase transitions are in particular concerned.

3.2 Statical Characterizations of Gibbs States

In quantum statistical mechanics, a quantum system at fixed temperature is not only characterized by its energy but also by its entropy, which is defined from the functional calculus as follows.

Definition 3.8 (von Neumann Entropy) Define the continuous function $\eta : [0, 1] \to \mathbb{R}_0^+$ by

$$\eta(s) \doteq -s \ln(s), \ s > 0, \quad \eta(0) \doteq 0 .$$

Let $\rho \in E(\mathbb{C}^n)$, $n \in \mathbb{N}$, and $D_\rho \in \mathcal{L}(\mathbb{C}^n)^+$ be the associated density matrix. The quantity

$$S(\rho) \doteq \mathrm{Tr}(\eta(D_\rho)) \in [0, \ln(n)]$$

is called the "von Neumann entropy" of ρ.

As $\sigma(D_\rho) \subseteq [0, 1]$ for any state $\rho \in E(\mathbb{C}^n)$, $\eta(D_\rho) \in \mathcal{L}(\mathbb{C}^n)$ is well-defined by the functional calculus of Definition 3.1. In fact, observe that the eigenvalues $p_1, \ldots, p_n$ (with multiplicity) of the density matrix D_ρ of any state $\rho \in E(\mathbb{C}^n)$ satisfy $p_i \geq 0$ and $p_1 + \cdots + p_n = 1$. That is, the sequence of numbers $p_1, \ldots, p_n$ defines a probability distribution in the finite set $\{1, 2, \ldots, n\}$. With this remark, the von Neumann entropy of ρ can be seen as the Shannon entropy of a classical probability distribution:

$$S(\rho) = - \sum_{i=1,\ p_i>0}^{n} p_i \ln(p_i) .$$

For this reason, one usually interprets the quantity $S(\rho)$ as the amount of "disorder" for the system in the state ρ.

In a sense that will be made precise below, the state of a quantum system in thermal equilibrium minimizes the energy while it maximizes the entropy, by the second law of thermodynamics. In fact, such a state minimizes the so-called free energy, which is defined at fixed temperature as follows.

Definition 3.9 (Free Energy) Fix $n \in \mathbb{N}$. Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$. For every state $\rho \in E(\mathbb{C}^n)$, we define the "free energy" associated with the Hamiltonian H at inverse temperature β by

$$F_{H,\beta}(\rho) \doteq \rho(H) - \beta^{-1} S(\rho) .$$

In the following we show that Gibbs states are free energy minimizers and that the mapping $\rho \mapsto F_{H,\beta}(\rho)$ from $E(\mathbb{C}^n)$ to $\mathbb{R}$ is convex. To this end, we first derive the well-known Peierls-Bogoliubov inequality.

Lemma 3.10 (Peierls-Bogoliubov Inequality) *Fix $n \in \mathbb{N}$. Let $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $f : \mathbb{R} \to \mathbb{R}$ be any convex function. For any Hilbert basis $\{\mathrm{e}_k\}_{k=1}^n$ of $\mathbb{C}^n$,*

$$\sum_{k=1}^{n} f(\langle \mathrm{e}_k, A(\mathrm{e}_k)\rangle) \leq \sum_{k=1}^{n} \langle \mathrm{e}_k, f(A)(\mathrm{e}_k)\rangle = \mathrm{Tr} f(A).$$

Proof

1. It suffices to show that

$$f(\langle x, A(x)\rangle) \leq \langle x, f(A)(x)\rangle$$

   for any $x \in \mathbb{C}^n$, $\|x\| = 1$ and any convex function $f : \mathbb{R} \to \mathbb{R}$.
2. Let $\{\mathrm{e}_k\}_{k=1}^n$ be any Hilbert basis of $\mathbb{C}^n$ of eigenvectors of A: $A(\mathrm{e}_k) = \sigma_k \mathrm{e}_k$ for some $\sigma_k \in \mathbb{R}$, $k = 1, \ldots, n$. Such a basis always exists and the corresponding eigenvalues are real numbers, for A is selfadjoint. Then, noting that $\{\mathrm{e}_k\}_{k=1}^n$ are also eigenvectors of the selfadjoint element $f(A) \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ with $f(A)(\mathrm{e}_k) = f(\sigma_k)\mathrm{e}_k$

$$\langle x, f(A)(x)\rangle = \sum_{k=1}^{n} a_k(x) f(\sigma_k),$$

   where $a_k(x) \doteq |\langle x, \mathrm{e}_k\rangle|^2 \geq 0$.
3. As $\|x\| = 1$, it follows that $\sum_{k=1}^{n} a_k(x) = 1$ (by Parseval's identity; see Corollary 7.226). Hence, by Jensen's inequality (Lemma 7.330), for any convex function $f : \mathbb{R} \to \mathbb{R}$ and every vector $x \in \mathbb{C}^n$ satisfying $\|x\| = 1$

$$\langle x, f(A)(x)\rangle \geq f\left(\sum_{k=1}^{n} a_k(x)\sigma_k\right) = f\left(\langle x, A(x)\rangle\right) .$$

□

Corollary 3.11 *Fix $n \in \mathbb{N}$. Let $D, D' \in \mathcal{L}(\mathbb{C}^n)^+$ be two density matrices (i.e., $\mathrm{Tr}(D) = \mathrm{Tr}(D') = 1$) and $f : [0, 1] \to \mathbb{R}$ be any convex function. Then, for all $\lambda \in [0, 1]$,*

$$\mathrm{Tr} f\left(\lambda D + (1-\lambda) D'\right) \leq \lambda \mathrm{Tr} f(D) + (1-\lambda)\mathrm{Tr} f\left(D'\right) .$$

Proof

1. Fix $n \in \mathbb{N}$, density matrices $D, D' \in \mathcal{L}(\mathbb{C}^n)^+$, a constant $\lambda \in [0, 1]$ and a convex function $f : [0, 1] \to \mathbb{R}$. Recall that the convex combination $\lambda D + (1-\lambda)D'$ is a new density matrix and that the spectrum of density matrices is always contained

in the interval [0, 1] (and thus $f(\lambda D + (1-\lambda)D')$, $f(D)$ and $f(D')$ are well-defined elements of $\mathcal{L}(\mathbb{C}^n)$, by the functional calculus).

2. Let $\mathrm{e}_1, \ldots, \mathrm{e}_n \in \mathbb{C}^n$ be an orthonormal basis of $\mathbb{C}^n$ consisting of eingenvectors of $\lambda D + (1-\lambda)D' \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. Then,

$$\begin{aligned}\mathrm{Tr} f(\lambda D + (1-\lambda)D') &= \sum_{k=1}^{n} \langle \mathrm{e}_k, f(\lambda D + (1-\lambda)D')(\mathrm{e}_k)\rangle \\ &= \sum_{k=1}^{n} f(\langle \mathrm{e}_k, (\lambda D + (1-\lambda)D')(\mathrm{e}_k)\rangle) \\ &= \sum_{k=1}^{n} f(\lambda\langle \mathrm{e}_k, D(\mathrm{e}_k)\rangle + (1-\lambda)\langle \mathrm{e}_k, D'(\mathrm{e}_k)\rangle) \ .\end{aligned}$$

3. By convexity of f

$$\mathrm{Tr} f(\lambda D + (1-\lambda)D') \leq \lambda \sum_{k=1}^{n} f(\langle \mathrm{e}_k, D(\mathrm{e}_k)\rangle) + (1-\lambda) \sum_{k=1}^{n} f(\langle \mathrm{e}_k, D'(\mathrm{e}_k)\rangle) \ .$$

Finally, by the Peierls-Bogoliubov inequality (Lemma 3.10)

$$\mathrm{Tr} f(\lambda D + (1-\lambda)D') \leq \lambda \mathrm{Tr} f(D) + (1-\lambda)\mathrm{Tr} f(D') \ .$$

□

Corollary 3.12 *Fix $n \in \mathbb{N}$. The von Neumann entropy defines a concave function from $E(\mathbb{C}^n)$ to $\mathbb{R}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and inverse temperature $\beta \in (0, \infty)$, the free energy functional $F_{H,\beta} : E(\mathbb{C}^n) \to \mathbb{R}$ is convex.*

Proof Exercise. □

We are now in a position to prove that Gibbs states are free energy minimizers.

Proposition 3.13 (Gibbs States as Free Energy Minimizers) *Fix $n \in \mathbb{N}$. For any state $\rho \in E(\mathbb{C}^n)$, every Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and inverse temperature $\beta \in (0, \infty)$,*

$$F_{H,\beta}(\rho) \geq F_{H,\beta}(\omega_{H,\beta}) = -\beta^{-1} \ln Z_{H,\beta} \ ,$$

where $\omega_{H,\beta} \in E(\mathbb{C}^n)$ is the Gibbs state (Definition 3.7) associated with the Hamiltonian H, at temperature $T = \beta^{-1}$. Here, $Z_{H,\beta}$ is the (normalization) constant of Definition 3.7.

Proof

1. Take any $\rho \in E(\mathbb{C}^n)$. Let $\{\mathrm{e}_k\}_{k=1}^n$ be a Hilbert basis of $\mathbb{C}^n$ of eigenvectors of the density matrix D_ρ. Let $p_k \geq 0$, $k = 1, \ldots, n$, be the corresponding eigenvalues, i.e., $D_\rho(\mathrm{e}_k) = p_k \mathrm{e}_k$. For simplicity, without loss of generality, we assume that $p_k > 0$ for all $k = 1, \ldots, n$:

$$\begin{aligned} -\beta F_{H,\beta}(\rho) &= -\beta\rho(H) + S(\rho) \\ &= -\beta \sum_{k=1}^n p_k \langle \mathrm{e}_k, H(\mathrm{e}_k)\rangle - \sum_{k=1}^n p_k \ln(p_k) \\ &= \sum_{k=1}^n p_k[-\beta\langle \mathrm{e}_k, H(\mathrm{e}_k)\rangle - \ln(p_k)] \,. \end{aligned}$$

2. In the special case $\rho = \omega_{H,\beta}$, observe that $\mathrm{e}_1, \ldots, \mathrm{e}_n$ are simultaneously eigenvectors of both H and D_ρ. In this case, for $k = 1, \ldots, n$, let h_k be the eigenvalue of H satisfying $H(\mathrm{e}_k) = h_k \mathrm{e}_k$. Then, for $\rho = \omega_{H,\beta}$,

$$p_k = Z_{H,\beta}^{-1} \exp(-\beta h_k) \,.$$

   In particular,

$$\begin{aligned} -\beta F_{H,\beta}(\omega_{H,\beta}) &= \sum_{k=1}^n p_k[-\beta h_k - (-\beta h_k - \ln Z_{H,\beta})] \\ &= \ln Z_{H,\beta} \,. \end{aligned}$$

   Recall that $p_1 + \cdots + p_n = 1$ because $\mathrm{Tr}\left(D_\rho\right) = 1$. Hence,

$$F_{H,\beta}(\omega_{H,\beta}) = -\beta^{-1} \ln Z_{H,\beta} \,.$$

3. By convexity of the function $\exp(\cdot)$ and Jensen's inequality (Lemma 7.330), for all $\rho \in E(\mathbb{C}^n)$, it follows that

$$\begin{aligned} \exp(-\beta F_{H,\beta}(\rho)) &\leq \sum_{k=1}^n p_k \exp[-\beta\langle \mathrm{e}_k, H(\mathrm{e}_k)\rangle - \ln(p_k)] \\ &= \sum_{k=1}^n p_k p_k^{-1} \exp(\langle \mathrm{e}_k, -\beta H(\mathrm{e}_k)\rangle) \,. \end{aligned}$$

Hence, by the Peierls-Bogoliubov inequality

$$\begin{aligned}\exp(-\beta F_{H,\beta}(\rho)) &\leq \sum_{k=1}^{n} \exp(\langle \mathrm{e}_k, -\beta H(\mathrm{e}_k)\rangle) \\ &\leq \mathrm{Tr}\exp(-\beta H) = \exp(\ln(\mathrm{Tr}\exp(-\beta H))) \ .\end{aligned}$$

4. Observing that $\exp(\cdot)$ is a strictly increasing function, one finally arrives at

$$F_{H,\beta}(\rho) \geq -\beta^{-1} \ln Z_{H,\beta} \ .$$

□

Remark 3.14 (Maximum Entropy Principle) Fix $n \in \mathbb{N}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and inverse temperature $\beta \in (0, \infty)$, define the expected value of the energy at the Gibbs state by $E_{H,\beta} \doteq \omega_{H,\beta}(H)$. Then,

$$\begin{aligned}\inf_{\rho \in E(\mathbb{C}^n)} \beta F_{H,\beta}(\rho) &= \inf_{\rho \in E(\mathbb{C}^n)} (\beta\rho(H) - S(\rho)) \\ &= \inf_{\rho \in E(\mathbb{C}^n),\ \rho(H)=E_{H,\beta}} (\beta\rho(H) - S(\rho)) \\ &= \beta E_{H,\beta} - \sup_{\rho \in E(\mathbb{C}^n),\ \rho(H)=E_{H,\beta}} S(\rho) \\ &= \beta E_{H,\beta} - S(\omega_{H,\beta}) \ .\end{aligned}$$

In other words, $\omega_{H,\beta}$ maximizes the von Neumann entropy in the set:

$$\{\rho \ : \ \rho(H) = E_{H,\beta}\} \subseteq E(\mathbb{C}^n)$$

of all states whose associated energy is exactly $E_{H,\beta}$. β is the "Lagrange multiplier" for the corresponding optimization problem with constraint.

Exercise 3.15 Show that, for any $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$,

$$\lim_{\beta \to 0} E_{H,\beta} = \frac{1}{n}\mathrm{Tr}\ H \geq \min \sigma(H) \ , \qquad \lim_{\beta \to \infty} E_{H,\beta} = \min \sigma(H) \ .$$

Deduce from this that, for all $E \in (\min \sigma(H), \frac{1}{n}\mathrm{Tr}\ H)$, there is $\beta \in (0, \beta)$ such that $E_{H,\beta} = E$.

In the following we show that Gibbs states can be seen as so-called tangent functionals to some convex function, which is called "pressure" of the corresponding quantum system. In particular, we will demonstrate from this property that the Gibbs states are the *unique* minimizers of the corresponding free energy functionals. Such an approach was initially proposed by Gibbs and turned out to be very fruitful for thermodynamics and statistical mechanics, also in the case of infinitely extended systems.

Definition 3.16 (Pressure) Fix $n \in \mathbb{N}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and inverse temperature $\beta \in (0, \infty)$, the "pressure" associated with the Hamiltonian H at temperature $T = \beta^{-1}$ is defined by the quantity:

$$P_{H,\beta} \doteq \beta^{-1} \ln Z_{H,\beta} \ .$$

The mapping $\mathcal{L}(\mathbb{C}^n)^{\mathbb{R}} \to \mathbb{R}$, $H \mapsto P_{H,\beta}$, is called "pressure function" at temperature $T = \beta^{-1}$.

The pressure function at any given temperature is a convex function of the Hamiltonian, as it is shown below. In particular it is continuous (for any convex function on a normed space of finite dimension is continuous), as a function of the energy observable.

Proposition 3.17 (Convexity of the Pressure Function) *For all $\beta \in (0, \infty)$, the mapping $H \mapsto P_{H,\beta}$ is convex, i.e., for all $\lambda \in [0, 1]$ and $H, H' \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$:*

$$P_{\lambda H+(1-\lambda)H',\beta} \leq \lambda P_{H,\beta} + (1-\lambda) P_{H',\beta} \ .$$

Proof For all $\rho \in E(\mathbb{C}^n)$, we define the (affine) mapping $a_\rho : \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}} \to \mathbb{R}$ by

$$a_\rho(H) \doteq -\rho(H) + \beta^{-1} S(\rho) = -F_{H,\beta}(\rho) \ .$$

By Theorem 3.13, for all $\lambda \in [0, 1]$ and $H, H' \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$,

$$\begin{aligned} P_{\lambda H+(1-\lambda)H',\beta} &= \sup_{\rho \in E(\mathbb{C}^n)} a_\rho(\lambda H + (1-\lambda)H') \\ &= \sup_{\rho \in E(\mathbb{C}^n)} (\lambda a_\rho(H) + (1-\lambda) a_\rho(H')) \\ &\leq \lambda \sup_{\rho \in E(\mathbb{C}^n)} a_\rho(H) + (1-\lambda) \sup_{\rho \in E(\mathbb{C}^n)} a_\rho(H') \\ &= \lambda P_{H,\beta} + (1-\lambda) P_{H',\beta} \ . \end{aligned}$$

□

Another important notion related with convex functions like the pressure function is the concept of tangent functionals on real linear spaces, which refers to the following definition.

Definition 3.18 (Tangent Functionals) Let V be a vector space over $\mathbb{R}$ and $g : V \to \mathbb{R}$ any function. We say that the linear functional $\varphi \in V'$ is "tangent to g at $v \in V$" if, for all $v' \in V$:

$$g(v + v') - g(v) \geq \varphi(v') \ .$$

If X is a separable real Banach space and g is convex and continuous, then it is well-known that g has, on each point $x \in X$, at least one continuous tangent functional, thanks to the Mazur theorem [53] (see Proposition 3.21 (iii)) and Lanford III-Robinson theorem [54, Theorem 1]. This will be discussed in more detail later on. Here, we first apply this notion of tangent functionals to the (convex) pressure function.

Lemma 3.19 *Fix $n \in \mathbb{N}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$, if $\rho \in E(\mathbb{C}^n)$ is a minimizer of the free energy $F_{H,\beta}$, then the linear functional $-\rho \in \mathcal{L}(\mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}; \mathbb{R})$ is tangent to the pressure function $\tilde{H} \mapsto P_{\tilde{H},\beta}$ at $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. (Recall that we canonically identify the real vector spaces* $\mathrm{Re}\{\mathcal{L}(\mathbb{C}^n)\}$ *and* $\mathcal{L}(\mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}; \mathbb{R})$ *and that* $E(\mathbb{C}^n) \subseteq \mathrm{Re}\{\mathcal{L}(\mathbb{C}^n)\}$.*)*

Proof As ρ minimizes $F_{H,\beta}$, by Proposition 3.13, for all $\tilde{H} \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$

$$\begin{aligned} P_{H+\tilde{H},\beta} - P_{H,\beta} &= - \inf_{\rho' \in E(\mathbb{C}^n)} F_{H+\tilde{H},\beta}(\rho') + F_{H,\beta}(\rho) \\ &\geq -F_{H+\tilde{H},\beta}(\rho) + F_{H,\beta}(\rho) \\ &= -\rho(\tilde{H}) - F_{H,\beta}(\rho) + F_{H,\beta}(\rho) \, . \end{aligned}$$

□

Corollary 3.20 (Gibbs States as Tangent Functionals) *Fix $n \in \mathbb{N}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and inverse temperature $\beta \in (0, \infty)$, the linear functional $-\omega_{H,\beta} \in \mathcal{L}(\mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}; \mathbb{R})$ is tangent to the pressure function $\tilde{H} \mapsto P_{\tilde{H},\beta}$ at $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$.*

Proof Combine Lemma 3.19 with Proposition 3.13. □

Let V be a vector space over $\mathbb{R}$ and $g : V \to \mathbb{R}$ any mapping. We say that g is differentiable at $v \in V$ in the "Gateaux sense" if, for all $v' \in V$, the limit

$$\lim_{\alpha \to 0} \frac{g(v + \alpha v') - g(v)}{\alpha} \doteq \mathrm{d}g(v)(v')$$

exists. In this case, the mapping $\mathrm{d}g(v) : V \to \mathbb{R}$, $v' \mapsto \mathrm{d}g(v)(v')$, is called the "Gateaux derivative" of g at v. Note that this mapping is positively homogeneous, i.e.,

$$\mathrm{d}g(v)(\lambda v') = \lambda \mathrm{d}g(v)(v')$$

for all $v' \in V$ and $\lambda \geq 0$, but it is *generally not linear*. In the special case of g being a *convex* function, the Gateaux derivative $\mathrm{d}g(v) : V \to \mathbb{R}$, $v \in V$, is *linear*, i.e., $\mathrm{d}g(v) \in V'$, when it exists. If X is a *complete normed* space (i.e., a Banach space) and $g : X \to \mathbb{R}$ a *convex continuous* function, then for all $x \in X$, $\mathrm{d}g(x) \in X^{\mathrm{td}}$, i.e., the Gateaux derivative of g at x is *linear and continuous*.

Let X be a normed space over $\mathbb{R}$ and $g : X \to \mathbb{R}$ any mapping. We say that g is differentiable at $x \in X$ in the "Fréchet sense" if there is $\mathrm{d}^F g(x) \in X^{\mathrm{td}}$ such that

$$\frac{g(x + x') - g(x) - [\mathrm{d}^F g(x)](x')}{\|x'\|} \to 0$$

as $x' \to 0$, $x' \neq 0$, in the normed space X. In this case, $\mathrm{d}^F g(x) \in X^{\mathrm{td}}$ is the "Fréchet derivative of g at x." It is apparent that $\mathrm{d}g(x) = \mathrm{d}^F g(x)$, whenever g is Fréchet differentiable at $x \in X$.

Gateaux derivatives in relation with tangent functionals have the following important properties in the Banach space case.

Proposition 3.21 *Let X be any* real Banach *space and $g : X \to \mathbb{R}$ an arbitrary* convex continuous *function. Then, the following assertions hold true:*

(i) *For all $x \in X$, there is at least one continuous functional $\varphi \in X^{\mathrm{td}}$, which is tangent to g at x.*
(ii) *g is differentiable at $x \in X$ in the Gateaux sense iff there is a unique continuous tangent functional to g at x, in which case* $\mathrm{d}g(x) \in X^{\mathrm{td}}$.
(iii) *(Mazur) If X is separable, then the set X_0 of all points at which the function g is differentiable in the Gateaux sense is dense in X. This refers to the so-called Asplund property of separable Banach spaces.*

For complete proofs of the three assertions of the last proposition, we recommend [12].

Exercise 3.22 Prove the "only if" part of the second point of the above proposition, i.e., that there is a unique continuous tangent functional to g at $x \in X$, whenever g has a Gateaux derivative in x.

The last proposition implies that the Gibbs state is the unique minimizer of the corresponding free energy functional.

Lemma 3.23 *Fix $n \in \mathbb{N}$. For all $H, H' \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$,*

$$P_{H+H',\beta} = P_{H,\beta} - \omega_{H,\beta}(H') + O(\|H'\|_{\mathrm{op}}^2) \, .$$

Proof Let $f(H') \doteq \mathrm{Tr}\exp(-\beta(H + H'))$. Then, by Definition 3.16, $P_{H+H',\beta} = \beta^{-1} \ln f(H')$. By using the cyclicity of the trace $\mathrm{Tr}(\cdot)$, one arrives at

$$\begin{aligned} f(H') &= \mathrm{Tr}\left(\sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!}(H + H')^n\right) \\ &= \sum_{k=0}^{\infty} \frac{(-\beta)^k}{k!} \mathrm{Tr}\left((H + H')^k\right) \end{aligned}$$

$$= f(0) - \beta \sum_{k=1}^{\infty} \mathrm{Tr}\left(k \frac{(-\beta)^{k-1}}{k!} H^{k-1} H' \right) + O(\|H'\|_{\mathrm{op}}^2)$$

$$= f(0) - \beta \mathrm{Tr}(\exp(-\beta H) H') + O(\|H'\|_{\mathrm{op}}^2) \; .$$

Hence,

$$f(H') - f(0) = -\beta Z_{H,\beta} \, \omega_{H,\beta}(H') + O(\|H'\|_{\mathrm{op}}^2) \; .$$

The lemma then follows by using the Taylor expansion of $\ln(x)$ around $x = f(0) = Z_{H,\beta}$, that is,

$$\ln(s) = \ln(Z_{H,\beta}) + Z_{H,\beta}^{-1}(s - f(0)) + O((s - f(0))^2) \; .$$

□

Corollary 3.24 *Fix $n \in \mathbb{N}$. The pressure function $H \mapsto P_{H,\beta}$ is differentiable in the sense of Fréchet on the whole space $\mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and, for all $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$,*

$$\mathrm{d}^F P_{H,\beta} = -\omega_{H,\beta} \; .$$

In particular, $-\omega_{H,\beta}$ is the unique linear (continuous) functional tangent to the pressure function at H, and, thus, $\omega_{H,\beta}$ is the unique minimizer of the free energy functional $F_{H,\beta}$.

Proof The differentiability in the sense of Fréchet of the pressure function with the Gibbs states as unique tangent functionals is a direct consequence of Lemma 3.23. Thus, the Gibbs is the unique minimizer of the free energy functional $F_{H,\beta}$, thanks to Lemma 3.19. □

Observe from the above corollary that there is a (trivial) one-to-one correspondence between minimizers of $F_{H,\beta}$ and linear functionals tangent to $P_{(\cdot),\beta}$ at H. We will see, later on, that this also holds true for infinite quantum systems, but the minimizers of the free energy (density) will not necessarily be unique anymore. The nonuniqueness of the free energy (density) minimizer refers to the appearance of a (first-order) phase transition.

We conclude our study of the pressure function by deriving the well-known Bogoliubov (convexity) inequality. This inequality is a standard estimate of statistical mechanics. See [39, Appendix D]. In the finite-dimensional situation studied here, one usually invokes a simpler version of this inequality, which refers to the global[2] Lipschitz continuity of the (convex) pressure function.

Corollary 3.25 (Bogoliubov Inequality) *Fix $n \in \mathbb{N}$. For all $H, H' \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$:*

[2] In fact, recall that any convex function on a normed space of finite dimension is locally Lipschitz continuous.

$$|P_{H',\beta} - P_{H,\beta}| \leq \|H' - H\|_{\mathrm{op}} .$$

Proof By Corollary 3.24

$$P_{H',\beta} - P_{H,\beta} = \int_0^1 \frac{d}{ds} P_{H+s(H'-H)} \mathrm{d}s = \int_0^1 \omega_{H+s(H'-H),\beta}(H - H')\mathrm{d}s .$$

Now, to prove the Bogoliubov inequality, recall that states are norm-one linear functionals. □

3.3 Dynamical Characterizations of Gibbs States

In the following we discuss the characterization of thermodynamical equilibrium of finite quantum systems via dynamical properties.

Definition 3.26 (Heisenberg Dynamics) Fix $n \in \mathbb{N}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, define the family $\{\tau_t^H\}_{t\in\mathbb{R}}$ of linear transformations $\mathcal{L}(\mathbb{C}^n) \to \mathcal{L}(\mathbb{C}^n)$ by

$$\tau_t^H(A) \doteq \exp(itH)A\exp(-itH)$$

for all $A \in \mathcal{L}(\mathbb{C}^n)$ and $t \in \mathbb{R}$. This family is called the "Heisenberg dynamics" associated with the Hamiltonian H.

Note that this family is called a "dynamics," because it is a so-called one-parameter group, i.e., for all $t_1, t_2 \in \mathbb{R}$, one has

$$\tau_{t_1}^H \circ \tau_{t_2}^H = \tau_{t_1+t_2}^H .$$

This directly follows from basic properties of the functional calculus (Definition 3.1).

Exercise 3.27 Prove that, for any $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$, the Heisenberg dynamics $\{\tau_t^H\}_{t\in\mathbb{R}}$ is a family of algebra $*$-automorphisms of $\mathcal{L}(\mathbb{C}^n)$ for which $\tau_t^H(A)^* = \tau_t^H(A^*)$ for all $A \in \mathcal{L}(\mathbb{C}^n)$ (i.e., $\{\tau_t^H\}_{t\in\mathbb{R}}$ is a family of $*$-automorphisms of $\mathcal{L}(\mathbb{C}^n)$). Prove additionally that this dynamics is unital, i.e., $\tau_t^H(\mathrm{id}_{\mathbb{C}^n}) = \mathrm{id}_{\mathbb{C}^n}$ for all $t \in \mathbb{R}$, and preserves positivity, i.e., for all $t \in \mathbb{R}$, $\tau_t^H \in \mathcal{L}^+(\mathcal{L}(\mathbb{C}^n); \mathcal{L}(\mathbb{C}^n))$.

For all $\rho \in E(\mathbb{C}^n)$ and $t \in \mathbb{R}$ define $\rho(t) \doteq \rho \circ \tau_t^H \in \mathcal{L}(\mathbb{C}^n)'$. From the last part of the above exercise, it immediately follows that $\{\rho(t)\}_{t\in\mathbb{R}}$ is a family of states of $\mathcal{L}(\mathbb{C}^n)$. This time evolution for states corresponds to the so-called Schrödinger picture of quantum mechanics.

Exercise 3.28 Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$, and $\beta \in (0, \infty)$. Prove that the corresponding Gibbs state is stationary, i.e., for all $t \in \mathbb{R}$, $\omega_{H,\beta}(t) = \omega_{H,\beta}$.

Of course, stationarity is an expected property for equilibrium states. However, general stationary states are not equilibrium states.

Note that the Schrödinger dynamics for states preserves the von Neumann entropy (Definition 3.8). This is a direct consequence of the following lemma.

Lemma 3.29 *Given $n \in \mathbb{N}$, for any state $\rho \in E(\mathbb{C}^n)$ and unitary*[3] *$U \in \mathcal{L}(\mathbb{C}^n)$, define the functional $\rho_U \in \mathcal{L}(\mathbb{C}^n)'$ by*

$$\rho_U(A) \doteq \rho(U^* A U) , \qquad A \in \mathcal{L}(\mathbb{C}^n) .$$

$\rho_U \in E(\mathbb{C}^n)$ (i.e., ρ_U is a state) and $S(\rho_U) = S(\rho)$ (i.e., ρ_U and ρ have exactly the same von Neumann entropy).

Proof Exercise. *Hint*: Remark that $D_{\rho_U} = U D_\rho U^*$ and use Exercise 3.3. □

Corollary 3.30 *The Schrödinger dynamics for states preserves the von Neumann entropy, that is, for all $\rho \in E(\mathbb{C}^n)$ and time $t \in \mathbb{R}$, $S(\rho(t)) = S(\rho)$, where $\rho(t) = \rho \circ \tau_t^H$ for some fixed Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$.*

Because of the last result, one says that the Heisenberg dynamics τ_t^H, as defined above, corresponds to a "closed" quantum system, in contrast to a so-called "open" quantum system, whose dynamics may change the (von Neumann) entropy of the initial state.

Observe that, for any $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $z \in \mathbb{C}$, the expression $\exp(izH)$ has a precise mathematical meaning, via the functional calculus (Definition 3.1). This allows us to generalize the notion of Heisenberg dynamics to "complex times," i.e., to extend the definition of τ_t^H to any complex $t \in \mathbb{C}$.

Definition 3.31 (Complex-Time Heisenberg Dynamics) Fix $n \in \mathbb{N}$. For any Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, define the family $\{\tau_z^H\}_{z\in\mathbb{C}}$ of mappings $\mathcal{L}(\mathbb{C}^n) \to \mathcal{L}(\mathbb{C}^n)$ by

$$\tau_z^H(A) \doteq \exp(izH) A \exp(-izH)$$

for all $A \in \mathcal{L}(\mathbb{C}^n)$ and $z \in \mathbb{C}$. This family is called the "complex-time Heisenberg dynamics" associated with the Hamiltonian H.

Exercise 3.32 Prove that the complex-time Heisenberg dynamics is still a one-parameter group of unital automorphisms of the algebra $\mathcal{L}(\mathbb{C}^n)$, i.e., that $\{\tau_z^H\}_{z\in\mathbb{C}}$ is a family of $*$-automorphisms of $\mathcal{L}(\mathbb{C}^n)$ for which $\tau_{z_1}^H(\mathrm{id}_{\mathbb{C}^n}) = \mathrm{id}_{\mathbb{C}^n}$ and $\tau_{z_1}^H \circ \tau_{z_2}^H = \tau_{z_1+z_2}^H$ for all $z_1, z_2 \in \mathbb{C}$.

Because of its one-parameter group property, we call the family $\{\tau_z^H\}_{z\in\mathbb{C}}$ a "complex-time dynamics." However, this dynamics is generally *not a family of $*$-automorphisms* (i.e., $\tau_z^H(A)^* \neq \tau_z^H(A^*)$), and it *does not preserve positivity*.

[3] That is, $UU^* = U^*U = \mathrm{id}_{\mathbb{C}^n}$.

Moreover, for $\rho \in E(\mathbb{C}^n)$ and $z \in \mathbb{C}$

$$\rho_z \doteq \rho \circ \tau_z^H \in \mathcal{L}(\mathbb{C}^n)'$$

is, in general, *not* a state. In fact, the "complex-time Heisenberg dynamics" is used here to show an important property of Gibbs states, namely, the celebrated KMS (Kubo-Martin-Schwinger) condition.

Proposition 3.33 (Gibbs States Are KMS States) *Fix $n \in \mathbb{N}$. Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$. The Gibbs state $\omega_{H,\beta}$ has the KMS property, i.e., for all $A, A' \in \mathcal{L}(\mathbb{C}^n)$:*

$$\omega_{H,\beta}(A\tau_{i\beta}^H(A')) = \omega_{H,\beta}(A'A) \ .$$

Proof By the definition of the Gibbs state and cyclicity of traces, one computes from Definitions 3.7 and 3.31 that

$$\begin{aligned}
\omega_{H,\beta}(A\tau_{i\beta}^H(A')) &= Z_{H,\beta}^{-1}\mathrm{Tr}(\exp(-\beta H)A\exp(i(i\beta H)A'\exp(-i(i\beta)H)) \\
&= Z_{H,\beta}^{-1}\mathrm{Tr}(\exp(-\beta H)A\exp(-\beta H)A'\exp(\beta H)) \\
&= Z_{H,\beta}^{-1}\mathrm{Tr}(A\exp(-\beta H)A') \\
&= Z_{H,\beta}^{-1}\mathrm{Tr}(\exp(-\beta H)A'A) = \omega_{H,\beta}(A'A) \ .
\end{aligned}$$

□

This last proposition shows that Gibbs states are examples of KMS states, which are defined as follows.

Definition 3.34 (KMS State) Fix $n \in \mathbb{N}$. Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$. We say that the state $\rho \in E(\mathbb{C}^n)$ is a KMS state with respect to the Hamiltonian H at inverse temperature β if, for all $A, A' \in \mathcal{L}(\mathbb{C}^n)$, one has

$$\rho(A\tau_{i\beta}^H(A')) = \rho(A'A) \ .$$

We will show in the following that, for fixed $n \in \mathbb{N}$, $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$, the only KMS state on $\mathcal{L}(\mathbb{C}^n)$ is the corresponding Gibbs state. Thus, equilibrium can be defined by properties of the time correlation functions of the given quantum system. This is one of the main approaches to the equilibrium of infinite quantum systems in mathematical physics, following original ideas of Haag, Hugenholtz, and Winnick. Because of their elegant definition and good mathematical properties (see, e.g., [55, Sections 5.3–5.4]), KMS states certainly constitute the most popular notion of equilibrium states for infinite systems.

In order to prove that KMS states of $\mathcal{L}(\mathbb{C}^n)$ are Gibbs states, we need the following well-known property of the algebra of bounded operators on any Hilbert space.

Lemma 3.35 *Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then, the center of the algebra $\mathcal{B}(H)$ (Definition 7.23) is trivial, i.e.,*

$$\mathcal{Z}(\mathcal{B}(H)) \doteq \{A \in \mathcal{B}(H) : [A, B] \doteq AB - BA = 0 \textit{ for all } B \in \mathcal{B}(H)\} = \mathbb{K}\mathrm{id}_H \ .$$

Proof For all $x, x' \in H$, define the bounded operator $\Theta_{x,x'} \in \mathcal{B}(H)$ by

$$\Theta_{x,x'}(x'') \doteq \langle x', x'' \rangle x \ , \qquad x'' \in H \ .$$

If $A \in \mathcal{Z}(\mathcal{B}(H))$, then, by the definition of the center of an algebra, for all $x, x' \in H$, $\Theta_{x,x'} A = A\Theta_{x,x'}$, which implies that

$$\langle x', x'' \rangle A(x) = \langle x', A(x'') \rangle x \ , \qquad x, x', x'' \in H \ .$$

Thus, for $x' = x'' \in H$ with $\|x'\| = 1$, one has

$$A(x) = \langle x', A(x') \rangle x \ , \qquad x \in H \ .$$

From this we conclude that the function $x' \mapsto \langle x', A(x') \rangle$ is constant on the unit sphere:

$$S_H \doteq \{x' \in H \ : \ \|x'\| = 1\} \ .$$

This fact implies, in turn, that $A = c \ \mathrm{id}_H$ for some constant $c \in \mathbb{K}$. □

Proposition 3.36 *Fix $n \in \mathbb{N}$. Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$. If $\rho \in E(\mathbb{C}^n)$ is a KMS state with respect to H at inverse temperature β, then* $\exp(\beta H)D_\rho \in \mathcal{Z}(\mathcal{L}(\mathbb{C}^n))$.

Proof

1. Fix $n \in \mathbb{N}$. By finite dimensionality, recall that any state $\rho \in E(\mathbb{C}^n)$ is uniquely defined by its density matrix D_ρ (Definition 2.41), thanks to Corollary 2.42. In particular, for any $A \in \mathcal{L}(\mathbb{C}^n)$

$$\rho(A) = \mathrm{Tr}(D_\rho A)$$

   with D_ρ being a positive operator satisfying $\mathrm{Tr}(D_\rho) = 1$.
2. If ρ is a KMS state (with respect to H, at inverse temperature β), then, by Definition 3.34 and the previous observation, for all $A, A' \in \mathcal{L}(\mathbb{C}^n)$:

$$\begin{aligned}\rho(A\tau^H_{i\beta}(A')) &= \mathrm{Tr}(D_\rho A \exp(i(i\beta H) A' \exp(-i(i\beta)H)) \\ &= \rho(A'A) = \mathrm{Tr}(D_\rho A'A) \ .\end{aligned}$$

That is, by cyclicity of the trace and by replacing A with A^* for all $A, A' \in \mathcal{L}(\mathbb{C}^n)$,

$$\mathrm{Tr}(A^* \exp(-\beta H) A' \exp(\beta H) D_\rho) = \mathrm{Tr}(A^* D_\rho A') .$$

3. Hence, for all $A, A' \in \mathcal{L}(\mathbb{C}^n)$,

$$\left\langle A, (\exp(-\beta H) A' \exp(\beta H) D_\rho - D_\rho A') \right\rangle_{HS} = 0 ,$$

where we recall that $\langle \cdot, \cdot \rangle_{\mathrm{HS}}$ is the unique scalar product associated with the Hilbert-Schmidt norm $\|\cdot\|_{\mathrm{HS}}$; see Lemma 2.36. But this implies that, for all $A' \in \mathcal{L}(\mathbb{C}^n)$,

$$\exp(-\beta H) A' \exp(\beta H) D_\rho = D_\rho A' .$$

4. The assertion then follows by left multiplying both sides the above equality with $\exp(\beta H)$.

□

Corollary 3.37 *Fix $n \in \mathbb{N}$. Let $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $\beta \in (0, \infty)$. The Gibbs state $\omega_{H,\beta} \in E(\mathbb{C}^n)$ (Definition 3.7) is the unique KMS state with respect to H at inverse temperature β.*

Proof Exercise. □

To conclude the section, we give a brief discussion on the "passivity" property of states (on $\mathcal{L}(\mathbb{C}^n)$) and its relation to the thermodynamical equilibrium.

Definition 3.38 (Passive State) Let $n \in \mathbb{N}$ and $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. We say that the state $\rho \in E(\mathbb{C}^n)$ is "passive" with respect to the Hamiltonian H if, for every unitary $U \in \mathcal{L}(\mathbb{C}^n)$,

$$\rho(U^*[H, U]) = \rho(U^* H U - U^* U H) = \rho(U^* H U - H) \geq 0 .$$

For a general C^*-algebra, the above condition defining passivity of states only involves unitaries U, which are path connected to the unity $\mathfrak{1}$ of the algebra, that is, those algebra elements U for which there is a continuous function f_U from $[0, 1]$ to the unitaries of the given algebra such that $f_U(0) = \mathfrak{1}$ and $f_U(1) = U$. Observe that in the algebra $\mathcal{L}(\mathbb{C}^n)$, any unitary is of the form $U = \exp(iA)$ for some $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and f_U can thus be chosen as $f_U(s) = \exp(isA)$. In other words, the path connectedness is trivially fulfilled by each unitary of $\mathcal{L}(\mathbb{C}^n)$, and this condition is consequently not included in the above definition.

Note at this point that, for any fixed Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, the corresponding passive states form a convex subset of $E(\mathbb{C}^n)$. Before discussing the consequences of passivity in more detail, we show first that equilibrium states are always passive.

Proposition 3.39 (Gibbs States Are Passive) *Let $n \in \mathbb{N}$ and $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. Then, for all $\beta \in (0, \infty)$, the Gibbs state $\omega_{H,\beta} \in E(\mathbb{C}^n)$ is a passive state.*

Proof For a state $\rho \in E(\mathbb{C}^n)$ minimizing the free energy $F_{H,\beta}$ (i.e., for $\rho = \omega_{H,\beta}$, thanks to Corollary 3.24) and any unitary $U \in \mathcal{L}(\mathbb{C}^n)$, one has

$$F_{H,\beta}(\rho_U) \doteq \rho_U\,(H) - \beta^{-1} S(\rho_U) \geq F_{H,\beta}(\rho) \doteq \rho\,(H) - \beta^{-1} S(\rho)\;,$$

where we recall that $\rho_U(\cdot) \doteq \rho(U^* \cdot U)$ is the state of Lemma 3.29. Noting that $S(\rho_U) = S(\rho)$ (Lemma 3.29), it follows that

$$\rho_U(H) = \rho(U^* H U) \geq \rho(H)\;,$$

that is, $\rho(U^* H U - H) \geq 0$. □

We now study the consequences of passivity in more detail, in particular the stationarity of passive states with respect to the corresponding Heisenberg dynamics. This is a consequence of the following lemma.

Lemma 3.40 *Fix $n \in \mathbb{N}$ and $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. Let the state $\rho \in E(\mathbb{C}^n)$ be passive with respect to the Hamiltonian H. Then, for every $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$*

$$\rho([H, A]) = \rho(HA - AH) = 0 \qquad \textit{and} \qquad \rho([A, H]A) = \rho(A[H, A]) \geq 0\;.$$

Proof

1. Take any $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and recall that $\exp(isA) \in \mathcal{L}(\mathbb{C}^n)$ is unitary for all $s \in \mathbb{R}$. Then, by passivity of the state $\rho \in E(\mathbb{C}^n)$, for all $s > 0$,

$$\begin{aligned} &s^{-1}\rho(\exp(-isA)[H, \exp(isA)]) \\ &\quad = s^{-1}\left(\rho(\exp(-isA)H\exp(isA)) - \rho(H)\right) \geq 0\;. \end{aligned}$$

2. By taking the limit $s \downarrow 0$, one arrives at

$$\frac{\mathrm{d}}{\mathrm{d}s}\rho(\exp(-isA)H\exp(isA))\bigg|_{s=0} = -i\rho([A, H]) \geq 0\;.$$

   Similarly, for all $s < 0$

$$\begin{aligned} -s^{-1}\rho(\exp(-isA)[H, \exp(isA)]) = (-s)^{-1}\,(\rho(\exp(-isA)H\exp(isA)) \\ -\rho(H)) \geq 0 \end{aligned}$$

   and, by taking the limit $s \uparrow 0$, one concludes that $-i\rho([A, H]) \leq 0$. By combining this with the preceding inequality, we obtain $-i\rho([A, H]) = 0$, i.e., $\rho([H, A]) = 0$.

3. Now, from this last equality and the passivity of $\rho \in E(\mathbb{C}^n)$, it follows that, for all $s > 0$,

$$s^{-2}\rho(\exp(-isA)H\exp(isA)) + is\rho([A,H]) - \rho(H) \geq 0 .$$

Taking the limit $s \downarrow 0$ yields

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\rho(\exp(-isA)H\exp(isA))\bigg|_{s=0} = -\rho\left([A,[A,H]]\right) \geq 0 .$$

4. But,

$$\begin{aligned} -[A,[A,H]] &= [[A,H],A] \\ &= (AH - HA)A - A(AH - HA) \\ &= AHA - HA^2 - A^2H + AHA \\ &= A^2H - HA^2 - 2A(AH - HA) \\ &= [A^2,H] - 2A[A,H] \\ &= 2A[H,A] - [H,A^2] . \end{aligned}$$

Hence, again by the first equality of the lemma

$$-\rho\left([A,[A,H]]\right) = \rho\left(2A[H,A]\right) - \rho\left(\left[H,A^2\right]\right) = 2\rho\left(A[H,A]\right) \geq 0 .$$

5. Finally, using the fact that states are Hermitian, we conclude that

$$\rho(A[H,A]) = \overline{\rho(A[H,A])} = \rho([H,A]^*A) = \rho([A,H]A) .$$

□

Corollary 3.41 (Passive States Are Stationary) *Fix $n \in \mathbb{N}$ and $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. Let the state $\rho \in E(\mathbb{C}^n)$ be passive with respect to the Hamiltonian H. Then, ρ is a stationary state, i.e., $\rho_t \doteq \rho \circ \tau_t^H = \rho$ for all $t \in \mathbb{R}$..*

Proof Take any $A \in \mathcal{L}(\mathbb{C}^n)$. For all times $t \in \mathbb{R}$, by Definition 3.26, one has

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\rho_t(A) &= \frac{\mathrm{d}}{\mathrm{d}t}\rho(\exp(itH)A\exp(-itH)) \\ &= i\rho([H,\exp(itH)A\exp(-itH)]) \\ &= i\rho([H,\tau_t^H(A)]) . \end{aligned}$$

As ρ is a passive state, we deduce from the last lemma that $\rho([H,\tau_t^H(A)]) = 0$. Consequently, $\rho_t(A)$ is constant in time for all $A \in \mathcal{L}(\mathbb{C}^n)$. □

Exercise 3.42 Prove that the second part of Lemma 3.40, i.e.,

$$\rho([A, H]A) = \rho(A[H, A]) \geq 0$$

for any passive state $\rho \in E(\mathbb{C}^n)$ and $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, also implies the stationarity of ρ.

Hint: Show that $\frac{\mathrm{d}}{\mathrm{d}t}\rho_t(A^2) = 0$ for arbitrary $A \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$ and $t \in \mathbb{R}$.

In the following we briefly discuss the passivity of states in the context of Carnot's notion of "cyclic processes."

Definition 3.43 (Discrete Cyclic Process) We call a "discrete cyclic process" any finite sequence of times $t_0 < t_1 < \cdots < t_m$, $m \in \mathbb{N}$, together with an associated sequence of (perturbed) Hamiltonians $H_1, \ldots, H_m \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, $n \in \mathbb{N}$. Given such a cyclic process and any fixed (unperturbed) Hamiltonian $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, for any state $\rho \in E(\mathbb{C}^n)$, we define its time evolution $(\rho_t)_{t\in\mathbb{R}}$ by

$$\rho_t \doteq \begin{cases} \rho \circ \tau_{t-t_0}^{H} & , \ t \leq t_0 \\ \rho \circ \tau_{t_1-t_0}^{H_1} \circ \cdots \circ \tau_{t_{k-1}-t_{k-2}}^{H_{k-1}} \circ \tau_{t-t_{k-1}}^{H_k} & , \ t \in (t_{k-1}, t_k] \quad \text{for} \quad k \in \{1, \ldots, m\} \ , \\ \rho \circ \tau_{t_1-t_0}^{H_1} \circ \cdots \circ \tau_{t_m-t_{m-1}}^{H_m} \circ \tau_{t-t_m}^{H} & , \ t > t_m \end{cases}$$

The physical meaning of the above definition is that for times prior to t_0 and after t_m, the time evolution of states is driven by the unperturbed Hamiltonian H, whereas between the times t_0 and t_m, the system suffers some sequence of external perturbations, represented by the perturbed Hamiltonians $H_1, \ldots, H_m$.

Proposition 3.44 *Fix $n \in \mathbb{N}$ and $H \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$. Let the state $\rho \in E(\mathbb{C}^n)$ be passive with respect to H. For any discrete cyclic process $t_0 < t_1 < \cdots < t_m$, $m \in \mathbb{N}$, $H_0, H_1, \ldots, H_m \in \mathcal{L}(\mathbb{C}^n)^{\mathbb{R}}$, one has that, for all $t \in \mathbb{R}$, $\rho_t(H) \geq \rho(H)$. In other words, there is no discrete cyclic process being able to extract energy from a system whose state is passive.*

Proof As passive states are stationary (Corollary 3.41), one has $\rho_t = \rho$ for all $t \leq t_0$, and the assertion trivially holds true in this case. So, assume that $t > t_0$ and define the operators:

$$U_t \doteq \begin{cases} \exp(-i(t - t_{k-1})H_k) \\ \exp(-i(t_{k-1} - t_{k-2})H_{k-1}) \cdots \exp(-i(t_1 - t_0)H_1) \ , \ t \in (t_{k-1}, t_k] \\ \exp(-i(t - t_m)H) \\ \exp(-i(t_m - t_{m-1})H_m) \cdots \exp(-i(t_1 - t_0)H_1) \qquad , \ t > t_m \end{cases}$$

with $k \in \{1, \ldots, m\}$. Note that $U_t \in \mathcal{L}(\mathbb{C}^n)$ is unitary (for products of unitaries are themselves unitary) and

$$\rho_t(H) - \rho(H) = \rho(U_t^* H U_t) - \rho(H) = \rho(U_t^*[H, U_t]) \ .$$

Finally, by passivity of the state ρ, we deduce that $\rho_t(H) - \rho(H) \geq 0$. □

As proven above, equilibrium states of finite quantum systems at any temperature are passive. However, not every passive state is an equilibrium state at some temperature. In fact, recall that any convex combination of passive states is again a passive state. In particular, convex combination of Gibbs states associated with the same fixed Hamiltonian, but at different temperatures, is passive. By the uniqueness of the equilibrium state, as (nontrivial) convex combinations of Gibbs states at different temperatures are generally not equal to a Gibbs state at some temperature (for the same Hamiltonian), such passive states cannot be (thermal) equilibrium states, in general.

There is a stronger notion of passivity for states, the so-called complete passivity, which is satisfied iff the given state $\rho \in E(\mathbb{C}^n)$ is a Gibbs state at some temperature, or a ground state (i.e., a minimizer of the total energy $\rho(H)$). For more details, see for instance [55, Section 5.3].

Chapter 4
Elements of C^*-Algebra Theory

4.1 Basic Notions

Algebraic quantum mechanics is an approach, starting in the 1940s (cf. the GNS construction), which takes the Heisenberg picture of quantum mechanics as the more fundamental one. Therefore, instead of starting with Hilbert spaces and the Schrödinger equation, one uses C^*-dynamical systems, that is, a pair constituted of a C^*-algebra and a group of $*$-automorphisms. The first generalizes the Banach space $\mathcal{B}(H)$ of all bounded (linear) operators acting on some Hilbert space H and, the second, the mapping of Definition 3.26. States of the systems are naturally defined as positive and normalized functionals on the C^*-algebra, generalizing in this way Definition 2.25.

Recall that $\mathcal{A}$ is a "$*$-algebra" whenever $\mathcal{A}$ is a *complex* vector space endowed with a product $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$, $(A, A') \mapsto AA'$ and a complex conjugation[1] $(\cdot)^* : \mathcal{A} \to \mathcal{A}$ in such a way that $\mathcal{A}$ is an algebra and $(AA')^* = A'^* A^*$. A "unital algebra" $\mathcal{A}$ is an algebra with a unit, which is always denoted here by $\mathfrak{1}$. See Definitions 7.15 and 7.60 for more details. Note also that a Banach algebra is not only a complete[2] normed algebra,[3] but it is additionally *associative*,[4] by definition. Then, as given in Definition 7.85, C^*-algebras are defined as follows.

---

[1] In this case, for any $A, B \in \mathcal{A}$ and $\alpha, \beta \in \mathbb{C}$, $(A^*)^* = A$, $(AB)^* = B^* A^*$ and $(\alpha A + \beta B)^* = \overline{\alpha} A^* + \overline{\beta} B^*$.

[2] Any Cauchy sequence in the normed space converges.

[3] $\mathcal{A}$ becomes a normed algebra whenever, for any $A, B \in \mathcal{A}$, $\|AB\| \leq \|A\| \, \|B\|$.

[4] That is, $A_1 A_2 A_3 = A_1(A_2 A_3) \doteq A_1 A_2 A_3$ for any $A_1, A_2, A_3 \in \mathcal{A}$.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_4

Definition 4.1 (C^*-Algebra) Let $\mathcal{A}$ be any $*$-algebra and $\|\cdot\|$ a norm in $\mathcal{A}$:

(i) We say that $(\mathcal{A}, \|\cdot\|)$ is a "normed $*$-algebra" if it is a normed algebra and $\|A\| = \|A^*\|$ for all $A \in \mathcal{A}$.
(ii) The normed $*$-algebra $\mathcal{A}$ is a "Banach $*$-algebra" if it is a Banach algebra (i.e., if it is associative and complete).
(iii) The Banach $*$-algebra $(\mathcal{A}, \|\cdot\|)$ is called a "C^*-algebra" if $\|A^*A\| = \|A\|^2$ for all $A \in \mathcal{A}$.
(iv) A $*$-subalgebra of the C^*-subalgebra $(\mathcal{A}, \|\cdot\|)$ is called "C^*-subalgebra" of $\mathcal{A}$ if it is norm-closed. In particular, it is itself a C^*-algebra.
(v) A C^*-subalgebra $\mathcal{B} \subseteq \mathcal{A}$ of a unital C^*-algebra $\mathcal{A}$ is a "unital C^*-subalgebra" if it contains the unit of $\mathcal{A}$. In particular, $\mathcal{B}$ is a unital C^*-algebra. (Remark that a C^*-subalgebra may be unital as a C^*-algebra without being necessarily a unital C^*-subalgebra.)

Exercise 4.2 Let $\mathcal{A}$ be any normed $*$-algebra. Show that:

(i) The involution $A \mapsto A^*$ is a continuous mapping.
(ii) If $\mathcal{A}$ is a unital C^*-algebra (i.e., $\mathcal{A}$ has a unit $\mathbb{1} \in \mathcal{A}$) then $\|\mathbb{1}\| = 1$.

In the following we give three important examples of C^*-algebra:

Example 4.3 (Algebra of Bounded Operators on a Hilbert Space) Let H be any complex Hilbert space. Then, the normed algebra $(\mathcal{B}(H), \|\cdot\|_{\text{op}})$, endowed with the complex conjugation $A \mapsto A^*$, where $A^* \in \mathcal{B}(H)$ is the adjoint operator of $A^* \in \mathcal{B}(H)$, is a unital C^*-algebra, the algebra unit being the identity operator.

Example 4.4 (Subalgebras of $\mathcal{B}(H)$) Let H be any complex Hilbert space and $\mathcal{A}$ a norm-closed subalgebra of $\mathcal{B}(H)$. $\mathcal{A}$ is a C^*-algebra, whenever it is self-conjugate (i.e., for any $A \in \mathcal{A}$, $A^* \in \mathcal{A}$).

As already mentioned above, the last example refers to the so-called concrete C^*-algebras. We will see later on that, by the Gelfand-Naimark theorem (Theorem 4.89), any C^*-algebra is equivalent to a C^*-algebra of this kind, as a $*$-algebra.

Example 4.5 (Continuous Functions on a Compact) Let K be a compact metric space (or, more generally, an arbitrary compact Hausdorff space). The normed algebra $(C(K;\mathbb{C}), \|\cdot\|_\infty)$ endowed with the usual complex conjugation for complex-valued functions $f \mapsto f^*$ (i.e., $f^*(p) \doteq \overline{f(p)}$, $p \in K$) is a unital commutative C^*-algebra. The self-conjugate elements of this $*$-algebra are precisely the real-valued continuous functions on K.

It turns out that any *unital* commutative C^*-algebra is equivalent, as a $*$-algebra, to the algebra of continuous functions on a compact Hausdorff[5] space, by the Gelfand theorem (see, e.g., [51, Theorem 2.1.11A] or Proposition 4.124 for the separable case). If the algebra is separable, this compact space can be chosen as

[5] Distinct points have disjoint neighborhoods.

being a metric space. We will prove this important result later on, in the separable case. Conversely, for any compact metric space K, the commutative unital C^*-algebra $C(K;\mathbb{C})$ is separable.

Lemma 4.6 *Let K be any compact metric space. Then, $(C(K;\mathbb{C}), \|\cdot\|_\infty)$ is a separable Banach space.*

Proof Note that, as K is compact, for any $n \in \mathbb{N}$, there is a finite subset $\Omega_n \subseteq K$ such that

$$K \subseteq \cup\{B_{1/n}(p) \ : \ p \in \Omega_n\} ,$$

where $B_{1/n}(p)$ denotes the open ball of radius $1/n$, centered at $p \in \Omega_n$. In fact, this is a consequence of the trivial inclusion

$$K \subseteq \cup\{B_{1/n}(p) \ : \ p \in K\}$$

together with Proposition 7.181. For all $n \in \mathbb{N}$, let $\mathcal{F}_n \subseteq C(K;\mathbb{C})$ be a collection of continuous functions $\mathcal{F}_n = \{h_p^{(n)} : p \in \Omega_n\}$ satisfying the following properties: (i) $h_p^{(n)}(K) \subseteq [0,1]$ for every $n \in \mathbb{N}$ and $p \in \Omega_n$; (ii) $h_p^{(n)}(K\backslash B_{1/n}(p)) = \{0\}$ for every $n \in \mathbb{N}$ and $p \in \Omega_n$, and

$$\sum_{q\in\Omega_n} h_q^{(n)}(p) = 1$$

for every $n \in \mathbb{N}$ and $p \in \Omega_n$. In other words, the family $\mathcal{F}_n$ is a so-called partition of unity subordinate to the family of open sets $\{B_{1/n}(p) : p \in \Omega_n\}$. Such families are known to exist in the present case. See, for instance, [19, 2.13 Theorem]. Now, by Proposition 7.176, it directly follows that the linear combination of functions of $\cup_{n\in\mathbb{N}}\mathcal{F}_n$ with rational coefficients is a dense countable subset of $(C(K;\mathbb{C}), \|\cdot\|_\infty)$. □

More generally, we have the following example of commutative C^*-algebras.

Example 4.7 (Continuous Functions Decaying at Infinity) Let M be an arbitrary metric space (or, more generally, an arbitrary topological space). The normed algebra $(C_0(M;\mathbb{C}), \|\cdot\|_\infty)$ of continuous functions decaying at infinity (see Definition 7.166 and Exercise 7.189) endowed with the usual complex conjugation for complex-valued functions $f \mapsto f^*$ is a (not necessarily unital) commutative C^*-algebra. Again, the self-conjugate elements of this $*$-algebra are precisely the real-valued continuous functions on M.

We will show later on that, up to a $*$-isomorphism, any separable commutative C^*-algebra has this form, up to a $*$-isomorphism, with M being a locally compact metric space. See Corollary 4.126. In fact, the same is true even if the commutative C^*-algebra is not separable, but M is then generally only a locally compact Hausdorff space, not necessarily a metric one.

The following example of commutative C^*-algebras is fundamental in the theory of Lebesgue integrals with respect to Borel measures.

Definition 4.8 (C^*-Algebras of Bounded Borel-Measurable Functions) Let M be any metric space. Let $\mathfrak{M}_b(M;\mathbb{R})$ be the smallest[6] Riesz subspace (Definition 7.267 (ii)) of $\mathcal{F}_b(M;\mathbb{R})$, which is σ-order-closed (Definition 1.20), an algebra (i.e., a subalgebra of $\mathcal{F}_b(M;\mathbb{R})$) and contains the algebra of bounded continuous functions $M \to \mathbb{R}$, i.e., $\mathfrak{M}_b(M;\mathbb{R}) \supseteq C_b(M;\mathbb{R})$. Then, we define the complex vector space:

$$\mathfrak{M}_b(M;\mathbb{C}) \doteq \{f_{\mathrm{Re}} + i f_{\mathrm{Im}} \ : \ f_{\mathrm{Re}}, f_{\mathrm{Im}} \in \mathfrak{M}_b(M;\mathbb{R})\} \subseteq \mathcal{F}_b(M;\mathbb{C}) \ .$$

The elements of $\mathfrak{M}_b(M;\mathbb{C})$ are called bounded "measurable functions"[7] on M.

Notice that $(\mathfrak{M}_b(M;\mathbb{C}), \|\cdot\|_\infty)$ is a unital commutative C^*-algebra, with $C_b(M;\mathbb{C}) \subseteq \mathfrak{M}_b(M;\mathbb{C})$ being a unital C^*-subalgebra. Observe additionally that, being a σ-order-closed subspace of the σ-order-complete space $\mathcal{F}_b(M;\mathbb{R})$, $\mathfrak{M}_b(M;\mathbb{R}) = \mathrm{Re}\{\mathfrak{M}_b(M;\mathbb{C})\}$ is σ-order-complete. See the second part of Exercise 1.21. In particular, by Corollary 7.289, $\mathfrak{M}_b(M;\mathbb{C})$ is closed with respect to the pointwise convergence of sequences.

Lemma 4.9 *Let M be an arbitrary metric space and $(f_n)_{n\in\mathbb{N}}$ a bounded sequence in $\mathfrak{M}_b(M;\mathbb{C})$ that pointwise converges to a function $f \in \mathcal{F}_b(M;\mathbb{C})$. Then, $f \in \mathfrak{M}_b(M;\mathbb{C})$.*

It turns out that the positive linear functionals on $\mathfrak{M}_b(M;\mathbb{C})$ that are σ-order-continuous, i.e., the integrals on $\mathfrak{M}_b(M;\mathbb{C})$ in the sense of Definition 7.309, are exactly the finite positive Borel measures on M:

Definition 4.10 (Borel Measures) Let M be any metric space. The *σ-order-continuous* elements of $\mathfrak{M}_b(M;\mathbb{C})'^{+}$, that is, the integrals on the Riesz-space $\mathfrak{M}_b(M;\mathbb{C})$, are called "Borel measures on M." If a Borel measure $\mu \in \mathfrak{M}_b(M;\mathbb{C})'^{+}$ is normalized, i.e., $\mu(1) = 1$, then we say that μ is a "probability measure on M."

In fact, given a Borel measure $\mu \in \mathfrak{M}_b(M;\mathbb{C})'^{+}$ in the above sense, the mapping $B \mapsto \mu(\chi_B)$ from the Borel sets of M to $\mathbb{R}_0^+$, χ_B being the characteristic function of B, is a finite positive Borel measure in the usual sense of measure theory. Recall that the characteristic functions of Borel sets are Borel-measurable functions, i.e., elements of $\mathfrak{M}_b(M;\mathbb{C})$. Conversely, if μ is a finite positive Borel measure (in the

[6] Note that such a minimum Riesz subspace under all those that are σ-order-closed algebras exists, for the intersection of σ-order-closed Riesz subspaces is again a σ-order-closed Riesz subspace and $\mathcal{F}_b(M;\mathbb{R}) \supseteq C_b(M;\mathbb{R})$ is (trivially) σ-order-closed.

[7] In fact, it is a standard exercise of measure theory to prove that $\mathfrak{M}_b(M;\mathbb{C})$ is nothing else than the space of all usual Borel-measurable bounded functions $M \to \mathbb{C}$. Thus, the present definition is equivalent to the usual one, which refers to preimages of Borel sets.

usual sense), then the linear functional $\mu \in \mathfrak{M}_b(M;\mathbb{C})'$, defined by the Lebesgue integral with respect to μ, i.e.,

$$\mu(f) \doteq \int f(p)\mu(\mathrm{d}p) \in \mathbb{C}\,, \qquad f \in \mathfrak{M}_b(M;\mathbb{C})\,,$$

is positive (by simple properties of the Lebesgue integrals) and σ-order-continuous (as a consequence of the Beppo Levi monotone convergence theorem).

From the point of view of positive linear functionals on spaces of Borel-measurable functions, the notion of support of Borel measures takes the following form.

Definition 4.11 (Supports of Borel Measures) Let M be any metric space and $\mu \in \mathfrak{M}_b(M;\mathbb{C})'^{+}$ a Borel measure. For a given subset $\Omega \subseteq M$, we say that μ "is supported in Ω" if, for every Borel-measurable function $f \in \mathfrak{M}_b(M;\mathbb{C})$ with $f|_\Omega = 0$, one has that $\mu(f) = 0$.

By Exercise 1.21, observe that in the Riesz space $\mathfrak{M}_b(M;\mathbb{R})$ of real-valued Borel-measurable functions on a metric space M, the order convergence of a sequence is equivalent to its pointwise convergence. Recall that $\mathfrak{M}_b(M;\mathbb{R})$ is, by definition, σ-order-closed in $\mathcal{F}_b(M;\mathbb{R})$. Thus, Proposition 7.313, which is an abstract version Lebesgue's dominated convergence for integrals on a general Riesz space, yields the usual form of Lebesgue's dominated convergence for Borel measures.

Proposition 4.12 (Lebesgue's Dominated Convergence for Borel Measures) *Let M be any metric space and $\mu \in \mathfrak{M}_b(M;\mathbb{C})'^{+}$ a Borel measure. Take any* bounded *sequence $(f_n)_{n\in\mathbb{N}}$ in $(\mathfrak{M}_b(M;\mathbb{C}), \|\cdot\|_\infty)$. If this sequence converges pointwise to a (bounded) function $f \in \mathcal{F}_b(M;\mathbb{R})$, then $f \in \mathfrak{M}_b(M;\mathbb{C})$ and*

$$\mu(f) = \lim_{n\to\infty} \mu(f_n)\,.$$

Proof The proposition is a direct consequence of the definition of the space of Borel-measurable functions $\mathfrak{M}_b(M;\mathbb{C})$ (Definition 4.8) and Borel measures (Definition 4.10), as well as Exercise 1.21 and Proposition 7.313. □

Here, C^*-algebras are by default unital. In all applications to quantum statistical mechanics discussed below, this is the relevant case. However, as in some situations (beyond the scope of this book) it is important to consider non-unital C^*-algebras, we make in the following some general remarks on this case. We start by showing that any non-unital C^*-algebra can be canonically extended to a unital one.

Exercise 4.13 Let $\mathcal{A}$ be any C^*-algebra (which may be possibly already unital) and consider its unitization $\tilde{\mathcal{A}} = \mathbb{C} \times \mathcal{A}$ as a $*$-algebra, as defined in Definition 7.63. For any $(\alpha, A) \in \tilde{\mathcal{A}}$, define

$$\begin{aligned}\|(\alpha, A)\|_{\tilde{\mathcal{A}}} &\doteq \sup\{\max\{|\alpha|, \|\alpha B + AB\|, \|\alpha B + BA\|\} : B \in \mathcal{A}, \ \|B\| = 1\} \\ &\leq |\alpha| + \|A\| .\end{aligned}$$

Show that:

(i) $(\tilde{\mathcal{A}}, \|\cdot\|_{\tilde{\mathcal{A}}})$ is a Banach $*$-algebra.
(ii) The new norm $\|\cdot\|_{\tilde{\mathcal{A}}}$ extends the norms of the C^*-algebras $\mathbb{C}$ and $\mathcal{A}$, that is, for all $\alpha \in \mathbb{C}$ and $A \in \mathcal{A}$, $\|(\alpha, 0)\|_{\tilde{\mathcal{A}}} = |\alpha|$ and $\|(0, A)\|_{\tilde{\mathcal{A}}} = \|A\|$.
(iii) $\mathcal{A} \subseteq \tilde{\mathcal{A}}$ is a closed $*$-ideal of $\tilde{\mathcal{A}}$.

Exercise 4.14 Prove that, for all $(\alpha, A) \in \tilde{\mathcal{A}}$,

$$\begin{aligned}\|(\alpha, A)\|_{\tilde{\mathcal{A}}} &= \sup\{\max\{|\alpha|, \|\alpha B + AB\|\} : B \in \mathcal{A}, \ \|B\| = 1\} \\ &= \sup\{\max\{|\alpha|, \|\alpha B + BA\|\} : B \in \mathcal{A}, \ \|B\| = 1\} .\end{aligned}$$

From Exercise 4.13, in order to prove that $\tilde{\mathcal{A}}$ is a C^*-algebra, it suffices to show that, for all $A \in \tilde{\mathcal{A}}$,

$$\left\|A^* A\right\|_{\tilde{\mathcal{A}}} \geq \|A\|_{\tilde{\mathcal{A}}}^2 .$$

This estimate is given in the following lemma.

Lemma 4.15 (C^*-Property of $\|\cdot\|_{\tilde{\mathcal{A}}}$) *For all $\alpha \in \mathbb{C}$ and $A \in \mathcal{A}$,*

$$\left\|(\alpha, A)^*(\alpha, A)\right\|_{\tilde{\mathcal{A}}} \geq \|(\alpha, A)\|_{\tilde{\mathcal{A}}}^2 .$$

Proof By Exercise 4.14 and the fact that $\|\cdot\|_{\tilde{\mathcal{A}}}$ extends the original norm of $\mathcal{A}$, for all $\alpha \in \mathbb{C}$ and $A \equiv (0, A) \in \mathcal{A} \subseteq \tilde{\mathcal{A}}$,

$$\|(\alpha, A)\|_{\tilde{\mathcal{A}}}^2 = \sup\left\{\max\{|\alpha|^2, \|(\alpha\mathbb{1} + A)B\|_{\tilde{\mathcal{A}}}^2\} : B \in \mathcal{A}, \ \|B\| = 1\right\} ,$$

while

$$\begin{aligned}\left\|(\alpha, A)^*(\alpha, A)\right\|_{\tilde{\mathcal{A}}} &= \sup_{B \in \mathcal{A}, \ \|B\|=1} \max\{|\alpha|^2, \left\|(\alpha\mathbb{1} + A)^*(\alpha\mathbb{1} + A)B\right\|_{\tilde{\mathcal{A}}}, \\ &\qquad \left\|B(\alpha\mathbb{1} + A)^*(\alpha\mathbb{1} + A)\right\|_{\tilde{\mathcal{A}}}\} \\ &= \sup\{\max\{|\alpha|^2, \left\|(\alpha\mathbb{1} + A)^*(\alpha\mathbb{1} + A)B\right\|_{\tilde{\mathcal{A}}}\} : B \in \mathcal{A}, \\ &\qquad \|B\| = 1\} .\end{aligned}$$

The last equality follows from

$$\begin{aligned}
&\sup\left\{\left\|B(\alpha\mathbb{1}+A)^*(\alpha\mathbb{1}+A)\right\| \,:\, B\in\mathcal{A},\ \|B\|=1\right\}\\
&\quad=\sup\left\{\left\|(\alpha\mathbb{1}+A)^*(\alpha\mathbb{1}+A)B^*\right\| \,:\, B\in\mathcal{A},\ \|B\|=1\right\}\\
&\quad=\sup\left\{\left\|(\alpha\mathbb{1}+A)^*(\alpha\mathbb{1}+A)B\right\| \,:\, B\in\mathcal{A},\ \|B\|=1\right\}\,.
\end{aligned}$$

Observing that, for $B \in \mathcal{A}$, $\|B\| = 1$,

$$\|(\alpha\mathbb{1}+A)B\|^2_{\tilde{\mathcal{A}}} = \left\|B^*(\alpha\mathbb{1}+A)^*(\alpha\mathbb{1}+A)B\right\|_{\tilde{\mathcal{A}}} \leq \left\|(\alpha\mathbb{1}+A)^*(\alpha\mathbb{1}+A)B\right\|$$

the assertion then follows. □

From the last results, we arrive at the following statement.

Proposition 4.16 (Unitization of C^*-Algebras) *For any C^*-algebra $\mathcal{A}$ (which may be possibly already unital), $(\tilde{\mathcal{A}}, \|\cdot\|_{\tilde{\mathcal{A}}})$ is a unital C^*-algebra with $\mathcal{A} \subseteq \tilde{\mathcal{A}}$ being a closed $*$-ideal. In particular, $\mathcal{A}$ is a C^*-subalgebra of its unitization.*

From Corollary 4.32, note moreover that the norm $\|\cdot\|_{\tilde{\mathcal{A}}}$ defined above is the unique norm of the unitization $\tilde{\mathcal{A}}$ of any C^*-algebra $\mathcal{A}$, for which the normed $*$-algebra $(\tilde{\mathcal{A}}, \|\cdot\|_{\tilde{\mathcal{A}}})$ is a (unital) C^*-algebra.

4.2 The Spectrum of an Algebra Element

Let $\mathcal{A}$ be any unital algebra, i.e., an algebra with a unit, always denoted by $\mathbb{1}$. We say that the element $A \in \mathcal{A}$ is "invertible" in $\mathcal{A}$ if there are $A_L^{-1}, A_R^{-1} \in \mathcal{A}$ such that

$$AA_R^{-1} = A_L^{-1}A = \mathbb{1}\,.$$

In this case note that $A_R^{-1} = A_L^{-1} \doteq A^{-1}$. Such an element $A^{-1} \in \mathcal{A}$ is called the "inverse element," or the "inverse," of $A \in \mathcal{A}$. It has the following properties.

Lemma 4.17 *Let $\mathcal{A}$ be a unital associative algebra (Definition 7.15):*

(i) *For all $A \in \mathcal{A}$, the inverse $A^{-1} \in \mathcal{A}$ is unique when it exists.*
(ii) *For any finite collection of invertible elements $A_1, \ldots, A_n \in \mathcal{A}$, the product $B \doteq A_1 \cdots A_n \in \mathcal{A}$ is invertible and*

$$B^{-1} = A_n^{-1} \cdots A_1^{-1}.$$

If $A_iA_j = A_jA_i$ for all $i, j \in \{1, \ldots, n\}$, then B is invertible only if $A_1, \ldots, A_n$ are all invertible.
(iii) *If $\mathcal{A}$ is a $*$-algebra and $A \in \mathcal{A}$ is invertible, then also A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$.*

Proof Exercise. *Hint*: To prove the second part of (ii), use the Ansätze

$$(A_i)_L^{-1} = B^{-1} A_1 \cdots A_{i-1} A_{i+1} \cdots A_n \ ,$$
$$(A_i)_R^{-1} = A_1 \cdots A_{i-1} A_{i+1} \cdots A_n B^{-1} \ .$$

□

We are now in a position to define the spectrum of an algebra element, and to give some of its main properties. To the end, recall that any non-unital algebra $\mathcal{A}$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ can be canonically extended to a unital one, denoted here by $\tilde{\mathcal{A}}$. This is the so-called unitization of an algebra; see Definition 7.16.

Definition 4.18 (Resolvent Set and Spectrum) Let $\mathcal{A}$ be a unital algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For any $A \in \mathcal{A}$, the "resolvent set" $R_{\mathcal{A}}(A) \subseteq \mathbb{K}$ of A is the set of all $\lambda \in \mathbb{K}$ such that $(\lambda \mathfrak{1} - A) \in \mathcal{A}$ is invertible. $\sigma_{\mathcal{A}}(A) \doteq \mathbb{K} \backslash R_{\mathcal{A}}(A)$ is called the "spectrum of A." For all $\lambda \in R_{\mathcal{A}}(A)$, we define

$$\mathcal{R}(A, \lambda) \doteq (\lambda \mathfrak{1} - A)^{-1} \in \mathcal{A} \ .$$

This element of the algebra $\mathcal{A}$ is called " the resolvent of A at λ." If $\mathcal{A}$ is a non-unital algebra over $\mathbb{K}$, then we define the spectrum of any $A \in \mathcal{A}$ as being $\sigma_{\mathcal{A}}(A) \doteq \sigma_{\tilde{\mathcal{A}}}(A)$, where $\tilde{\mathcal{A}}$ is the unitization of $\mathcal{A}$.

Example 4.19 Let $\mathcal{A} = C(K; \mathbb{C})$, where K is any compact metric space. Then, for all $f \in \mathcal{A}$, $\sigma(f) = f(K)$.

Exercise 4.20 Let $\mathcal{A}$ be any *non-unital associative* algebra. Show that, for all $A \in \mathcal{A}$, one has that $0 \in \sigma_{\mathcal{A}}(A)$.

Hint: Show that if $0 \notin \sigma_{\mathcal{A}}(A) \doteq \sigma_{\tilde{\mathcal{A}}}(A)$ for some $A \in \mathcal{A} \subseteq \tilde{\mathcal{A}}$, then the unit of $\tilde{\mathcal{A}}$ must be an element of $\mathcal{A}$.

The following lemma gathers important properties of the spectrum with respect to algebra and $*$-algebra operations.

Lemma 4.21 *Let $\mathcal{A}$ be any associative algebra:*

(i) *If $\mathcal{A}$ is a $*$-algebra, then, for all $A \in \mathcal{A}$*

$$\sigma_{\mathcal{A}}(A^*) = \overline{\sigma_{\mathcal{A}}(A)} \doteq \{\bar{z} \ : \ z \in \sigma_{\mathcal{A}}(A)\} \ .$$

(ii) *If $\mathcal{A}$ is unital and $A \in \mathcal{A}$ is invertible, then $0 \notin \sigma_{\mathcal{A}}(A) \cup \sigma_{\mathcal{A}}(A^{-1})$ and*

$$\sigma_{\mathcal{A}}(A^{-1}) = \sigma_{\mathcal{A}}(A)^{-1} \doteq \{z^{-1} \ : \ z \in \sigma_{\mathcal{A}}(A)\} \ .$$

(iii) *For all $A, B \in \mathcal{A}$*

$$\sigma_{\mathcal{A}}(AB) \cup \{0\} = \sigma_{\mathcal{A}}(BA) \cup \{0\} \ .$$

Proof Exercise. *Hint:* We may assume that $\mathcal{A}$ is unital. In order to show (iii), prove first the identity

$$(\lambda 1 - BA)\left(B(\lambda 1 - AB)^{-1}A + 1\right) = \lambda 1$$

for all $A, B \in \mathcal{A}$ and $\lambda \in R_{\mathcal{A}}(AB)$. □

Let $\mathcal{A}$ be an arbitrary unital algebra. For all $A \in \mathcal{A}$ and $k \in \mathbb{N}_0$, let $A^k \in \mathcal{A}$ be recursively defined by $A^0 \doteq 1$ and

$$A^{k+1} \doteq AA^k \ , \qquad k \in \mathbb{N}_0 \ .$$

Using the above definition of monomials of algebra elements, we define polynomials of such elements.

Definition 4.22 (Polynomials of Algebra Elements) Let $\mathcal{A}$ be an arbitrary unital associative algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For any polynomial $\mathcal{P}(s) = \sum_{k=0}^{n} c_k s^k$, $c_1, \ldots, c_n \in \mathbb{K}$ ($n \in \mathbb{N}_0$) and all $A \in \mathcal{A}$, define

$$\mathcal{P}(A) \doteq \sum_{k=0}^{n} c_k A^k \in \mathcal{A} \ .$$

If $\mathcal{A}$ is non-unital, we restrict the above definition of $\mathcal{P}(A)$ to the special case of polynomials $\mathcal{P}$ for which the constant term c_0 is zero.

Observe that the following identities hold true for any pair of polynomials $\mathcal{P}_1, \mathcal{P}_2$ and every algebra element $A \in \mathcal{A}$:

$$(\mathcal{P}_1 + \mathcal{P}_2)(A) = \mathcal{P}_1(A) + \mathcal{P}_2(A) \ ,$$
$$(\mathcal{P}_1\mathcal{P}_2)(A) = \mathcal{P}_1(A)\mathcal{P}_2(A) \ ,$$
$$\mathcal{P}_1 \circ \mathcal{P}_2(A) = \mathcal{P}_1(\mathcal{P}_2(A)) \ .$$

For a complex algebra $\mathcal{A}$, note also that if the polynomial $\mathcal{P}$ has real coefficients (i.e., $c_1, \ldots, c_n \in \mathbb{R}$), then $\mathcal{P}(A) \in \mathrm{Re}\{\mathcal{A}\}$ whenever $A \in \mathrm{Re}\{\mathcal{A}\}$.

In the following proposition, we show that, in any *complex* unital algebra, taking the spectrum and building a polynomial of an element are interchangeable operations.

Proposition 4.23 (Spectral Mapping Property) *Let $\mathcal{A}$ be any (not necessarily unital) associative algebra over $\mathbb{C}$. For every $A \in \mathcal{A}$ such that $\sigma_{\mathcal{A}}(A) \neq \emptyset$, and any polynomial $\mathcal{P}$,*

$$\sigma_{\mathcal{A}}(\mathcal{P}(A)) = \mathcal{P}(\sigma_{\mathcal{A}}(A)) \doteq \{\mathcal{P}(z) \ : \ z \in \sigma_{\mathcal{A}}(A)\} \ .$$

Proof Assume for simplicity that $\mathcal{A}$ is unital (otherwise one has to take its unitization $\tilde{\mathcal{A}}$). If $\mathcal{P} = c \in \mathbb{C}$, the assertion is trivial. Thus, fix any polynomial $\mathcal{P} \notin \mathbb{C}$ and $A \in \mathcal{A}$ with nonempty spectrum $\sigma_{\mathcal{A}}(A) \neq \emptyset$.

1. For any $\lambda \in \mathbb{C}$, by the "fundamental theorem of algebra," there are constants

$$a(\lambda), b_1(\lambda), \ldots, b_n(\lambda) \in \mathbb{C}\ ,\ a(\lambda) \neq 0\ ,$$

such that, for any $s \in \mathbb{C}$,

$$\mathcal{P}(s) - \lambda = a(\lambda)(b_1(\lambda) - s) \cdots (b_n(\lambda) - s)\ .$$

In particular

$$\mathcal{P}(A) - \lambda\mathfrak{1} = a(\lambda)(b_1(\lambda)\mathfrak{1} - A) \cdots (b_n(\lambda)\mathfrak{1} - A)\ .$$

2. As $a(\lambda) \neq 0$ and

$$(b_i(\lambda)\mathfrak{1} - A)(b_j(\lambda)\mathfrak{1} - A) = (b_j(\lambda)\mathfrak{1} - A)(b_i(\lambda)\mathfrak{1} - A)$$

for all $i, j \in \{1, \ldots, n\}$, $(\mathcal{P}(A) - \lambda\mathfrak{1})$ is non-invertible iff $(b_i(\lambda)\mathfrak{1} - A)$ is non-invertible for some $i \in \{1, \ldots, n\}$, thanks to Lemma 4.17 (ii). Thus, $\lambda \in \sigma(\mathcal{P}(A))$ iff, for some $i \in \{1, \ldots, n\}$, $b_i(\lambda) \in \sigma(A)$. But, by construction, $\mathcal{P}(s) = \lambda$ iff $s \in \{b_1(\lambda), \ldots, b_n(\lambda)\}$. Hence, $\lambda \in \sigma(\mathcal{P}(A))$ iff $\lambda \in \mathcal{P}(\sigma(A))$.

□

In the following we study the spectrum of elements of Banach algebras. Recall that a Banach algebra is not only a complete normed algebra, but it is also associative, by definition.

Lemma 4.24 (Neumann Series) *Let $\mathcal{A}$ be a unital Banach algebra. For all $A, B \in \mathcal{A}$ such that A is invertible with $\|BA^{-1}\| < 1$, $(A - B) \in \mathcal{A}$ is also invertible and*

$$(A - B)^{-1} = A^{-1} + A^{-1} \sum_{n=1}^{\infty} (BA^{-1})^n\ .$$

This last series is known as the "Neumann series" for $(A - B)^{-1}$.

Proof Exercise. □

The above lemma has the following important consequence for the spectrum of elements of Banach algebras.

Corollary 4.25 (Spectrum in Banach Algebras) *Let $\mathcal{A}$ be a unital Banach algebra. For all $A \in \mathcal{A}$, the resolvent set $R_{\mathcal{A}}(A)$ is an* open *subset of $\mathbb{C}$ and the spectrum $\sigma_{\mathcal{A}}(A) \subseteq \mathbb{C}$ is* compact.

Proof Exercise. □

The above lemma on Neumann series also yields that the spectrum of an element of a complex Banach algebra is never empty.

Lemma 4.26 *Let $\mathcal{A}$ be a unital Banach algebra over $\mathbb{C}$. For all $A \in \mathcal{A}$, $\sigma_{\mathcal{A}}(A) \neq \emptyset$.*

Idea of Proof

1. Fix $A \in \mathcal{A}$ and assume that $\sigma_{\mathcal{A}}(A) = \emptyset$, that is, $R_{\mathcal{A}}(A) = \mathbb{C}$. For any continuous linear function $\varphi \in \mathcal{A}^{\text{td}}$, define the function $f_\varphi : \mathbb{C} \to \mathbb{C}$ by

$$f_\varphi(z) \doteq \varphi\left(\mathcal{R}(A, z)\right) .$$

By using the Neumann series for

$$\mathcal{R}(A, z) = (z\mathbb{1} - A)^{-1} = (z_0\mathbb{1} - A + (z - z_0)\mathbb{1}))^{-1} ,$$

one concludes that f_φ is an entire function, i.e., it is analytic over the whole complex plane $\mathbb{C}$.

2. Again by the Neumann series, one additionally has

$$\lim_{|z|\to\infty} f_\varphi(z) = 0 .$$

As f_φ is continuous (because it is analytic), it follows that $f_\varphi \in C_0(\mathbb{C}; \mathbb{C})$. In particular it is uniformly bounded in $\mathbb{C}$.

3. Hence, from the Liouville theorem, it has to be constant. For it tends to zero, as $|z| \to \infty$, $f_\varphi = 0$. Thus, one has

$$\varphi\left((z\mathbb{1} - A)^{-1}\right) = 0$$

for all $z \in \mathbb{C}$ and $\varphi \in \mathcal{A}^{\text{td}}$.

4. Continuous linear functionals separate elements of any normed space, by the Hahn-Banach theorem. See Theorem 7.40 and Corollary 7.41. It follows that $(z\mathbb{1} - A)^{-1} = 0$ for all $z \in \mathbb{C} = R_{\mathcal{A}}(A)$. But this would imply that, for all $z \in \mathbb{C}$,

$$\mathbb{1} = (z\mathbb{1} - A)^{-1}(z\mathbb{1} - A) = 0 ,$$

which is a contradiction and one thus has $\sigma_{\mathcal{A}}(A) \neq \emptyset$.

□

As the spectrum $\sigma_{\mathcal{A}}(A) \subseteq \mathbb{C}$ of any element A of a complex Banach algebra $\mathcal{A}$ is always nonempty and compact (in particular, it is bounded), we can define its "spectral radius":

$$r_{\mathcal{A}}(A) \doteq \max\{|\lambda| \ : \ \lambda \in \sigma_{\mathcal{A}}(A)\} \in \mathbb{R}_0^+ .$$

This positive number can be characterized by the so-called Gelfand radius, which will be defined from the following lemma.

Lemma 4.27 *Let $A \in \mathcal{A}$ be an arbitrary element of a normed associative algebra $\mathcal{A}$. Then,*

$$\lim_{n\to\infty} \|A^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|A^n\|^{1/n} \le \|A\| \ .$$

Proof

1. Fix $A \in \mathcal{A}$. Note that $\|A^n\|^{1/n} \le \|A\|$ and if $\|A^n\| = 0$ for some $n \in \mathbb{N}$ then the assertion trivially follows, from the submultiplicativity[8] of the norm of a normed algebra. Thus, we can assume without loss of generality that $\|A^n\| > 0$ for all $n \in \mathbb{N}$. In this case, define the sequence of real numbers:
$$c_n \doteq \ln(\|A^n\|^{\frac{1}{n}}) \le \ln(\|A\|) \ , \qquad n \in \mathbb{N} \ .$$
2. If $\inf_{n\in\mathbb{N}} c_n = -\infty$, then $\lim_{n\to\infty} \|A^n\|^{1/n} = 0$ and again the assertion trivially follows. So, we assume that $\inf_{n\in\mathbb{N}} c_n \in \mathbb{R}$.
3. By the submultiplicativity of the norm, one has that
$$c_{n_1+\cdots+n_m} \le \frac{n_1}{n_1+\cdots+n_m} c_{n_1} + \cdots + \frac{n_m}{n_1+\cdots+n_m} c_{n_m}$$
for all $m \in \mathbb{N}$ and $n_1, \ldots, n_m \in \mathbb{N}_0$ with $n_1 + \cdots + n_m \ge 1$, where $c_0 \doteq 1$. In particular, for all $n, N \in \mathbb{N}$ with $N \ge n$,
$$c_N \le \frac{N \bmod n}{N} \ln(\|A\|) + \frac{N - (N \bmod n)}{N} c_n \ .$$
Here, $(N \bmod n)$ denotes the rest of the division of N by n in $\mathbb{N}_0$.
4. From this estimate we conclude that, for all $\varepsilon > 0$, there is $N_\varepsilon \in \mathbb{N}$ such that, for all $N \ge N_\varepsilon$,
$$c_N \le \varepsilon + \inf_{n\in\mathbb{N}} c_n \in \mathbb{R} \ .$$
This implies that
$$\lim_{n\to\infty} \ln(\|A^n\|^{\frac{1}{n}}) = \inf_{n\in\mathbb{N}} \ln(\|A^n\|^{\frac{1}{n}}) \ .$$

[8] $\|AB\| \le \|A\|\|B\|$ for all $A, B \in \mathcal{A}$.

For $\exp(\cdot)$ is a continuous monotonically increasing function $\mathbb{R} \to \mathbb{R}$,

$$\lim_{n\to\infty} \|A^n\|^{1/n} = \exp\left(\lim_{n\to\infty} \ln(\|A^n\|^{\frac{1}{n}})\right)$$
$$= \exp\left(\inf_{n\in\mathbb{N}} \ln(\|A^n\|^{\frac{1}{n}})\right) = \inf_{n\in\mathbb{N}} \|A^n\|^{1/n} .$$

□

Because of the above lemma, one defines the so-called Gelfand radius of an element of a normed associative algebra as follows.

Definition 4.28 (Gelfand Radius) For any element A of a normed associative algebra $\mathcal{A}$, we define its "Gelfand radius" by

$$r_{\mathcal{A}}^{\mathrm{Gelf}}(A) \doteq \lim_{n\to\infty} \|A^n\|^{1/n} = \inf_{n\in\mathbb{N}} \|A^n\|^{1/n} \leq \|A\| .$$

In Banach algebras, the Gelfand and spectral radii of an element coincide, as stated in the next theorem.

Theorem 4.29 (Spectral Radius of a Banach Algebra Element) *Let $\mathcal{A}$ be a unital Banach algebra. For all $A \in \mathcal{A}$,*

$$r_{\mathcal{A}}(A) = r_{\mathcal{A}}^{\mathrm{Gelf}}(A) .$$

This assertion refers for instance to [51, Proposition 2.2.2]. The lower bound $r_{\mathcal{A}}(A) \geq r_{\mathcal{A}}^{\mathrm{Gelf}}(A)$ follows from arguments using Neumann series. We omit the details of the proof. The upper bound $r_{\mathcal{A}}(A) \leq r_{\mathcal{A}}^{\mathrm{Gelf}}(A)$ is a direct consequence of the following fact.

Exercise 4.30 Let $\mathcal{A}$ be a unital Banach algebra. Take any invertible element $A \in \mathcal{A}$ as well as $B \in \mathcal{A}$ such that

$$r_{\mathcal{A}}^{\mathrm{Gelf}}(BA^{-1}) < 1 .$$

Show that $(A - B) \in \mathcal{A}$ is invertible.

Theorem 4.29 can then be used to study the spectrum of a C^*-algebra element, recalling that a C^*-algebra is a particular case of a Banach algebra.

Proposition 4.31 (Spectrum of a C^*-Algebra Element) *Let $\mathcal{A}$ be a unital C^*-algebra:*

(i) *If $A \in \mathcal{A}$ is normal, i.e., if $AA^* = A^*A$, then $r_{\mathcal{A}}(A) = \|A\|$.*
(ii) *If $A \in \mathcal{A}$ is unitary, i.e., if $AA^* = A^*A = \mathfrak{1}$, then*

$$\sigma_{\mathcal{A}}(A) \subseteq \{z \in \mathbb{C} \ : \ |z| = 1\} .$$

In particular, from (i) *and* $\|\mathbb{1}\| = 1$ *(Exercise 4.2), for any unitary* $A \in \mathcal{A}$, $\|A\| = 1$.

(iii) *If A is self-conjugate, i.e., if $A = A^*$, then*

$$\sigma_{\mathcal{A}}(A) \subseteq [-\|A\|, \|A\|] \subseteq \mathbb{R} .$$

Proof

1. As $\mathcal{A}$ is a C^*-algebra, for all $A \in \mathcal{A}$,

$$r_{\mathcal{A}}(A) = \lim_{n\to\infty} \|A^n\|^{\frac{1}{n}} = \lim_{n\to\infty} \|A^{2^n}\|^{2^{-n}} = \lim_{n\to\infty} \|(A^{2^n})^* A^{2^n}\|^{2^{-n-1}} ,$$

thanks to Theorem 4.29. Thus, if A is normal, then

$$r_{\mathcal{A}}(A) = \lim_{n\to\infty} \|(A^*A)^{2^n}\|^{2^{-n-1}} .$$

2. Using again the fact that $\mathcal{A}$ is a C^*-algebra and that A^*A is self-conjugate, one has the identities

$$\|(A^*A)^{2^n}\| = \|(A^*A)^{2^{n-1}}\|^2 = \cdots = \|A^*A\|^{2^n}$$

and, hence, $r_{\mathcal{A}}(A) = \|A^*A\|^{2^{-1}} = \|A\|$. This proves (i).
3. To prove (ii), note that unitaries are normal elements. Hence, from the above computation, for any unitary $A \in \mathcal{A}$

$$r_{\mathcal{A}}(A) = \|A^*A\|^{2^{-1}} = \|\mathbb{1}\|^{2^{-1}} = 1 .$$

Recall that $\|\mathbb{1}\| = 1$ for the unity of any C^*-algebra, by Exercise 4.2. Thus, take $z \in \sigma(A)$, implying in particular $|z| \leq 1$. In order to prove that $|z|$ cannot be smaller than 1, we observe that, as A is unitary, A is invertible and $A^{-1} = A^*$. In particular,

$$\sigma(A^*) = \{\bar{z} \ : \ z \in \sigma(A)\} = \sigma(A^{-1}) = \{z^{-1} \ : \ z \in \sigma(A)\} .$$

From this identity and observing that $r_{\mathcal{A}}(A^*) = r_{\mathcal{A}}(A) = 1$, one has that $|z^{-1}| = |z|^{-1} \leq 1$ and, hence, $|z| \geq 1$ for all $z \in \sigma(A)$.
4. Assume now that $A = A^*$. Then, $\pm i(1 + \|A\|) \in R_{\mathcal{A}}(A)$ (because the spectral radius of A is $\|A\|$) and the element

$$\begin{aligned}&(i(\|A\| + 1)\mathbb{1} - A)(i(\|A\| + 1)\mathbb{1} + A)^{-1}\\ &\quad = (i(\|A\| + 1)\mathbb{1} + A)^{-1}(i(\|A\| + 1)\mathbb{1} - A) \in \mathcal{A}\end{aligned}$$

is unitary, by Lemma 4.17 (iii). In particular, from (ii), for any $\alpha \in \mathbb{C}$ so that $\mathrm{Im}\{\alpha\} \neq 0$ and $|\alpha| \leq \|A\|$, the element

$$\frac{(\|A\|+1)i-\alpha}{(\|A\|+1)i+\alpha}\mathbb{1}-(i(\|A\|+1)\mathbb{1}-A)(i(\|A\|+1)\mathbb{1}+A)^{-1} \in \mathcal{A}$$

is invertible, because

$$\left|\frac{(\|A\|+1)i-\alpha}{(\|A\|+1)i+\alpha}\right| \neq 1 .$$

5. But

$$\begin{aligned}\frac{(\|A\|+1)i-\alpha}{(\|A\|+1)i+\alpha}&\mathbb{1}-(i(\|A\|+1)\mathbb{1}-A)(i(\|A\|+1)\mathbb{1}+A)^{-1}\\ &= \frac{2i(\|A\|+1)}{(\|A\|+1)i+\alpha}(A-\alpha\mathbb{1})(i(\|A\|+1)\mathbb{1}+A)^{-1} .\end{aligned}$$

Noting that

$$(A-\alpha\mathbb{1})(i(1+\|A\|)\mathbb{1}+A)^{-1} = (i(1+\|A\|)\mathbb{1}+A)^{-1}(A-\alpha\mathbb{1})$$

together with Lemma 4.17 (ii), it follows that $(\alpha\mathbb{1}-A)$ is invertible for all $\alpha \in \mathbb{C}$ so that $\mathrm{Im}\{\alpha\} \neq 0$ and $|\alpha| \leq \|A\|$. Hence, as the spectral radius of A is $\|A\|$, $\sigma(A) \subseteq [-\|A\|, \|A\|]$. This proves (iii).

□

Corollary 4.32 (Uniqueness of the Norm of a C^*-Algebra) *Let $\mathcal{A}$ be an arbitrary (not necessarily unital) $*$-algebra. There is at most one norm $\|\cdot\|$ in $\mathcal{A}$ such that $(\mathcal{A}, \|\cdot\|)$ is a C^*-algebra.*

Proof Observe that, for all $A \in \mathcal{A}$, the spectrum $\sigma_{\mathcal{A}}(A)$ and, in particular, the spectral radius of this element are completely determined by the algebra structure of $\mathcal{A}$. If $(\mathcal{A}, \|\cdot\|)$ is a unital C^*-algebra, then for all $A \in \mathcal{A}$

$$\|A\| = \|A^*A\|^{1/2} = r_{\mathcal{A}}(A^*A)^{1/2} ,$$

because A^*A is a normal element. See Proposition 4.31 (i). If $\mathcal{A}$ is non-unital, then the corollary follows from the first part of the proof combined with Proposition 4.16.

□

Note that the spectrum $\sigma_{\mathcal{A}}(B)$ of an algebra element $B \in \mathcal{A}$ a priori depends on the whole algebra $\mathcal{A}$. Observe also, directly from the definition of spectra, that the inclusion $\sigma_{\mathcal{A}}(B) \subseteq \sigma_{\mathcal{B}}(B)$ holds true, for any element $B \in \mathcal{B}$ of a unital subalgebra

$\mathcal{B} \subseteq \mathcal{A}$ of an arbitrary unital algebra. If B belongs to a unital C^*-subalgebra[9] $\mathcal{B} \subseteq \mathcal{A}$ of a C^*-algebra $\mathcal{A}$, then one may wonder whether $\sigma_{\mathcal{A}}(B)$ and $\sigma_{\mathcal{B}}(B)$ can be different sets. It turns out that this is not possible, as consequence of the following proposition.

Proposition 4.33 *Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{B} \subseteq \mathcal{A}$ a unital C^*-subalgebra. If $B \in \mathcal{B}$ is invertible in $\mathcal{A}$, then $B^{-1} \in \mathcal{B}$.*

The above proposition follows from the last theorem and arguments involving the Neumann series. We omit again the proof. See for instance [51, Proof of Proposition 2.2.7].

Corollary 4.34 *Let $\mathcal{A}$ be a unital C^*-algebra and $\mathcal{B} \subseteq \mathcal{A}$ a unital C^*-subalgebra. For all $B \in \mathcal{B}$, $\sigma_{\mathcal{A}}(B) = \sigma_{\mathcal{B}}(B)$.*

By the last corollary, given a unital C^*-algebra $\mathcal{A}$, for all $A \in \mathcal{A}$, the spectrum $\sigma_{\mathcal{A}}(A)$ of A does not depend on the whole algebra $\mathcal{A}$, but only on the smallest unital C^*-subalgebra of $\mathcal{A}$ containing the element A. This fact motivates the following definition for the spectrum of elements of a C^*-algebra.

Definition 4.35 (Spectrum of Elements of General C^*-Algebras) Let $\mathcal{A}$ be any unital (non-unital) C^*-algebra. For all $A \in \mathcal{A}$, $\sigma(A)$, the "spectrum" of A, is defined to be $\sigma(A) \doteq \sigma_{\mathcal{B}}(A)$, where $\mathcal{B}$ is any unital C^*-algebra that contains $\mathcal{A}$ (the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$) as a unital C^*-subalgebra. (Recall that $\tilde{\mathcal{A}}$ is a unital C^*-algebra, thanks to Proposition 4.16.)

4.3 C^*-Algebras as $*$-Ordered Vector Spaces

For simplicity of exposition, in the first part of this section, we mainly focus on unital C^*-algebras, the non-unital case being discussed afterward. In order to see a C^*-algebra $\mathcal{A}$ as an $*$-ordered vector space, we need to define a self-conjugate convex cone in $\mathcal{A}$, as explained in Sect. 1.1. It corresponds to define the positive elements of $\mathcal{A}$. Motivated by the equivalent definition of positive operators on Hilbert spaces stated in Corollary 2.5, one defines positivity of elements of unital C^*-algebras as follows.

Definition 4.36 (Positive Elements of a C^*-Algebra) Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. We say that $A \in \mathcal{A}$ is "positive" if $A = A^*$ and $\sigma(A) \subseteq \mathbb{R}_0^+$. The set of all positive elements of $\mathcal{A}$ is denoted by $\mathcal{A}^+ \subseteq \mathrm{Re}\{\mathcal{A}\}$.

To get a preorder relation in $\mathcal{A}$, we need now to prove that $\mathcal{A}^+$ is a convex cone. This is relatively obvious in the following example.

[9] $\mathcal{B}$ is a self-conjugate subalgebra of $\mathcal{A}$ which is itself a C^*-algebra and contains the unit of $\mathcal{A}$, which is thus also the unit of $\mathcal{B}$.

Example 4.37

$$C(K;\mathbb{C})^+ = \{f \in C(K) \ : \ f(p) \geq 0 \quad \text{for all } p \in K\} ,$$

where K is any compact metric space and $C(K;\mathbb{C})$ is seen as a unital C^*-algebra.

In the above example, the positive elements are exactly those whose spectrum is positive. Note, however, that, in general, the positivity of the spectrum of a C^*-algebra element does not imply that the element is self-conjugate. In this particular example, we clearly see that $\mathcal{A}^+$ is a convex cone. In fact, it is the usual positive cone for complex-valued functions. Nevertheless, in the general case, i.e., for a general C^*-algebra $\mathcal{A}$, the fact that $\mathcal{A}^+$ is a convex cone is not completely obvious, and its proof requires some preliminary technical results.

We first observe in the following lemma that any unital C^*-algebra is linearly generated by its positive elements, like in Proposition 2.6 for the special case of the unital C^*-algebra $\mathcal{B}(H)$ of bounded operators on a complex Hilbert space H.

Lemma 4.38 *Let $\mathcal{A}$ be any unital C^*-algebra. For all $A \in \mathrm{Re}\{\mathcal{A}\}$, there are $A^+, A^- \in \mathcal{A}^+$ such that*

$$A = A^+ - A^- , \qquad \|A^+\|, \|A^-\| \leq \|A\| .$$

More generally, for all (not necessarily self-conjugate) $A \in \mathcal{A}$, there are $A^+_{\mathrm{Re}}, A^-_{\mathrm{Re}}, A^+_{\mathrm{Im}}, A^-_{\mathrm{Im}} \in \mathcal{A}^+$ such that

$$A = A^+_{\mathrm{Re}} - A^-_{\mathrm{Re}} + i(A^+_{\mathrm{Im}} - A^-_{\mathrm{Im}}) , \qquad \|A^+_{\mathrm{Re}}\|, \|A^-_{\mathrm{Re}}\|, \|A^+_{\mathrm{Im}}\|, \|A^-_{\mathrm{Im}}\| \leq \|A\| .$$

In particular, for any $\varphi, \varphi' \in \mathcal{A}'$, $\varphi = \varphi'$ if $\varphi(A) = \varphi'(A)$ for all $A \in \mathcal{A}^+$.

Proof If $A = 0$, then the assertion is trivial. For all $A \in \mathcal{A}$, $A \neq 0$,

$$\begin{aligned} A &= \frac{\|A\|}{4}\left(\|A\|^{-1} A + \mathbf{1}\right)^2 - \frac{\|A\|}{4}\left(\|A\|^{-1} A - \mathbf{1}\right)^2 \\ &= \mathcal{P}_1(A) - \mathcal{P}_2(A) , \end{aligned}$$

where the polynomials $\mathcal{P}_1, \mathcal{P}_2$ are defined by

$$\mathcal{P}_1(s) \doteq \frac{\|A\|}{4}\left(\|A\|^{-1} s + 1\right)^2 , \qquad \mathcal{P}_2(s) \doteq \frac{\|A\|}{4}\left(\|A\|^{-1} s - 1\right)^2 .$$

In particular, $\|\mathcal{P}_1(A)\|, \|\mathcal{P}_2(A)\| \leq \|A\|$. If $A \in \mathrm{Re}\{\mathcal{A}\}$, $\sigma(A) \subseteq \mathbb{R}$ and thus $\sigma(\mathcal{P}_1(A)), \sigma(\mathcal{P}_2(A)) \subseteq \mathbb{R}_0^+$, by the spectral mapping property. Additionally, in this case, $\mathcal{P}_1(A), \mathcal{P}_2(A) \in \mathrm{Re}\{\mathcal{A}\}$. Hence, $\mathcal{P}_1(A), \mathcal{P}_2(A) \in \mathcal{A}^+$. This proves the first part of the lemma. In order to prove the second part, write any (possibly non-self-conjugate) C^*-algebra element A as $A = \mathrm{Re}\{A\} + i\mathrm{Im}\{A\}$, and apply the first part to $\mathrm{Re}\{A\}, \mathrm{Im}\{A\} \in \mathrm{Re}\{\mathcal{A}\}$, observing that $\|\mathrm{Re}\{A\}\|, \|\mathrm{Im}\{A\}\| \leq \|A\|$. □

We give now a property of non-zero elements of an arbitrary unital C^*-algebra, which is equivalent to the positivity of these elements. This generalizes Proposition 2.12, which refers to the special case $\mathcal{B}(H)$:

Proposition 4.39 *Let $\mathcal{A}$ be a unital C^*-algebra. The self-conjugate element $A \in \mathrm{Re}\{\mathcal{A}\}\backslash\{0\}$ is positive iff*

$$\left\| \mathbf{1} - \|A\|^{-1} A \right\| \leq 1 \, .$$

If $A \in \mathcal{A}$ is self-conjugate and $\|\mathbf{1} - A\| \leq 1$, then A is positive and $\|A\| \leq 2$.

Proof Assume that $A \in \mathcal{A}$, $A \neq 0$, is positive. Then, $\sigma(\|A\|^{-1} A) \subseteq [0, 1]$, thanks to Proposition 4.31 (iii). Conversely, if $A \in \mathcal{A}$, $A \neq 0$, is self-conjugate and $\sigma(\|A\|^{-1} A) \subseteq [0, 1]$, then A is positive. In fact, the inclusion $\sigma(\|A\|^{-1} A) \subseteq [0, 1]$ is equivalent to

$$\sigma(\mathbf{1} - \|A\|^{-1} A) \cap (1, \infty) = \emptyset \, .$$

But, as $\mathbf{1} - \|A\|^{-1} A$ is positive for all $A \in \mathrm{Re}\{\mathcal{A}\}$, $A \neq 0$, the last condition is equivalent to the inequality $\left\| \mathbf{1} - \|A\|^{-1} A \right\| \leq 1$, thanks to Proposition 4.31 (i). The second part of the proposition is proven in a similar way. □

This last result directly yields that the set of positive elements of any unital C^*-algebra is a convex cone, as desired.

Proposition 4.40 *Let $\mathcal{A}$ be a unital C^*-algebra. $\mathcal{A}^+ \subseteq \mathrm{Re}\{\mathcal{A}\}$ is a closed, pointed, convex cone.*

Proof

1. Clearly, the set

$$\{A \in \mathrm{Re}\{\mathcal{A}\} \ : \ A \neq 0, \ \left\| \mathbf{1} - \|A\|^{-1} A \right\| \leq 1\} \cup \{0\} = \mathcal{A}^+$$

   is a closed cone. See also Proposition 4.39.
2. Let $A, A' \in \mathcal{A}^+$. If $0 \in \{A, A', A + A'\}$, then, trivially, $A + A' \in \mathcal{A}^+$. Assume thus that $\|A\|, \|A'\| > 0$. For $A, A' \in \mathcal{A}^+$, $\|\|A\|\,\mathbf{1} - A\| \leq \|A\|$ and $\|\|A'\|\mathbf{1} - A'\| \leq \|A'\|$. In particular, by the triangle inequality

$$\begin{aligned} \left\|(\|A\| + \left\|A'\right\|)\mathbf{1} - (A + A')\right\| &\leq \|\|A\|\,\mathbf{1} - A\| + \left\|\left\|A'\right\|\mathbf{1} - A'\right\| \\ &\leq \|A\| + \left\|A'\right\| \, . \end{aligned}$$

   Hence,

$$\left\| \mathbf{1} - (\|A\| + \left\|A'\right\|)^{-1}(A + A') \right\| \leq 1$$

and $(\|A\| + \|A'\|)^{-1}(A + A')$ is thus positive, by Proposition 4.39. As $\mathcal{A}^+$ is a cone, it follows that $(A + A')$ is positive, and this proves that $\mathcal{A}^+$ is a convex cone.

3. Suppose finally that $A \in (-\mathcal{A}^+) \cap \mathcal{A}^+$, i.e., that $-A, A \in \mathcal{A}^+$. Then, $A \in \mathrm{Re}\{\mathcal{A}\}$ and $\sigma(A) = \{0\}$, because $-A, A \in \mathcal{A}^+$ implies that $\sigma(A) \subseteq -\mathbb{R}_0^+ \cap \mathbb{R}_0^+ = \{0\}$ and $\sigma(A) \neq \emptyset$ (Lemma 4.26). But this yields $\|A\| = 0$, by Proposition 4.31 (i). Hence, the cone $\mathcal{A}^+$ is pointed.

□

Proposition 4.40 implies that the set $\mathcal{A}^+$ of all positive elements of a unital C^*-algebra $\mathcal{A}$ naturally defines a partial order in $\mathcal{A}$, denoted, as is usual, by $\geq$. Since $\mathcal{A}^+ \subseteq \mathrm{Re}\{\mathcal{A}\}$, it allows us to see any unital C^*-algebras as a $*$-ordered vector space (see Definitions 1.6 and 1.27).

Definition 4.41 (Unital C^*-Algebras as $*$-Ordered Vector Spaces) Any unital C^*-algebra $\mathcal{A}$ is canonically seen as the $*$-ordered vector space whose positive (pointed) cone is $\mathcal{A}^+$, as defined above.

Remark from Exercise 1.8 that the fact, proven above, that, for any unital C^*-algebra $\mathcal{A}$, the corresponding positive cone $\mathcal{A}^+$ is generating implies that any *positive* linear functional $\varphi \in \mathcal{A}'^+$ is real, i.e., it is self-conjugate in $\mathcal{A}'$.

Since any preordered normed space (Definition 1.32) is Archimedean (Exercise 1.33), we deduce from Proposition 4.40 that any unital C^*-algebra is an *Archimedean* $*$-ordered vector space. In particular, order units of a C^*-algebra define norms on the (real) subspace of self-conjugate elements of this algebra (see Definition 1.41). Like in the special case of the algebra of bounded operators on a complex Hilbert space, the norm of self-conjugate elements of any unital C^*-algebra is nothing else than the norm associated with the algebra unit, seen as an order unit, like in the special case of $\mathcal{B}(H)$ (Exercise 2.4):

Lemma 4.42 *Let $\mathcal{A}$ be a unital C^*-algebra. The algebra unit $\mathfrak{1} \in \mathcal{A}$ is also an order unity for $\mathcal{A}$ as a $*$-(pre)ordered vector space and, for all $A \in \mathrm{Re}\{\mathcal{A}\}$, $\|A\|_{\mathfrak{1}} = \|A\|$, where $\|\cdot\|$ is the unique norm of $\mathcal{A}$ as a C^*-algebra.*

Proof

1. Recall that, for any self-conjugate element $A \in \mathcal{A}$, the spectrum $\sigma(A)$ is a compact subset of the real line, by Corollary 4.25 and Lemma 4.21 (i). In particular, as $\sigma(\alpha\mathfrak{1} - A) = \alpha - \sigma(A)$, for sufficiently large $\alpha \in \mathbb{R}$, one has $\sigma(\alpha\mathfrak{1} - A) \subseteq \mathbb{R}_0^+$, that is, $\alpha\mathfrak{1} \geq A$, $\alpha\mathfrak{1} - A \in \mathcal{A}$ being self-conjugate for any real number $\alpha \in \mathbb{R}$. This proves that the algebra unit $\mathfrak{1} \in \mathcal{A}$ is also an order unit for $\mathcal{A}$; see Definition 1.2.
2. By Lemma 4.21 (i), Proposition 4.31 (i), and Corollary 4.25, for any self-conjugate element $A \in \mathcal{A}$, $\|A\|$ or $-\|A\|$ is in the spectrum $\sigma(A) \subseteq \mathbb{R}$. Thus, as

$$\sigma(\alpha\mathfrak{1} - A) = \alpha - \sigma(A) , \qquad \sigma(A - (-\alpha\mathfrak{1})) = \alpha + \sigma(A) ,$$

we obtain from Definitions 1.4 and 1.41 that

$$\begin{aligned}\|A\|_{\mathfrak{1}} &\doteq \inf\{\alpha > 0 \ : \ \alpha\mathfrak{1} \geq A \geq -\alpha\mathfrak{1}\} \\ &= \inf\{\alpha > 0 \ : \ (\alpha\mathfrak{1} - A), \big(A - (-\alpha\mathfrak{1})\big) \in \mathcal{A}^+\} \\ &\geq \|A\| \ .\end{aligned}$$

Similarly, by Proposition 4.31 (iii), i.e., by $\sigma(A) \subseteq [-\|A\|, \|A\|]$, we conclude that $\|A\|_{\mathfrak{1}} \leq \|A\|$.

□

As any unital C^*-algebra $\mathcal{A}$, canonically seen as a $*$-ordered vector space (Definition 4.41), is Archimedean and has an order unit (by Exercise 1.33, Proposition 4.40 and Lemma 4.42), we can apply Proposition 1.47 to $V = \mathcal{A}$ and arrive at the following important fact, which corresponds to a generalization of Corollary 2.8 proven, so far, only for the (unital) C^*-algebra $\mathcal{B}(H)$ of bounded operators on a complex Hilbert space:

Corollary 4.43 *For any unital C^*-algebra $\mathcal{A}$, the topological and the order dual spaces of $\mathcal{A}$ coincide, i.e., $\mathcal{A}^{\mathrm{td}} = \mathcal{A}^{\mathrm{od}}$.*

Observe from the last lemma combined with Corollary 1.48 that any unital C^*-algebra $\mathcal{A}$ is a semisimple (pre)ordered vector space, that is, for any $A, A' \in \mathcal{A}$, $A \neq A'$ only if, for some positive $\varphi \in \mathcal{A}'$, one has $\varphi(A) \neq \varphi(A')$. Recall meanwhile that continuous linear functionals separate elements of any normed space, by the Hahn-Banach theorem. See Theorem 7.40 and Corollary 7.41.

Like in the special case of $\mathcal{B}(H)$ (Corollary 2.14), we have the following important identity for the operator norm of *positive* linear functionals on any unital C^*-algebra.

Proposition 4.44 *Let $\mathcal{A}$ be a unital C^*-algebra. $\varphi \in \mathcal{A}'$ is positive iff $\|\varphi\| = \varphi(\mathfrak{1})$.*

Proof Exercise. *Hint:* By using Exercise 7.77, adapt the proof of Corollary 2.13 to show that, for any $\varphi \in \mathrm{Re}\{A'\}$, φ is positive iff $\|\varphi\|_{\mathrm{op}} = \varphi(\mathfrak{1})$. Finally, adapt the proof of Corollary 2.14 to prove that, for any $\varphi \in A'$, $\|\varphi\|_{\mathrm{op}} = \varphi(\mathfrak{1})$ only if $\varphi \in \mathrm{Re}\{A'\}$. □

Observe from Exercise 7.77 (ii) that, given two unital C^*-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, the space $\mathrm{Re}\{\mathcal{B}(\mathcal{A}_1; \mathcal{A}_2)\}$ can be canonically identified (by restriction) with $\mathcal{B}(\mathrm{Re}\{\mathcal{A}_1\}; \mathrm{Re}\{\mathcal{A}_2\})$. Then, we can generalize Corollary 2.10, which is done for unital C^*-algebras of bounded operators on complex Hilbert spaces, to all unital C^*-algebras.

Corollary 4.45 *Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two unital C^*-algebras. Any element of $\mathcal{L}_{\mathrm{ob}}(\mathrm{Re}\{\mathcal{A}_1\}; \mathrm{Re}\{\mathcal{A}_2\})$, i.e., any order-bounded*[10] *(real) linear transformation*

[10] It maps any order-bounded subset of $\mathrm{Re}\{\mathcal{A}_1\}$ into an order-bounded subset of $\mathrm{Re}\{\mathcal{A}_2\}$. See Definitions 1.16 and 1.23.

$\mathrm{Re}\{\mathcal{A}_1\} \to \mathrm{Re}\{\mathcal{A}_2\}$, *is the restriction of a unique continuous linear transformation* $\mathcal{A}_1 \to \mathcal{A}_2$. *In particular,*

$$\mathcal{L}^+(\mathcal{A}_1; \mathcal{A}_2) \subseteq \mathcal{B}(\mathcal{A}_1; \mathcal{A}_2) ,$$

i.e., every positive linear transformation $\mathcal{A}_1 \to \mathcal{A}_2$ *is continuous. Conversely, the restriction to* $\mathcal{A}_1$ *of any* self-conjugate *continuous linear transformation* $\mathcal{A}_1 \to \mathcal{A}_2$ *defines an order-bounded (real) linear transformation* $\mathrm{Re}\{\mathcal{A}_1\} \to \mathrm{Re}\{\mathcal{A}_2\}$.

Proof By Exercises 1.33, 1.40, and 1.42 together with Proposition 4.40, Definition 4.41, and Lemma 4.42, the real space $\mathrm{Re}\{\mathcal{A}\}$ of self-conjugates elements of a unital C^*-algebra $\mathcal{A}$ is an Archimedean real ordered vector space with an order unit and a locally full norm. Therefore, it suffices to apply Corollary 1.45 on $V_1 = \mathrm{Re}\{\mathcal{A}_1\}$ and $V_2 = \mathrm{Re}\{\mathcal{A}_2\}$ to get the assertion. □

In the following proposition, we summarize the most important properties of the positive cone of any unital C^*-algebra, as well as the properties of the dual positive cone of such algebras. They are essentially the same as the ones gathered in Propositions 2.15 and 2.16 for the special of bounded operators on a complex Hilbert space.

Proposition 4.46 (The Positive Cone of a Unital C^*-Algebra) *Let* $\mathcal{A}$ *be any unital* C^**-algebra. The convex cone* $\mathcal{A}^+$ *has the following properties:*

(i) *It is a norm-closed pointed cone. In particular,* $\mathcal{A}$ *is an Archimedean ordered vector space.*
(ii) *It is (by definition) a subset of* $\mathrm{Re}\{\mathcal{A}\}$. *Thus,* $\mathcal{A}$ *is a* $*$*-(pre)ordered vector space.*
(iii) *For all* $A \in \mathrm{Re}\{\mathcal{A}\}$, $A^2 \in \mathcal{A}^+$.
(iv) *For any* $A \in \mathcal{A}$, *there are* $A^+_{\mathrm{Re}}, A^-_{\mathrm{Re}}, A^+_{\mathrm{Im}}, A^-_{\mathrm{Im}} \in \mathcal{A}^+$ *such that*

$$A = A^+_{\mathrm{Re}} - A^-_{\mathrm{Re}} + i(A^+_{\mathrm{Im}} - A^-_{\mathrm{Im}}) , \qquad \|A^+_{\mathrm{Re}}\|, \|A^-_{\mathrm{Re}}\|, \|A^+_{\mathrm{Im}}\|, \|A^-_{\mathrm{Im}}\| \leq \|A\| .$$

In particular, $\mathcal{A}^+$ *is generating for* $\mathcal{A}$.
(v) *If* $A \in \mathrm{Re}\{\mathcal{A}\}\backslash\{0\}$, *then* $A \in \mathcal{A}^+$ *iff*

$$\left\| \mathbb{1} - \|A\|^{-1} A \right\| \leq 1 .$$

Proposition 4.47 (The Dual Positive Cone of a Unital C^*-Algebra) *Let* $\mathcal{A}$ *be any unital* C^**-algebra. The convex cone* $\mathcal{A}'^+$ *of positive linear functionals on* $\mathcal{A}$ *has the following properties:*

(i) *It is a pointed cone whose elements are all bounded linear functionals, that is,* $\mathcal{A}'^+ \subseteq \mathcal{A}^{\mathrm{td}}$. *In particular, the dual space* $\mathcal{A}'$ *is an ordered vector space. In fact, note that* $\mathcal{A}'^+$ *is pointed, because* $\mathcal{A}^+$ *is generating for* $\mathcal{A}$. *See Exercise 1.29.*
(ii) *It is a subset of* $\mathrm{Re}\{\mathcal{A}'\}$, *i.e., any positive linear functional on* $\mathcal{A}$ *is necessarily self-conjugate. Thus,* $\mathcal{A}'$ *is a* $*$*-(pre)ordered vector space.*
(iii) $\mathcal{A}^{\mathrm{td}} = \mathcal{A}^{\mathrm{od}}$. *In particular,* $\mathcal{A}'^+ \subseteq \mathcal{A}^{\mathrm{td}}$ *is generating for* $\mathcal{A}^{\mathrm{td}}$.

(iv) *Any* $\varphi \in \mathcal{A}'$ *is a positive linear functional iff* $\|\varphi\|_{\text{op}} = \varphi(\mathbb{1})$.
(v) $\mathcal{A}'^+$ *separates the elements of* $\mathcal{A}$, *that is, for any* $A, A' \in \mathcal{A}$, $A \neq A'$, *there is* $\varphi \in \mathcal{A}'^+$ *such that* $\varphi(A) \neq \varphi(A')$. *(Recall that this property refers to the semisimplicity of* $\mathcal{A}$ *as a preordered vector space, deduced from Corollary* 1.48.*)*

We introduced above the notion of positivity of C^*-algebra elements (Definition 4.36), but studied the corresponding positive and dual positive cones only in the unital case, so far. To conclude this section and for completeness, we make a few important, albeit simple, remarks on *non*-unital C^*-algebras as $*$-ordered vector spaces.

In fact, observe from Proposition 4.16 that any non-unital C^*-algebra $\mathcal{A}$ can be canonically seen as a $*$-ideal of its unitization $\tilde{\mathcal{A}}$, which is a unital C^*-algebra. See Definitions 7.16 and 7.63, as well as Definitions 7.24 and 7.69. In particular, $\mathcal{A}$, being a self-conjugate subspace of a $*$-ordered vector space, is canonically a $*$-ordered vector space. Observe that proceeding in this way, we end up with a positive cone for $\mathcal{A}$ that is nothing else than the set $\mathcal{A}^+$ of self-conjugate elements with positive spectrum, as defined in Definition 4.36.

A version of Propositions 4.46 and 4.47 holds true in non-unital case. To prove this statement, we need the continuous functional calculus, given below by Proposition 4.100 and Corollary 4.101. We postpone this study to Sect. 4.6.1, which can be performed *independently* of the assertions presented below. Thus, we first give a version of Proposition 4.46, which gather important properties of positive cones of unital C^*-algebras, which hold also in the non-unital case.

Proposition 4.48 (The Positive Cone of a C^*-Algebra) *Let* $\mathcal{A}$ *be any (not necessarily unital)* C^**-algebra. The convex cone* $\mathcal{A}^+$ *has the following properties:*

(i) *It is a norm-closed pointed cone. In particular,* $\mathcal{A}$ *is an Archimedean ordered vector space.*
(ii) $\mathcal{A}$ *is a* $*$*-ordered vector space.*
(iii) *For all* $A \in \mathrm{Re}\{\mathcal{A}\}$, $A^2 \in \mathcal{A}^+$.
(iv) *For any* $A \in \mathcal{A}$, *there are* $A_{\mathrm{Re}}^+, A_{\mathrm{Re}}^-, A_{\mathrm{Im}}^+, A_{\mathrm{Im}}^- \in \mathcal{A}^+$ *such that*

$$A = A_{\mathrm{Re}}^+ - A_{\mathrm{Re}}^- + i(A_{\mathrm{Im}}^+ - A_{\mathrm{Im}}^-) , \qquad \|A_{\mathrm{Re}}^+\|, \|A_{\mathrm{Re}}^-\|, \|A_{\mathrm{Im}}^+\|, \|A_{\mathrm{Im}}^-\| \leq \|A\| .$$

In particular, $\mathcal{A}^+$ *is generating for* $\mathcal{A}$.

Proof In fact, in the non-unital case, (i)–(iii) are direct consequences of Proposition 4.46 (i)–(iii) applied to the unitization $\tilde{\mathcal{A}}$ of an arbitrary non-unital C^*-algebra $\mathcal{A}$. By contrast, (iv) cannot be proven as in Lemma 4.38, for the proof of the latter essentially uses the existence of a unit in the algebra itself. To prove this point for an arbitrary (not necessarily unital) C^*-algebra $\mathcal{A}$, one uses Proposition 4.102 (ii). We omit the details. □

Similarly, a version of Proposition 4.47, which gather important properties of dual positive cones of unital C^*-algebras, holds true also in the non-unital case. To prove this more general version, that is, Proposition 4.53, aside from the functional calculus, we need to introduce the concept of approximate units.

Definition 4.49 (Approximate Units of Subspaces) Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $\mathcal{V} \subseteq \mathcal{A}$ a vector subspace, which is also seen as an ordered space, being a subspace of the ordered vector space $\mathcal{A}$. A net $(E_i)_{i\in I}$ in $\mathcal{V}$ is called "approximate unit[11]" for $\mathcal{V}$ if the following properties hold true:

(i) $(E_i)_{i\in I} \subseteq \mathcal{A}^+$ is an increasing net of positive elements.
(ii) $\|E_i\| \leq 1$ for all $i \in I$.
(iii) For all $B \in \mathcal{V}$, $\lim_{i\in I} E_i B = B$.

The existence of approximate units in non-unital C^*-algebras is a priori unclear. We show below that any C^*-algebra does possess an approximate unity. This is a particular case of the following assertion on right ideals of C^*-algebras (Definition 7.24):

Proposition 4.50 (Approximate Units of Ideals) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. Every right ideal $\mathcal{I} \subseteq \mathcal{A}$ possesses an approximate unit.*

Proof We can assume, without loss of generality, that $\mathcal{A}$ is unital, because any right ideal of a non-unital algebra is also a right ideal of the unitization of this algebra. See Definitions 7.16 and 7.24. Take any right ideal $\mathcal{I} \subseteq \mathcal{A}$. Let I be the set of all nonempty finite subsets of $\mathcal{I}$. It is an upward directed set with respect to inclusion; see Definition 1.13. For all $i \in I$, we define the positive element

$$B_i \doteq \sum_{B\in i} BB^* \in \mathcal{A} .$$

Since $\mathcal{I}$ is a right ideal of $\mathcal{A}$, we have $B_i \in \mathcal{I}$ for all $i \in I$. Then, for any $i \in I$, let

$$E_i \doteq |i| B_i (\mathbb{1} + |i| B_i)^{-1} \in \mathcal{I} .$$

Note from Proposition 4.100 (the functional calculus) and Corollary 4.103 that E_i is positive and, using additionally Proposition 4.31 (i),

$$\|E_i\| = r_{\mathcal{A}}(A) \doteq \max \sigma(E_i) \leq \sup \left\{ |i| x (\mathbb{1} + |i| x)^{-1} : x \in \mathbb{R}_0^+ \right\} = 1 .$$

Moreover, again by Proposition 4.100 and Corollary 4.103, for any $B \in \mathcal{I}$ and $i \in I$ with $B \in i$, we have that

$$\begin{aligned}(E_i B - B)(E_i B - B)^* &= (E_i - \mathbb{1}) BB^* (E_i - \mathbb{1})^* \\ &\leq \sum_{C\in i} (E_i - \mathbb{1}) CC^* (E_i - \mathbb{1})^*\end{aligned}$$

[11] This is also called "approximate identity" in the literature; see, e.g., [51, Definition 2.2.17].

$$= (E_i - \mathbb{1})B_i(E_i - \mathbb{1})^*$$
$$= (|i|B_i(\mathbb{1} + |i|B_i)^{-1} - \mathbb{1})B_i(|i|B_i(\mathbb{1} + |i|B_i)^{-1} - \mathbb{1})^*$$
$$= B_i(\mathbb{1} + |i|B_i)^{-2} \leq \mathbb{1}|i|^{-1} .$$

Thus, for every $B \in \mathcal{I}$, one has that $\|E_i B - B\| \leq |i|^{-\frac{1}{2}}$ for all $i \in I$ such that $B \in i$. In particular, $\lim_{i\in I} E_i B = B$ for all $B \in \mathcal{I}$. As a consequence, $\mathcal{I}$ possesses an approximate unit, since one can always extract an increasing subnet from an upward directed set. □

Corollary 4.51 *Every non-unital C^*-algebra possesses an approximate unit.*

We now exploit the existence of an approximate unit in any C^*-algebra to show the following result, which is interesting in itself and will have applications later on.

Proposition 4.52 (The Dual Positive Cone of Non-unital C^*-Algebras) *Let $\mathcal{A}$ be any C^*-algebra with its unitization denoted by $\tilde{\mathcal{A}}$:*

(i) *The mapping $\varphi \mapsto \varphi|_{\mathcal{A}}$ from $\tilde{\mathcal{A}}'^{+}$ to $\mathcal{A}'^{+}$ is surjective, i.e., any positive functional on $\mathcal{A}$ is the restriction of a positive functional on $\tilde{\mathcal{A}}$.*
(ii) *The restriction mapping has a right inverse $\mathcal{A}'^{+} \to \tilde{\mathcal{A}}'^{+}$, $\varphi \mapsto \tilde{\varphi}$, which preserves the norm of functionals, i.e., for all $\varphi \in \mathcal{A}'^{+}$, $\|\tilde{\varphi}\|_{\mathrm{op}} = \|\varphi\|_{\mathrm{op}}$ and $\varphi = \tilde{\varphi}|_{\mathcal{A}}$. Additionally, for all $\varphi \in \mathcal{A}'^{+}$, $\tilde{\varphi}$ is the infimum of the set of all positive functionals on $\tilde{\mathcal{A}}$ that extend φ.*

Proof

1. Note first from Proposition 4.48 (iv) together with Corollary 1.35 (applied to $\mathrm{Re}\{\mathcal{A}\}$) that $\mathcal{A}'^{+} \subseteq \mathcal{A}^{\mathrm{td}}$.
2. Take any $\varphi \in \mathcal{A}'^{+}$ and let $(E_i)_{i\in I}$ be any approximate unit of $\mathcal{A}$; see Definition 4.49. Recall that $\mathcal{A}$ is a $*$-ideal (Definitions 7.24 and 7.69) of $\tilde{\mathcal{A}}$ (Definitions 7.16 and 7.63). Observing that $\mathcal{A}$ is a self-conjugate subspace of $\tilde{\mathcal{A}}$, we infer from Proposition 4.46 (iv) together with Exercise 1.8 that $\varphi \in \mathrm{Re}\{\mathcal{A}'\}$. Hence, the sesquilinear form $(A, B) \mapsto \varphi(A^*B)$ from $\mathcal{A} \times \mathcal{A}$ to $\mathbb{C}$ (Definition 7.200) is a (complex) scalar semiproduct on $\mathcal{A}$. Thus, from the Cauchy-Schwarz inequality (Proposition 7.202) and $\mathcal{A}'^{+} \subseteq \mathcal{A}^{\mathrm{td}}$, for all $A \in \mathcal{A}$,
$$|\varphi(A)|^2 = \lim_{i\in I} \varphi(E_i A)^2 \leq \sup_{i\in I} \varphi(E_i^2)\varphi(A^*A) \leq \varphi(A^*A)\,\|\varphi\|_{\mathrm{op}} .$$
Recall that $\|E_i\| \leq 1$ by definition of approximating units.
3. Define the linear functional $\tilde{\varphi} \in \tilde{\mathcal{A}}'$ by
$$\tilde{\varphi}((\alpha, A)) \doteq \alpha\,\|\varphi\|_{\mathrm{op}} + \varphi(A) , \qquad (\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}} .$$
Observe that $\|\tilde{\varphi}\|_{\mathrm{op}} \geq \|\varphi\|_{\mathrm{op}}$, by definitions of the operator norm of linear functionals.

4. Keeping in mind the definition of the product in $\tilde{\mathcal{A}}$, given in Definition 7.16, we remark that for all $(\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}}$,

$$(\alpha, A)^*(\alpha, A) = (|\alpha|^2, \bar{\alpha}A + \alpha A^* + A^*A) .$$

Then, we deduce from $\varphi \in \mathcal{A}'^+ \subseteq \mathrm{Re}\{\mathcal{A}'\}$ and the inequality of point 2. that for all $(\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}}$,

$$\begin{aligned}\tilde{\varphi}((\alpha, A)^*(\alpha, A)) &= |\alpha|^2 \|\varphi\|_{\mathrm{op}} + \bar{\alpha}\varphi(A) + \alpha\varphi(A^*) + \varphi(A^*A) \\ &\geq |\alpha|^2 \|\varphi\|_{\mathrm{op}} - 2|\alpha|\sqrt{\varphi(A^*A)\|\varphi\|_{\mathrm{op}}} + \varphi(A^*A) \\ &= \left(|\alpha|\sqrt{\|\varphi\|_{\mathrm{op}}} - \sqrt{\varphi(A^*A)}\right)^2 \geq 0 .\end{aligned}$$

5. From Corollary 4.103, $\tilde{\varphi} \in \tilde{\mathcal{A}}'^+$, i.e., it is a positive linear functional on the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$. From Proposition 4.44, $\|\tilde{\varphi}\|_{\mathrm{op}} = \tilde{\varphi}(\mathfrak{1}) = \|\varphi\|_{\mathrm{op}}$, because the unit of $\tilde{\mathcal{A}} \doteq \mathbb{C} \times \mathcal{A}$ is $\mathfrak{1} = (1, 0)$.
6. Let $\tilde{\varphi}' \in \tilde{\mathcal{A}}'^+$ be any positive extension of φ. Then, for some constant $C \in \mathbb{R}$,

$$\tilde{\varphi}'((\alpha, A)) \doteq \alpha C + \varphi(A) , \qquad (\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}} .$$

By Proposition 4.44, $C = \tilde{\varphi}'(\mathfrak{1}) = \|\tilde{\varphi}'\|_{\mathrm{op}} \geq \|\varphi\|_{\mathrm{op}}$. Finally, we remark that for all $(\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}}$,

$$\tilde{\varphi}'((\alpha, A)^*(\alpha, A)) - \tilde{\varphi}((\alpha, A)^*(\alpha, A)) = |\alpha|^2 \left(C - \|\tilde{\varphi}\|_{\mathrm{op}}\right) \geq 0$$

and we thus conclude that $\tilde{\varphi}' \geq \tilde{\varphi}$, thanks to Corollary 4.103.

□

We are now in a position to prove a general version of Proposition 4.47, which gather important properties of dual positive cones of (not necessarily unital) C^*-algebras.

Proposition 4.53 (The Dual Positive Cone of a C^*-Algebra) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. The convex cone $\mathcal{A}'^+$ of positive linear functionals on $\mathcal{A}$ has the following properties:*

(i) *It is a pointed cone, i.e., the dual space $\mathcal{A}'$ is an ordered vector space.*
(ii) *It is a subset of* $\mathrm{Re}\{\mathcal{A}'\}$*, i.e., any positive linear functional on $\mathcal{A}$ is necessarily self-conjugate. Thus, $\mathcal{A}'$ is a $*$-(pre)ordered vector space.*
(iii) $\mathcal{A}^{\mathrm{td}} = \mathcal{A}^{\mathrm{od}}$*. In particular, $\mathcal{A}'^+ \subseteq \mathcal{A}^{\mathrm{td}}$ is generating for $\mathcal{A}^{\mathrm{td}}$.*
(iv) *$\mathcal{A}'^+$ separates the elements of $\mathcal{A}$, that is, for any $A, A' \in \mathcal{A}$, $A \neq A'$, there is $\varphi \in \mathcal{A}'^+$ such that $\varphi(A) \neq \varphi(A')$. (Recall that this property refers to the semisimplicity of $\mathcal{A}$ as a preordered vector space.)*

Proof In fact, in the non-unital case, (i) holds true, because $\mathcal{A}^+$ is generating for $\mathcal{A}$, thanks to Exercise 1.29 and Proposition 4.48 (iv). Assertions (ii) and (iv) follow from Proposition 4.52 combined with Proposition 4.47 (ii and v) applied to the unitization $\tilde{\mathcal{A}}$ of an arbitrary non-unital C^*-algebra $\mathcal{A}$, by observing that $\mathcal{A}$ is a self-conjugate subspace of $\tilde{\mathcal{A}}$. To prove (iii), let $\varphi \in \mathcal{A}^{\mathrm{td}}$. Then, by the Hahn-Banach extension theorem (Theorem 7.40), there is a continuous linear functional $\tilde{\varphi} \in \tilde{\mathcal{A}}^{\mathrm{td}}$ whose restriction to $\mathcal{A} \subseteq \tilde{\mathcal{A}}$ is φ. By Proposition 4.47 (iii), $\tilde{\varphi}$ is a linear combination of positive linear functionals. Thus, as restrictions of positive functional functionals are positive, one has that

$$\mathcal{A}^{\mathrm{td}} \subseteq \mathcal{A}^{\mathrm{od}} \doteq \mathrm{span}(\mathcal{A}'^{+}) ;$$

see Definition 1.25. As positive functionals on $\tilde{\mathcal{A}}$ are continuous (Proposition 4.47 (iii)), by Proposition 4.52, it follows that $\mathcal{A}^{\mathrm{od}} \subseteq \mathcal{A}^{\mathrm{td}}$. □

4.4 Ideals and Quotients of C^*-Algebras

Ideals and quotients are standard notions and are defined in Sect. 7.1. The ideal of an algebra refers to Definition 7.24, while a $*$-ideal of a $*$-algebra is given by Definition 7.69. By Exercise 7.68, a $*$-ideal is just a self-conjugate ideal of the $*$-algebra.

The quotient of a vector space by one of its subspace is defined in Definition 7.7. By taking ideals $\mathcal{I}$ as such subspaces in an algebra $\mathcal{A}$, the corresponding quotient $\mathcal{A}/\mathcal{I}$ can be made into an algebra, named the "quotient of the algebra $\mathcal{A}$ by its ideal $\mathcal{I}$." See Definition 7.28. Remark that $\mathcal{A}/\mathcal{I}$ is unital whenever $\mathcal{A}$ is unital. Mutatis mutandis for an associative or commutative algebra $\mathcal{A}$. For $*$-algebras $\mathcal{A}$ the same construction can of course be done. By taking now a $*$-ideal $\mathcal{I}$, one defines the so-called quotient $\mathcal{A}/\mathcal{I}$ of the $*$-algebra $\mathcal{A}$ by its self-conjugate ideal $\mathcal{I}$, which is now a $*$-algebra; see Definition 7.72.

In the particular context of C^*-algebras, the concepts of closed ideals and closed $*$-ideals turn out to be equivalent. In other words, any closed ideal of a C^*-algebra is automatically self-conjugate. This fact is another important consequence of the existence of approximate units for ideals given by Proposition 4.50.

Lemma 4.54 (Ideals Are $*$-Ideals) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. A closed subspace $\mathcal{I} \subseteq \mathcal{A}$ is a $*$-ideal (of $\mathcal{A}$ seen as a $*$-algebra) iff it is an ideal (of $\mathcal{A}$ seen as an algebra).*

Proof Recall that any $*$-ideal of a $*$-algebra is (by a direct consequence of its definition) an ideal; see Definition 7.69 and Exercise 7.68. Conversely, assume that $\mathcal{I} \subseteq \mathcal{A}$ is a closed ideal. By Proposition 4.50, the (right) ideal $\mathcal{I}$ has an approximate unit $(E_i)_{i\in I}$. Observe that, for any $B \in \mathcal{I}$,

$$\lim_{i\in I} \left\| B^* E_i - B^* \right\| = \lim_{i\in I} \| E_i B - B \| = 0 \, .$$

As $\mathcal{I}$ is a left ideal and $E_i \in \mathcal{I}$, $B^* E_i \in \mathcal{I}$. Hence, as $\mathcal{I}$ is closed, $B^* \in \mathcal{I}$. ◻

It follows from Lemma 4.54 that the quotient $\mathcal{A}/\mathcal{I}$ of the C^*-algebra $\mathcal{A}$ by any closed ideal $\mathcal{I}$ is a $*$-algebra. Note in this case that $\mathcal{A}/\mathcal{I}$ is also a Banach $*$-algebra with norm equal to

$$\|[A]\|_{\mathcal{A}/\mathcal{I}} \doteq \inf\left\{ \|A'\| \ : \ A' \in [A] \right\} = \inf_{B\in\mathcal{I}} \|A - B\| \ , \qquad A \in \mathcal{A} \, ,$$

thanks to Exercises 7.34 and 7.87. In fact, in the particular context of C^*-algebras, this quotient is even a C^*-algebra.

Proposition 4.55 (Quotients of C^*-Algebras Are C^*-Algebras) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed ideal. For any approximate unit $(E_i)_{i\in I}$ of $\mathcal{I}$,*

$$\|[A]\|_{\mathcal{A}/\mathcal{I}} = \lim_{i\in I} \|A - E_i A\| \ .$$

Additionally, the quotient $*$-algebra $\mathcal{A}/\mathcal{I}$ is a C^*-algebra.

Proof By Exercise 4.13 (ii), we can assume without loss of generality that $\mathcal{A}$ is unital. Let $(E_i)_{i\in I}$ be an approximate unit of a closed ideal $\mathcal{I} \subseteq \mathcal{A}$. Clearly, for any $A \in \mathcal{A}$,

$$\|[A]\|_{\mathcal{A}/\mathcal{I}} = \inf_{B\in\mathcal{I}} \|A - B\| \leq \lim_{i\in I} \inf_{j\geq i} \left\| A - E_j A \right\| \ .$$

On the other hand, for any $A \in \mathcal{A}$ and $B \in \mathcal{I}$,

$$\lim_{i\in I} \sup_{j\geq i} \left\| A - E_j A \right\| = \lim_{i\in I} \sup_{j\geq i} \left\| (\mathfrak{1} - E_j)(A - B) \right\| \leq \|A - B\| \ .$$

Note that $\left\| \mathfrak{1} - E_j \right\| \leq 1$ because E_j is positive and $\left\| E_j \right\| \leq 1$, by definition of approximate units (Definition 4.49). Hence,

$$\lim_{i\in I} \sup_{j\geq i} \left\| A - E_j A \right\| \leq \inf_{B\in\mathcal{I}} \|A - B\| = \|[A]\|_{\mathcal{A}/\mathcal{I}} \leq \lim_{i\in I} \inf_{j\geq i} \left\| A - E_j A \right\|$$

and it follows that, for any $A \in \mathcal{A}$,

$$\|[A]\|_{\mathcal{A}/\mathcal{I}} = \lim_{i\in I} \|A - E_i A\| \ .$$

Now, by Exercise 7.87 and Lemma 4.54, $\mathcal{A}/\mathcal{I}$ is a Banach $*$-algebra. Thus, in order to prove that $\mathcal{A}/\mathcal{I}$ is a C^*-algebra, we only have to show that, for all $A \in \mathcal{A}$,

$$\left\|[A]^*[A]\right\|_{\mathcal{A}/\mathcal{I}} = \|[A]\|^2_{\mathcal{A}/\mathcal{I}} .$$

See Definition 4.1 (iii). Observe first that, for all $A \in \mathcal{A}$,

$$\begin{aligned}
\|[A]\|^2_{\mathcal{A}/\mathcal{I}} &= \lim_{i\in I} \|A - E_i A\|^2 \\
&= \lim_{i\in I} \left\|(A^* - A^* E_i)(A - E_i A)\right\| \\
&= \lim_{i\in I} \left\|(A^* A - 2A^* E_i A - A^* E_i^2 A\right\| \\
&\geq \left\|[A^* A]\right\|_{\mathcal{A}/\mathcal{I}} = \left\|[A]^*[A]\right\|_{\mathcal{A}/\mathcal{I}} ,
\end{aligned}$$

because $2A^* E_i A + A^* E_i^2 A \in \mathcal{I}$, as $\mathcal{I}$ is an ideal. On the other hand, as the closed ideal $\mathcal{I}$ is self-conjugate (thanks to Lemma 4.54), for all $B \in \mathcal{I}$,

$$\begin{aligned}
\|[A]\|^2_{\mathcal{A}/\mathcal{I}} &= \inf_{C\in\mathcal{I}} \left\|A^* - C^*\right\| = \left\|[A^*]\right\|^2_{\mathcal{A}/\mathcal{I}} \\
&= \lim_{i\in I} \left\|(A^* - E_i A^*)^*\right\|^2 \\
&= \lim_{i\in I} \left\|(A^* - E_i A^*)(A - A E_i)\right\| \\
&= \lim_{i\in I} \left\|(\mathbf{1} - E_i) A^* A (\mathbf{1} - E_i)\right\| \\
&= \lim_{i\in I} \left\|(\mathbf{1} - E_i)(A^* A + B)(\mathbf{1} - E_i)\right\| \\
&\leq \left\|A^* A + B\right\| .
\end{aligned}$$

Recall that $\|\mathbf{1} - E_i\| \leq 1$ and note that, for any $B \in \mathcal{I}$,

$$\lim_{i\in I} (\mathbf{1} - E_i) B = 0 ,$$

by definition of approximating units of the (right) ideal $\mathcal{I}$. Finally, by taking the infimum

$$\|[A]\|^2_{\mathcal{A}/\mathcal{I}} \leq \inf_{B\in\mathcal{I}} \left\|A^* A + B\right\| = \left\|[A^* A]\right\|_{\mathcal{A}/\mathcal{I}} = \left\|[A]^*[A]\right\|_{\mathcal{A}/\mathcal{I}} .$$

□

Recall from Exercise 7.70 that, for any two $*$-algebras $\mathcal{A}$, $\mathcal{B}$ and any $*$-homomorphism $\Theta : \mathcal{A} \to \mathcal{B}$, $\ker(\Theta) \subseteq \mathcal{A}$ is a $*$-ideal of $\mathcal{A}$. It follows from

Proposition 4.55 that any closed ideal of an arbitrary (not necessarily unital) C^*-algebra is of this following form.

Corollary 4.56 (Closed Ideals as Kernels of $*$-Homomorphisms) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. Take $\mathcal{I} \subseteq \mathcal{A}$. $\mathcal{I}$ is a closed ideal of $\mathcal{A}$ iff there is a C^*-algebra $\mathcal{B}$ and a $*$-homomorphism $\Theta : \mathcal{A} \to \mathcal{B}$ such that* $\ker(\Theta) = \mathcal{I}$.

Proof Let $\mathcal{A}$, $\mathcal{B}$ be two C^*-algebras and $\Theta : \mathcal{A} \to \mathcal{B}$ any $*$-homomorphism. By Exercise 7.70, $\ker(\Theta) \subseteq \mathcal{A}$ is a ($*$-)ideal of $\mathcal{A}$. To show that the kernel $\ker(\Theta)$ is closed, notice from Exercise 7.64 that there is a $*$-homomorphism $\tilde{\Theta} : \tilde{\mathcal{A}} \to \tilde{\mathcal{B}}$, whose restriction to $\mathcal{A}$ is Θ, keeping in mind that the unitizations $\tilde{\mathcal{A}}$, $\tilde{\mathcal{B}}$ of $\mathcal{A}$, $\mathcal{B}$ are C^*-algebras (Proposition 4.16). From Lemma 4.96 (ii) $\tilde{\Theta}$ is continuous and, consequently, Θ is also continuous. Thus, $\ker(\Theta)$ is a closed subspace of $\mathcal{A}$. Conversely, if $\mathcal{I} \subseteq \mathcal{A}$ is a closed ideal, then, by Proposition 4.55, the quotient $\mathcal{A}/\mathcal{I}$ is a C^*-algebra. Let $\mathfrak{q} : \mathcal{A} \to \mathcal{A}/\mathcal{I}$ be defined as in Definition 7.7. Then, by Exercise 7.73 (i), $\mathfrak{q}$ is a $*$-homomorphism and, by construction, $\mathcal{I} = \ker(\mathfrak{q})$. □

The following proposition contributes a useful criterion to prove simplicity of C^*-algebras. In particular, it will be used later on to show that quantum spin and fermion algebras are simple. See Lemma 4.153 and Corollary 4.167, as well as Sect. 5.1. Recall that this property refers to Definition 7.48 (iii) for normed algebras: A normed algebra $\mathcal{A}$ is "simple" if $\{0\}$ and $\mathcal{A}$ are the only *closed* ideals of $\mathcal{A}$. One must avoid confusion of this particular definition with the one of general (i.e., not normed) simple algebras, for which the closeness of the ideals is not meaningful. See Definition 7.24 (iii).

Proposition 4.57 (Simplicity of C^*-Algebras via Simplicity of Dense Nets of C^*-Sublgebras) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed ideal. Let $(\mathcal{A}_i)_{i\in I}$ be any increasing (with respect to inclusion of the subalgebras) of C^*-subalgebras of $\mathcal{A}$ such that the union $\cup_{i\in I}\mathcal{A}_i \subseteq \mathcal{A}$ is dense in $\mathcal{A}$. Define, for every $i \in I$, the ideal $\mathcal{I}_i \doteq \mathcal{I} \cap \mathcal{A}_i$ of $\mathcal{A}_i$. Then, the union $\cup_{i\in I}\mathcal{I}_i \subseteq \mathcal{I}$ is dense in $\mathcal{I}$. Additionally, all C^*-subalgebras $\mathcal{A}_i$, $i \in I$, are simple only if the full algebra $\mathcal{A}$ is itself simple.*

Proof Assume, by contradiction, that, for some $A \in \mathcal{I}$,

$$\inf\{\|A - B\| : \ B \in \cup_{i\in I}\mathcal{I}_i\} = \varepsilon > 0\,.$$

Choose a sequence $A_n \in \mathcal{A}_{i_n}$, $n \in \mathbb{N}$, converging to A such that $\|A - A_n\| \le \varepsilon/2$. Then, in particular, for all $n \in \mathbb{N}$ and $B_n \in \mathcal{I}_{i_n} \subseteq \mathcal{I}$, one has that $\|A_n - B_n\| \ge \varepsilon/2$. Invoking Corollary 4.56, take any C^*-algebra $\mathcal{B}$ and a $*$-homomorphism $\Theta : \mathcal{A} \to \mathcal{B}$ such that $\ker(\Theta) = \mathcal{I}$. By construction, $\ker(\Theta|_{\mathcal{A}_i}) = \mathcal{I} \cap \mathcal{A}_i \doteq \mathcal{I}_i$ for any $i \in I$. Note that one can prove that, for all $i \in I$ and $A \in \mathcal{A}_i$, one has that

$$\|\Theta(A)\| = \inf\{\|A - B\| : B \in \ker(\Theta|_{\mathcal{A}_i})\} \doteq \|[A]\|_{\mathcal{A}_i/\ker(\Theta|_{\mathcal{A}_i})}$$

and Θ is continuous. This is explicitly proven in Proposition 4.97, applied for instance to the $*$-homomorphism $\Theta|_{\mathcal{A}_i} : \mathcal{A}_i \to \mathcal{B}$. In particular, for all $n \in \mathbb{N}$, $\|\Theta(A_n)\| \geq \varepsilon/2$ and by continuity of Θ, it follows that $\|\Theta(A)\| \geq \varepsilon/2$, i.e., $A \notin \ker(\Theta) = \mathcal{I}$. As a consequence, the union $\cup_{i\in I}\mathcal{I}_i \subseteq \mathcal{I}$ is dense in $\mathcal{I}$.

Finally, assume now that all C^*-subalgebras $\mathcal{A}_i$, $i \in I$, are simple. Take any closed ideal $\mathcal{I}$ of $\mathcal{A}$. Using the above construction, $\mathcal{I}_i \doteq \mathcal{I} \cap \mathcal{A}_i$ is a closed ideal of $\mathcal{A}_i$ for all $i \in I$. So, if $\mathcal{A}_i$ is simple for all $i \in I$, then $\mathcal{I}_i = \{0\}$ or $\mathcal{A}_i$. If there is $i \in I$ such that $\mathcal{I}_i = \mathcal{A}_i \neq \{0\}$, then, for all $j \geq i$, $\mathcal{I}_j \supseteq \mathcal{I}_i \neq \{0\}$ and $\mathcal{I}_j = \mathcal{A}_j$. As $\cup_{i\in I}\mathcal{A}_i \subseteq \mathcal{A}$ and $\cup_{i\in I}\mathcal{I}_i \subseteq \mathcal{I}$ are dense in closed sets $\mathcal{A}$ and $\mathcal{I}$, it follows that $\mathcal{I} = \mathcal{A}$ when there is $i \in I$ such that $\mathcal{I}_i = \mathcal{A}_i \neq \{0\}$. In other words, if all C^*-subalgebras $\mathcal{A}_i$, $i \in I$, are simple, then the full algebra $\mathcal{A}$ is also simple. □

The following proposition is a version of Proposition 4.44 for non-unital C^*-algebras.

Proposition 4.58 *Let $\mathcal{A}$ be a (not necessarily unital) C^*-algebra. A linear functional $\varphi \in \mathcal{A}'$ is positive iff it is continuous and, for any approximate unity $(E_i)_{i\in I}$ for $\mathcal{A}$, one has that*

$$\|\varphi\|_{\mathrm{op}} = \lim_{i\in I} \varphi(E_i^2) \ .$$

Proof

1. Assume that $\varphi \in \mathcal{A}'$ is positive. By Proposition 4.53 (iii), $\mathcal{A}'^{+} \subseteq \mathcal{A}^{\mathrm{td}}$. Thus, φ is continuous. Additionally, by Proposition 4.53 (ii), $\varphi \in \mathrm{Re}\{A'\}$. Hence, the sesquilinear form $(A, B) \mapsto \varphi(A^*B)$ from $\mathcal{A} \times \mathcal{A}$ to $\mathbb{C}$ is a scalar semiproduct on $\mathcal{A}$. Take any approximate unity $(E_i)_{i\in I}$ for $\mathcal{A}$. Recall that $\mathcal{A}$ is a $*$-ideal of $\tilde{\mathcal{A}}$, while bearing in mind Definition 4.49 and Proposition 4.50. In particular, we can take an approximate unity $(E_i)_{i\in I}$ for $\mathcal{A}$. From the Cauchy-Schwarz inequality (Proposition 7.202), for all $A \in \mathcal{A}$,

$$|\varphi(A)|^2 = \lim_{i\in I} \varphi(E_i A)^2 = \lim_{i\in I} \inf_{j\geq i} \varphi(E_i A)^2$$
$$\leq \varphi(A^*A) \lim_{i\in I} \inf_{j\geq i} \varphi(E_j^2) \leq \left\|A^*A\right\|^2 \|\varphi\|_{\mathrm{op}} \lim_{i\in I} \inf_{j\geq i} \varphi(E_j^2) \ .$$

By taking a sequence $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, with $\|A_n\| = 1$ and $\lim_{n\to\infty} |\varphi(A)| = \|\varphi\|_{\mathrm{op}}$, we infer from the above estimate that

$$\|\varphi\|_{\mathrm{op}} \leq \lim_{i\in I} \inf_{j\geq i} \varphi(E_j^2) \leq \lim_{i\in I} \sup_{j\geq i} \varphi(E_j^2) \leq \|\varphi\|_{\mathrm{op}} \ .$$

Therefore,

$$\lim_{i\in I} \varphi(E_i^2) = \|\varphi\|_{\mathrm{op}}$$

2. Assume now that $\varphi \in \mathcal{A}'$ satisfies

$$\lim_{i\in I} \varphi(E_i^2) = \|\varphi\|_{\text{op}} \tag{4.1}$$

for some approximate unity $(E_i)_{i\in I}$ for $\mathcal{A}$. For simplicity and without loss of generality, we may additionally assume that $\|\varphi\|_{\text{op}} = 1$. Note that if $\mathcal{A}$ is unital, one may take $E_i = \mathfrak{1}$, and the positivity of φ follows from Proposition 4.44. Thus, let $\mathcal{A}$ be a non-unital C^*-algebra. We define an extension $\tilde{\varphi} \in \tilde{\mathcal{A}}'$ of the linear functional $\varphi \in \mathcal{A}'$ to the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$ as follows:

$$\tilde{\varphi}((\alpha, A)) \doteq \alpha + \varphi(A) , \qquad (\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}} .$$

See for instance the point 3 of the proof of Proposition 4.52, keeping in mind that here $\|\varphi\|_{\text{op}} = 1$. Then, for all $(\alpha, A) \in \mathbb{C} \times \mathcal{A} \doteq \tilde{\mathcal{A}}$,

$$|\tilde{\varphi}((\alpha, A))| = \lim_{i\in I} |\varphi(\alpha E_i^2 + E_i^2 A)| \leq \|(\alpha, A)\|_{\tilde{\mathcal{A}}} ,$$

thanks to (4.1). Note that we used that, for all $A \in \mathcal{A}$,

$$\lim_{i\in I} E_i^2 A = \lim_{i\in I} (E_i A + E_i(E_i A - A)) = \lim_{i\in I} E_i A = A ,$$

by Definition 4.49. See also Exercise 4.13. From the last estimate, we conclude that $\|\tilde{\varphi}\|_{\text{op}} = \tilde{\varphi}(\mathfrak{1}) = 1$. Thus, from Proposition 4.44, $\tilde{\varphi}$ is positive. Hence, φ is positive, being the restriction of a positive functional on $\tilde{\mathcal{A}}$.

□

One important, albeit simple, consequence of the last proposition is the fact that the norm of positive functionals on (not necessarily unital) C^*-algebras behaves additively.

Corollary 4.59 *Let $\mathcal{A}$ be a (not necessarily unital) C^*-algebra. For any pair $\varphi_1, \varphi_2 \in \mathcal{A}'^+$ of positive linear functionals, one has that*

$$\|\varphi_1 + \varphi_2\|_{\text{op}} = \|\varphi_1\|_{\text{op}} + \|\varphi_2\|_{\text{op}} .$$

In particular, as norms are positively homogeneous, $\|\cdot\|_{\text{op}}$ is an affine function on $\mathcal{A}'^+$, i.e., for all $\varphi_1, \varphi_2 \in \mathcal{A}'^+$ and $\lambda \in [0, 1]$,

$$\|\lambda\varphi_1 + (1-\lambda)\varphi_2\|_{\text{op}} = \lambda \|\varphi_1\|_{\text{op}} + (1-\lambda) \|\varphi_2\|_{\text{op}} .$$

Exercise 4.60 Let $\mathcal{A}$ be an arbitrary C^*-algebra (which could already be unital) and $\tilde{\mathcal{A}}$ its unitization. Show that the transformation $\varphi \mapsto \tilde{\varphi}$ of Proposition 4.52, from $\mathcal{A}'^+$ to $\tilde{\mathcal{A}}'^+$, behaves additively, i.e., for any pair $\varphi_1, \varphi_2 \in \mathcal{A}'^+$ of positive linear functionals, one has

$$\widetilde{\varphi_1 + \varphi_2} = \tilde{\varphi}_1 + \tilde{\varphi}_2 .$$

4.5 States

We discuss now the general notion of states of C^*-algebras. This is a key notion in the algebraic formulation of quantum mechanics.

Definition 4.61 (States of a C^*-Algebra) Let $\mathcal{A}$ be a (not necessarily unital) C^*-algebra.

(i) A linear functional $\rho \in \mathcal{A}'$ is a "state" of $\mathcal{A}$ if it is positive (i.e., $\rho \in \mathcal{A}'^{+}$) and norm-one (i.e., $\|\rho\|_{\text{op}} = 1$). $E(\mathcal{A})$ denotes the set of all states of $\mathcal{A}$.
(ii) The state $\rho \in E(\mathcal{A})$ is said to be an "extreme state" of $\mathcal{A}$ when it is an extreme point of the convex set $E(\mathcal{A})$, i.e., if $\rho = \lambda\rho' + (1-\lambda)\rho''$ for $\lambda \in (0, 1)$ and $\rho', \rho'' \in E(\mathcal{A})$ only if $\rho' = \rho'' = \rho$. The set of all extreme states of $\mathcal{A}$ is denoted by $\mathcal{E}(\mathcal{A}) \subseteq E(\mathcal{A})$.
(iii) A state $\rho \in E(\mathcal{A})$ is "faithful" if, for all $A \in \mathcal{A}^+$, $\rho(A) = 0$ only if $A = 0$.

Observe from Corollary 4.59 that the set of states $E(\mathcal{A})$ is always convex. Additionally, by Proposition 4.44, states of *unital* C^*-algebras can be equivalently defined as follows.

Definition 4.62 (States of a Unital C^*-Algebra) Let $\mathcal{A}$ be a unital C^*-algebra. A linear functional $\rho \in \mathcal{A}'$ is a "state" of $\mathcal{A}$ if it is positive (i.e., $\rho \in \mathcal{A}'^{+}$) and normalized (i.e., $\rho(\mathbb{1}) = 1$).

Exercise 4.63 Let $\mathcal{A}$ be any non-unital C^*-algebra. Show that any state on $\mathcal{A}$ has a unique extension to a state on the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$. In particular, $E(\mathcal{A})$ is canonically seen as a subset of $E(\tilde{\mathcal{A}})$.

Hint: Consider Proposition 4.52.

As in the special case of the algebras of bounded operators on complex Hilbert spaces, note that any positive linear functional on an arbitrary C^*-algebra is a positive multiple of a state on this algebra. Observe also that $\mathcal{E}(\mathcal{A})$ is nonempty for any C^*-algebra $\mathcal{A}$. This follows from the Krein-Milman theorem [18, Theorem 3.23] (see Proposition 7.334 in the case of a separable C^*-algebra $\mathcal{A}$) for locally convex spaces, and this will be discussed in more detail below.

Exercise 4.64 (Separation Properties of States—I) Let $\mathcal{A}$ be a (not necessarily unital) C^*-algebra. Prove that:

(i) $\rho(A) = 0$ for all $\rho \in E(\mathcal{A})$ iff $A = 0$.
(ii) $\rho(A) \in \mathbb{R}$ for all $\rho \in E(\mathcal{A})$ iff $A^* = A$.

The last exercise generalizes to arbitrary C^*-algebras Corollary 7.234, which only refers to the algebra of bounded operators on a Hilbert space.

Example 4.65 Let $C(K;\mathbb{C})$ be the C^*-algebra of complex-valued continuous functions on a compact metric space K. For all $p \in K$, define the linear functional $\rho_p : C(K;\mathbb{C}) \to \mathbb{C}$ by $\rho_p(f) \doteq f(p)$. Observe that $\rho_p(\mathfrak{1}) = 1$ and $\rho_p(f) \geq 0$ for all $f \in C(K;\mathbb{C})^+$ (see Example 4.37). In other words, $\rho_p \in E(C(K;\mathbb{C}))$.

If the metric space M is compact, then observe that $C_{\mathrm{b}}(M;\mathbb{C}) = C(M;\mathbb{C})$. Thus, from Lemma 4.75 (i), any state on $C(M;\mathbb{C})$ extends to a state on $\mathfrak{M}_{\mathrm{b}}(M;\mathbb{C}) \supseteq C(M;\mathbb{C})$. In fact, it is easy to see that any state $\rho_p \in E(C(M;\mathbb{C}))$, $p \in M$, as in the last example, extends to a probability measure $\mu_p \in E(\mathfrak{M}_{\mathrm{b}}(M;\mathbb{C}))$ (in the sense of Definition 4.10).

Exercise 4.66 Let M be any metric space and define, for all $p \in M$, the state $\mu_p \in E(\mathfrak{M}_{\mathrm{b}}(M;\mathbb{C}))$ by

$$\mu_p(f) \doteq f(p) , \qquad f \in \mathfrak{M}_{\mathrm{b}}(M;\mathbb{C}) .$$

Show that μ_p is a probability measure (see Definition 4.10), i.e., it is σ-order-continuous, positive, and normalized.

This fact motivates the following generalization of the states of Example 4.65:

Example 4.67 Let K be any compact metric space and μ a probability measure (i.e., a state on $\mathfrak{M}_{\mathrm{b}}(K;\mathbb{C})$; see Definition 4.10). We define the state $\rho_\mu \in E(C(K;\mathbb{C}))$ as being the restriction of μ to $C(K;\mathbb{C})$.

It turns out that any state on the unital C^*-algebra $C(K;\mathbb{C})$ of complex-valued continuous functions on a compact metric space K is of this form. This refers to the celebrated Riesz-Markov theorem (in the special case of compact metric spaces), rephrased in terms of states on C^*-algebras of functions.

Theorem 4.68 (Riesz-Markov) *Let $\rho \in E(C(K;\mathbb{C}))$, where K is an arbitrary compact metric space. There is a unique probability measure $\mu_\rho \in E(\mathfrak{M}_{\mathrm{b}}(K;\mathbb{C}))$ that extends ρ to $\mathfrak{M}_{\mathrm{b}}(K;\mathbb{C}) \supseteq C(K;\mathbb{C})$. The mapping $\rho \mapsto \mu_\rho$ from $E(C(K;\mathbb{C}))$ to $E(\mathfrak{M}_{\mathrm{b}}(K;\mathbb{C}))$ is an affine one-to-one correspondence.*

In other words, the states of $C(K;\mathbb{C})$ are naturally identified with the probability measures on K. Observe that Example 4.67 and Theorem 4.68 can be extended to the more general case of compact Hausdorff spaces, if one defines probability measures as being not only normalized positive Borel measures but also regular.[12] See, e.g., [57, Theorem 29.1]. In fact, in a compact metric space, any finite positive Borel measure is regular. For a complete exposition of this subject, along with a

[12] For a measure μ on a compact Hausdorff space K, it means that, for any Borel set $B \subseteq K$, $\mu(B) = \sup\{\mu(C) : C \subseteq B,\ C \text{ compact}\} = \inf\{\mu(O) : B \subseteq O,\ O \text{ open}\}$.

proof of the Riesz-Markov theorem done in great generality, we recommend [3, Section 14.3].

Notice that in Sect. 4.7, we show that any separable commutative C^*-algebras is equivalent (i.e., $*$-isomorphic) to the algebra of complex functions on a compact metric space. See Proposition 4.124. Thus, because of the above theorem, the theory of C^*-algebras and its states is frequently seen as a noncommutative extension of classical probability theory.[13]

For any (not necessarily unital) C^*-algebra $\mathcal{A}$, it turns out that extremality in the set $E(\mathcal{A})$ of states, which is a geometric property, is equivalent to a condition exclusively linked to the order structure of the dual space $\mathcal{A}'$.

Definition 4.69 (Pure State) Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. We say that the state $\rho \in E(\mathcal{A})$ is "pure" if

$$[0, \rho] \doteq \mathcal{A}'^{+} \cap (\rho - \mathcal{A}'^{+}) = \{\alpha\rho \ : \ \alpha \in [0, 1]\} \subseteq \mathcal{A}'^{+} \ ;$$

see Definition 1.4. In other words, ρ is pure when the only positive linear functionals $\mathcal{A} \to \mathbb{C}$ below ρ are the multiples $\alpha\rho$, $\alpha \in [0, 1]$, of ρ.

For any state $\rho \in E(\mathcal{A})$, note that, obviously,

$$\{\alpha\rho \ : \ \alpha \in [0, 1]\} \subseteq \{\rho' \in \mathcal{A}' \ : \ \rho \succeq \rho' \succeq 0\} = [0, \rho] \ .$$

The converse of this inclusion is however only satisfied if the state is extreme. To prove this fact, we use the following result.

Lemma 4.70 *Let $\mathcal{A}$ be any C^*-algebra and define $\overline{B}_1(0) \subseteq \mathcal{A}'^{+}$ as being the unit closed ball in $\mathcal{A}'^{+}$, i.e.,*

$$\overline{B}_1(0) \doteq \{\varphi \in \mathcal{A}'^{+} \ : \ \|\varphi\|_{\mathrm{op}} \leq 1\} \supseteq E(\mathcal{A}) \ .$$

A positive functional $\varphi \in \mathcal{A}'^{+}$ is an extreme point of $\overline{B}_1(0)$ iff it is either zero or a pure state.

Proof

1. Take $\varphi \in \mathcal{A}'^{+}$. The positive functional $\varphi = 0 \in \mathcal{A}'^{+}$ is clearly extreme in $\overline{B}_1(0)$. Thus, we assume that $\varphi \neq 0$.
2. Observe additionally that if $0 < \|\varphi\|_{\mathrm{op}} < 1$, then $\|\varphi\|_{\mathrm{op}}^{-1} \varphi \in \overline{B}_1(0)$ and

$$\varphi = \|\varphi\|_{\mathrm{op}} \left(\|\varphi\|_{\mathrm{op}}^{-1} \varphi\right) + (1 - \|\varphi\|_{\mathrm{op}})0 \ .$$

   In other words, φ is not extreme in this case, and we may assume that $\varphi \in E(\mathcal{A})$, i.e., $\|\varphi\|_{\mathrm{op}} = 1$.

[13] In fact, for technical reasons, like the normality of expectation values, one rather considers, in this context, von Neumann algebras instead of general C^*-algebras.

3. Assume that $\rho \in E(\mathcal{A})$ is pure and suppose that $\rho = \lambda\rho' + (1-\lambda)\rho''$ for some $\lambda \in (0, 1)$ and $\rho', \rho'' \in E(\mathcal{A})$. Then, clearly,

$$\rho \succeq \lambda\rho' \succeq 0 \qquad \text{and} \qquad \rho \succeq (1-\lambda)\rho'' \succeq 0 \, .$$

Hence, by purity of ρ, for some $\alpha', \alpha'' \in [0, 1]$, one has $\lambda\rho' = \alpha'\rho$ and $(1-\lambda)\rho'' = \alpha''\rho$. By the fact that states are norm-one functionals and Corollary 4.59, it then follows that $\alpha' = \lambda$ and $\alpha'' = (1-\lambda)$. Hence, $\rho' = \rho'' = \rho$ and ρ is thus extreme.

4. Assume now that $\rho \in E(\mathcal{A})$ is not pure. Then, there is a state $\rho' \in E(\mathcal{A})$, $\rho' \neq \rho$, and a constant $\lambda \in (0, 1)$ such that $\rho \succeq \lambda\rho'$. Define the positive functional

$$\rho'' \doteq \frac{1}{1-\lambda}(\rho - \lambda\rho') \, .$$

By construction

$$\rho = (1-\lambda)\rho'' + \lambda\rho' \, .$$

In particular, as $\rho' \neq \rho$, one also has that $\rho'' \neq \rho'$. Moreover, by Corollary 4.59, $\left\|\rho''\right\|_{\mathrm{op}} = 1$, i.e., $\rho'' \in E(\mathcal{A})$. Hence, ρ is not extreme.

□

Observe that the above lemma implies the existence of pure states for any C^*-algebra $\mathcal{A}$. In fact, the set $\overline{B}_1(0) \subseteq \mathcal{A}'^+$ of positive functionals is weak*-compact and convex. Thus, from the Krein-Milman theorem [18, Theorem 3.23] (see Proposition 7.334 in the case of a separable C^*-algebra $\mathcal{A}$), it is the closure of the linear combinations of its extreme points. Hence, as $\overline{B}_1(0) \neq \{0\}$, there must be pure states for $\mathcal{A}$. See Sect. 4.5.1 for more details.

Proposition 4.71 (Pure = Extreme) *Let $\mathcal{A}$ be any C^*-algebra. The state $\rho \in E(\mathcal{A})$ is pure iff it is extreme in $E(\mathcal{A})$.*

Proof Note that $\rho \in \overline{B}_1(0) \subseteq \mathcal{A}'^+$ is a state iff it maximizes the function $\varphi \mapsto \|\varphi\|_{\mathrm{op}}$ in $\overline{B}_1(0)$. By Corollary 4.59 this function is affine in $\overline{B}_1(0)$, and, hence, $E(\mathcal{A})$ is a face[14] of $\overline{B}_1(0)$, i.e., $\rho \in E(\mathcal{A})$ is extreme in $E(\mathcal{A})$ iff it is extreme in $\overline{B}_1(0)$. With this observation, the proposition directly follows from the last lemma. □

Again by the Krein-Milman theorem [18, Theorem 3.23] (see Proposition 7.334 in the case of a separable C^*-algebra $\mathcal{A}$), note that the last proposition implies the existence of extreme states for any C^*-algebra $\mathcal{A}$.

[14] A face F of a convex set K is defined to be a subset of K with the property that, if $\rho = \lambda_1\rho_1 + \cdots + \lambda_n\rho_n \in F$ with $\rho_1, \ldots, \rho_n \in K$, $\lambda_1, \ldots, \lambda_n \in (0, 1)$ and $\lambda_1 + \cdots + \lambda_n = 1$, then $\rho_1, \ldots, \rho_n \in F$.

Corollary 4.72 *Let $C(K;\mathbb{C})$ be the unital C^*-algebra of continuous complex-valued functions on a compact metric space K. Then, the states of Example 4.65 are extreme states:*

$$\{\rho_p \;:\; p \in K\} \subseteq \mathcal{E}(C(K;\mathbb{C}))\ .$$

Proof

1. For all $p \in K$, let $C_p(K;\mathbb{C}) \subseteq C(K;\mathbb{C})$ be the subspace of continuous functions $f : K \to \mathbb{C}$ such that $f(p) = 0$. For any fixed $p \in K$, let $\rho \in E(C(K;\mathbb{C}))$ and $\alpha \in (0, 1]$ be such that $\rho_p \succeq \alpha\rho$. Then, for all $f \in C_p(K;\mathbb{C})^+$, $\alpha\rho(f) = 0$, that is $\rho(f) = \rho_p(f) = 0$.
2. Recall that

$$\mathrm{Re}\{C_p(K;\mathbb{C})\} = C_p(K;\mathbb{C})^+ - C_p(K;\mathbb{C})^+\ .$$

Thus, as

$$C_p(K;\mathbb{C}) = \mathrm{Re}\{C_p(K;\mathbb{C})\} + i\mathrm{Re}\{C_p(K;\mathbb{C})\}\ ,$$

by linearity of ρ, $\rho(f) = \rho_p(f) = 0$ for all $f \in C_p(K;\mathbb{C})$.
3. Observe that any function $f \in C(K;\mathbb{C})$ is the sum of an element of $C_p(K;\mathbb{C})$ and a constant function:

$$f = (f - f(p)\mathbb{1}) + f(p)\mathbb{1}\ .$$

From the normalization of states, one then arrives at

$$\begin{aligned}\rho(f) &= \rho((f - f(p)\mathbb{1}) + f(p)\mathbb{1}) \\ &= \rho(f(p)\mathbb{1}) = f(p) = \rho_p(f)\end{aligned}$$

for all $f \in C(K;\mathbb{C})$. That is, $\rho = \rho_p$ and, hence, ρ_p is pure. By Proposition 4.71, ρ_p is extreme.
4. Observe that we also have shown that, given $p \in K$, for any state $\rho \in E(\mathcal{A})$, one has $\rho = \rho_p$ iff $\rho(C_p(K;\mathbb{C})^+) = \{0\}$.

□

By using the Riesz-Markov theorem, one can improve Corollary 4.72 as follows.

Corollary 4.73 *Let $C(K;\mathbb{C})$ be the unital C^*-algebra of complex-valued continuous functions on a compact metric space K. Then, the extremes states are the ones of Example 4.65:*

$$\{\rho_p \;:\; p \in K\} = \mathcal{E}(C(K;\mathbb{C}))\ .$$

Idea of Proof

1. Note that any probability μ measure on a compact metric space K, being a regular measure, has a well-defined "support," i.e., there is a smallest *closed* subset $\mathrm{supp}(\mu) \subseteq K$, for which $\mu(\mathrm{supp}(\mu)) = \mu(K) = 1$. Here, for any Borel set $\Omega \subseteq K$, $\mu(\Omega) \doteq \mu(\chi_\Omega)$. See remarks following Definition 4.10.
2. If $\mathrm{supp}(\mu)$ consists of one single point, i.e., $\mathrm{supp}(\mu) = \{p\}$ for some $p \in K$, then clearly $\rho_\mu = \rho_p$. If $\mathrm{supp}(\mu)$ is not such a singleton, then it splits into two *disjoint* Borel sets $\Omega', \Omega'' \subseteq K$ with $\mu(\Omega'), \mu(\Omega'') > 0$. Hence, if the state $\rho \in E(C(K;\mathbb{C}))$ is not of the form $\rho = \rho_p$ for some $p \in K$, then there are disjoint Borel sets $\Omega', \Omega'' \subseteq K$ with $\mu_\rho(\Omega'), \mu_\rho(\Omega'') \in (0,1)$ and $\mu_\rho(\Omega') + \mu_\rho(\Omega'') = 1$.
3. This implies, in turn, that

$$\rho = \mu_\rho(\Omega')\rho' + \mu_\rho(\Omega'')\rho'' ,$$

where $\rho', \rho'' \in E(C(K;\mathbb{C}))$, $\rho' \neq \rho''$, are defined by

$$\rho'(f) \doteq \mu_\rho(\Omega')^{-1} \int_{\Omega'} f(p)\, \mu_\rho(\mathrm{d}p) , \quad \rho''(f) \doteq \mu_\rho(\Omega'')^{-1} \int_{\Omega''} f(p)\, \mu_\rho(\mathrm{d}p) .$$

In other words, ρ is not in this case an extreme state.

□

Observe from the last corollary that the extreme (or pure) states ρ on $C(K;\mathbb{C})$ are exactly those for which

$$\rho((f - \rho(f))^2) = 0 \tag{4.2}$$

for all $f \in \mathrm{Re}\{C(K;\mathbb{C})\}$, that is, the variance of the distribution of any "physical quantity," represented by the (arbitrary) real-valued continuous function f on the compact metric space K, is zero. The extreme states of $C(K;\mathbb{C})$ are so-called characters, which are, by definition, linear functionals that are multiplicative and self-conjugate. See Definition 4.117. In fact, notice from Lemma 4.122 that the extreme states of any commutative C^*-algebra (like $C(K;\mathbb{C})$) are exactly its characters. By contrast, in the case of an arbitrary (not necessarily commutative) C^*-algebra $\mathcal{A}$, any character of $\mathcal{A}$, i.e., a non-zero linear mapping $\varphi : \mathcal{A} \to \mathbb{C}$ satisfying

$$\varphi(A_1 A_2) = \varphi(A_1)\varphi(A_2) , \ \varphi(A_1^*) = \overline{\varphi(A_1)} , \qquad A_1, A_2 \in \mathcal{A} ,$$

is indeed an extreme state of $E(\mathcal{A})$, thanks to Lemma 4.118, but not every extreme state on $\mathcal{A}$ is of this type. Another simple example of a character, for both commutative and noncommutative cases, is given as follows.

Example 4.74 (Unitizations Admit Characters) Let $\mathcal{A}$ be any C^*-algebra and $\tilde{\mathcal{A}} \doteq \mathbb{C} \times \mathcal{A}$ its unitization. Recall that $\mathcal{A}$ is a (closed) $*$-ideal of $\tilde{\mathcal{A}}$ and observe that the quotient $\tilde{\mathcal{A}}/\mathcal{A}$ may be identified with $\mathbb{C}$. Clearly, the mapping $(\alpha, A) \mapsto \alpha$ from $\tilde{\mathcal{A}}$ to $\tilde{\mathcal{A}}/\mathcal{A} \equiv \mathbb{C}$ is a character, being a non-zero $*$-homomorphism $\tilde{\mathcal{A}} \to \mathbb{C}$. See Exercise 7.74.

Notice additionally that, again in contrast to the commutative case, the existence of characters (like in Examples 4.65 and 4.74) in the *noncommutative* case is generally wrong, by the Bell-Kochen-Specker theorem [58, 59]. However, for some noncommutative C^*-algebras, the extremality in sets of so-called invariant states turns out to be equivalent to a property similar to (4.2), also corresponding to vanishing statistical dispersions of states, albeit in a weaker sense than for characters. This refers to the notion of ergodicity, which appears in many important cases and is conceptually very natural in statistical mechanics. This point will be discussed in detail in Sect. 5.4.

We come back now to the general (i.e., not necessarily commutative) case of C^*-algebras $\mathcal{A}$. Given a fixed element $A \in \mathcal{A}$, we show in the sequel that, for any value $a \in \sigma(A)$ in the spectrum of A, there is a state that implements a as the expected value for A. This is reminiscent of a trivial property of extreme states given in Example 4.65, except that in the general case this can only be done for *fixed* algebra elements. Important consequences of this fact will be also discussed. We start with the following lemma.

Lemma 4.75 *Let $\mathcal{A}$ be any C^*-algebra:*

(i) *If $\mathcal{B} \subseteq \mathcal{A}$ is a C^*-subalgebra of $\mathcal{A}$ and $\tilde{\rho} \in E(\mathcal{B})$ (i.e., $\tilde{\rho}$ is a state on $\mathcal{B}$), then there is a state $\rho \in E(\mathcal{A})$ whose restriction to $\mathcal{B}$ is $\tilde{\rho}$.*
(ii) *If $\mathcal{A}$ is unital then, for all $A \in \mathcal{A}$ and $a \in \sigma(A)$, there is a state $\rho \in E(\mathcal{A})$ such that $\rho(A) = a$.*
(iii) *If $\mathcal{A}$ is non-unital then, for all $A \in \mathrm{Re}\{\mathcal{A}\}$ and $a \in \sigma(A)\backslash\{0\}$, there is a state $\rho \in E(\mathcal{A})$ such that $\rho(A) = a$. Additionally, there is a sequence of states $\rho_n \in E(\mathcal{A})$, $n \in \mathbb{N}$, such that $\lim_{n\to\infty} \rho_n(A) = 0 \in \sigma(A)$. (Recall from Exercise 4.20 that 0 must be a spectral value of A, $\mathcal{A}$ being non-unital.)*

Proof

(i) Assume first that $\mathcal{B}$ is a unital C^*-subalgebra of $\mathcal{A}$. Take any $\tilde{\rho} \in E(\mathcal{B})$. As $\tilde{\rho}$ is a state, $\|\tilde{\rho}\|_{\mathrm{op}} = \tilde{\rho}(\mathfrak{1}) = 1$. See Proposition 4.44 and Definition 4.62. By the Hahn-Banach extension theorem (Theorem 7.40), there is $\rho \in \mathcal{A}^{\mathrm{td}}$ extending $\tilde{\rho}$ to $\mathcal{A}$ in such a way that $\|\rho\|_{\mathrm{op}} = \|\tilde{\rho}\|_{\mathrm{op}}$. In particular, $\rho(\mathfrak{1}) = 1$ and therefore, $\|\rho\|_{\mathrm{op}} = \rho(\mathfrak{1}) = 1$, which in turn implies that ρ is a state on $\mathcal{A}$. Now, if the C^*-subalgebra $\mathcal{B} \subseteq \mathcal{A}$ is not unital, then we use the unitizations $\tilde{\mathcal{B}}$ and $\tilde{\mathcal{A}}$, of respectively $\mathcal{B}$ and $\mathcal{A}$. Note in this case that $\tilde{\mathcal{B}}$ can be canonically seen as a unital C^*-subalgebra of $\tilde{\mathcal{A}}$. By Exercise 4.63, $\tilde{\rho} \in E(\mathcal{B})$ (uniquely) extends to a state on $\tilde{\mathcal{B}}$, and using the unital case just explained, we can extend it to a state on $\tilde{\mathcal{A}}$. By construction, the restriction of the latter to $\mathcal{A}$ is positive and norm-one, that is, it is a state on $\mathcal{A}$ extending $\tilde{\rho}$.

(ii) Take any $A \in \mathcal{A}$. Note that if $A = \alpha\mathbb{1}$, $\alpha \in \mathbb{C}$, then $\sigma(A) = \{\alpha\}$ and $\rho(A) = \alpha$ for any state $\rho \in E(\mathcal{A})$. Therefore, suppose that $A \notin \mathbb{C}\mathbb{1}$. Let $a \in \sigma(A)$ and $X_A \subseteq \mathcal{A}$ be the two-dimensional subspace span$\{\mathbb{1}, A\} \subseteq \mathcal{A}$. Define the linear functional $\tilde{\rho} : X_A \to \mathbb{C}$ by

$$\tilde{\rho}(\alpha\mathbb{1} + \beta A) \doteq \alpha + \beta a, \quad \alpha, \beta \in \mathbb{C} .$$

With this definition, for all $\alpha, \beta \in \mathbb{C}$, $\alpha + \beta a \in \sigma(\alpha\mathbb{1} + \beta A)$. Hence,

$$|\tilde{\rho}(\alpha\mathbb{1} + \beta A)| = |\alpha + \beta a| \leq r_{\mathcal{A}}(\alpha\mathbb{1} + \beta A) \leq \|\alpha\mathbb{1} + \beta A\| ,$$

where we recall that $r_{\mathcal{A}}(B)$ is the spectral radius of any algebra element $B \in \mathcal{A}$, defined by

$$r_{\mathcal{A}}(B) \doteq \max\{|\lambda| \ : \ \lambda \in \sigma(B)\} \leq \|B\| .$$

See also Definition 4.28 and Theorem 4.29. In particular, $\|\tilde{\rho}\|_{\mathrm{op}} \leq 1$. For $\tilde{\rho}(\mathbb{1}) = 1$, one has $\|\tilde{\rho}\|_{\mathrm{op}} = 1$. By the Hahn-Banach extension theorem (Theorem 7.40), there is $\rho \in \mathcal{A}^{\mathrm{td}}$ extending $\tilde{\rho}$ in such a way that $\|\rho\|_{\mathrm{op}} = \|\tilde{\rho}\|_{\mathrm{op}}$. In particular, $\|\rho\|_{\mathrm{op}} = \rho(\mathbb{1}) = 1$ (i.e., $\rho \in E(\mathcal{A})$) and $\rho(A) = a$.

(iii) This assertion refers to Lemma 4.121.

□

Given a fixed element $A \in \mathcal{A}$ of a unital C^*-algebra $\mathcal{A}$, we show in Lemma 4.75 (ii) that we can construct a state that implements any value in the spectrum $\sigma(A)$ of A as its expectation value. In fact, Lemma 4.75 can be improved by imposing the purity (or extremality) of states.

Theorem 4.76 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra.*

(i) *If $\mathcal{B} \subseteq \mathcal{A}$ is a C^*-subalgebra of $\mathcal{A}$ and $\tilde{\rho} \in \mathcal{E}(\mathcal{B})$ (i.e., $\tilde{\rho}$ is a* pure *state on $\mathcal{B}$), then there is a pure state $\rho \in \mathcal{E}(\mathcal{A})$ whose restriction to $\mathcal{B}$ is $\tilde{\rho}$.*
(ii) *If $\mathcal{A}$ is unital then, for all $A \in \mathcal{A}$ and $a \in \sigma(A)$, there is a* pure *state $\rho \in \mathcal{E}(\mathcal{A})$ such that $\rho(A) = a$.*
(iii) *If $\mathcal{A}$ non-unital, then, for all $A \in$ Re$\{\mathcal{A}\}$ and $a \in \sigma(A)\backslash\{0\}$, there is a* pure *state $\rho \in E(\mathcal{A})$ such that $\rho(A) = a$.*

The proof of this theorem uses the Krein-Milman theorem [18, Theorem 3.23] (see Proposition 7.334 in the case of a separable C^*-algebra $\mathcal{A}$), and it is omitted. For more details, see the proof of [51, Proposition 2.3.24] yielding Theorem 4.76 (i). Note also that Lemma 4.121 implies Lemma 4.75 (iii) and Theorem 4.76 (iii).

One may use the first part of the above theorem to extend Corollary 4.73 to a class of non-unital C^*-algebras of continuous functions.

Corollary 4.77 *Let $C_0(K_0; \mathbb{C})$ be the (generally non-unital) C^*-algebra of complex-valued continuous functions that "decay at infinity" on the metric space $K_0 \subseteq K\backslash\{p_0\}$, where K is a compact metric space and $p_0 \in K$ is arbitrary. Then, the extreme states of $C_0(K_0; \mathbb{C})$ are like those of Example 4.65:*

$$\mathcal{E}(C_0(K_0;\mathbb{C})) \subseteq \{\rho_p \ : \ p \in K_0\} \ .$$

Proof Exercise. *Hint:* Note that $C_0(K_0;\mathbb{C})$ can be canonically identified with a C^*-subalgebra of $C(K;\mathbb{C})$. □

In fact, notice that it directly follows from Lemma 4.118 that

$$\mathcal{E}(C_0(K_0;\mathbb{C})) \supseteq \{\rho_p : p \in K_0\}$$

and, thus, the extreme states of the (generally non-unital) C^*-algebra $C_0(K_0;\mathbb{C})$, exactly as in the case of the (unital) C^*-algebra $C(K;\mathbb{C})$, can be identified with the points of the domain K_0.

Since Lemma 4.75 (ii) links states to spectra of algebra elements, states can be used to characterize properties that are related with the spectrum of an element, via, for instance, Proposition 4.31. This yields the following important result.

Proposition 4.78 (Separation Properties of States—II) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $A \in \mathcal{A}$:*

(i) $\rho(A) \geq 0$ *for all* $\rho \in E(\mathcal{A})$ *iff* $A \in \mathcal{A}^+$.
(ii) *If A is normal (i.e., $AA^* = A^*A$), then there is $\rho \in E(\mathcal{A})$ such that* $|\rho(A)| = \|A\|$.

Proof Exercise. Hint: In the non-unital case, consider the unitization $\tilde{\mathcal{A}}$ of the C^*-algebra $\mathcal{A}$. □

We now extend the notion of Gelfand transform of Definition 2.33 to general C^*-algebras.

Definition 4.79 (Gelfand Transform in General C^*-Algebras) Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $A \in \mathcal{A}$. We define a linear transformation $\Xi : \mathcal{A} \to \mathcal{F}(E(\mathcal{A});\mathbb{C})$ by

$$\Xi(A)\,(\rho) \doteq \rho(A) \ , \qquad A \in \mathcal{A},\ \rho \in E(\mathcal{A}) \ .$$

The function $\Xi(A) : E(\mathcal{A}) \to \mathbb{C}$ is called here the "Gelfand transform" of $A \in \mathcal{A}$.

Recall that, strictly speaking, what is most commonly called "Gelfand transform" refers rather to the case of *commutative* C^*-algebras and the domain of such transforms is not the set of all states on the given algebra but only the set of *extreme* states. This special and very important case will be discussed at the end of the present section, in Sect. 4.7.

$\Xi : \mathcal{A} \to \mathcal{F}(E(\mathcal{A});\mathbb{C})$ is clearly a positive linear transformation, i.e., a morphism of (pre)ordered vector spaces. In particular, it is self-conjugate, as $\mathcal{A}$ is linearly generated by its positive cone (Proposition 4.48 (iv)). From the first part of Exercise 4.64, Ξ is injective. Because of Proposition 4.78 (i), it is bipositive. Moreover, like in the special case of bounded operators on complex Hilbert spaces, its image is a vector subspace of $C(E(\mathcal{A});\mathbb{C})$, where $E(\mathcal{A})$ is endowed with the metric associated with the operator norm for linear functionals. Thus, any unital C^*-algebra $\mathcal{A}$ is equivalent, as a $*$-ordered vector space, to the subspace of $C(E(\mathcal{A}); \mathbb{C})$ formed by the Gelfand transforms of all elements of $\mathcal{A}$. We show below that, if $\mathcal{A}$ is a *separable* (unital) C^*-algebra, it is possible to define a metric on $E(\mathcal{A})$ such that $E(\mathcal{A})$ becomes a compact metric space and the Gelfand transforms of elements of $\mathcal{A}$ are still continuous functions. In fact, a similar result holds for all (unital) C^*-algebras, but the topology making $E(\mathcal{A})$ a compact Hausdorff space is generally not metrizable. We will not discuss this more general case here.

Given any unital C^*-algebra $\mathcal{A}$, observe from Lemma 4.75 (ii) that, for any $A \in \mathcal{A}$

$$\sigma(A) \subseteq [\Xi(A)](E(\mathcal{A})) ,$$

and not only

$$\sigma(A) \subseteq \overline{[\Xi(A)](E(\mathcal{A}))} .$$

Compare this with Corollary 2.31. Similarly, if $\mathcal{A}$ is a non-unital C^*-algebra then, from Lemma 4.75 (iii), for all $A \in \mathrm{Re}\{\mathcal{A}\}$,

$$\sigma(A) \subseteq [\Xi(A)](E(\mathcal{A})) \cup \{0\} \subseteq \overline{[\Xi(A)](E(\mathcal{A}))} .$$

In fact, the images $[\Xi(A)](E(\mathcal{A}))$, $A \in \mathcal{A}$, play a similar role in the theory of abstract C^*-algebra, to the role played by the numerical range of bounded linear operators on Hilbert spaces. See Definition 7.241 and Proposition 7.242.

4.5.1 Weak Topology for States in the Separable Case*

In the present paragraph, $\mathcal{A}$ is any *separable* C^*-algebra, i.e., $\mathcal{A}$ possesses a *countable* dense subset $\mathcal{A}_0 \subseteq \mathcal{A}$. In this case, any countable dense subset of $\mathcal{A}$ naturally defines a metric in the set $E(\mathcal{A})$ of states. Note additionally that the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$ is separable, whenever $\mathcal{A}$ is separable. It turns out that the respective metrics of dense sets always yield the same topology. Moreover, $E(\mathcal{A})$ is compact with respect to any of such a metric, whenever the C^*-algebra $\mathcal{A}$ is *unital*. Observe that many C^*-algebras that are relevant for quantum statistical mechanics are separable. This is the case, for instance, for quantum lattices (spins and fermions). However, other important C^*-algebras, like the algebra of bounded operators on a complex Hilbert space, are not separable, as soon as the Hilbert space has infinite dimension (even if the Hilbert space is itself separable).

In fact, the referred metrics are first defined in the whole topological dual of any separable normed space.

Definition 4.80 Let X be any normed space. For any sequence $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ in X, we define the mapping $d_{\mathcal{S}} : X^{\mathrm{td}} \times X^{\mathrm{td}} \to \mathbb{R}_0^+$ by

$$d_{\mathcal{S}}(\varphi, \varphi') \doteq \max_{n\in\mathbb{N}} \frac{2^{-n}|(\varphi - \varphi')(x_n)|}{1 + |(\varphi - \varphi')(x_n)|}\,, \qquad \varphi, \varphi' \in X^{\mathrm{td}}\,.$$

Exercise 4.81 Prove that $d_{\mathcal{S}}$ is a "pseudometric," that is, for all $\varphi, \varphi', \varphi'' \in X^{\mathrm{td}}$, one has

$$d_{\mathcal{S}}(\varphi, \varphi') = d_{\mathcal{S}}(\varphi', \varphi) \qquad \text{and} \qquad d_{\mathcal{S}}(\varphi, \varphi'') \le d_{\mathcal{S}}(\varphi, \varphi') + d_{\mathcal{S}}(\varphi', \varphi'')\,.$$

Note that, for a general sequence $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ in X, $d_{\mathcal{S}}$ is only a pseudometric but not a metric, i.e., $d_{\mathcal{S}}(\varphi, \varphi') = 0$ does not necessarily imply that $\varphi = \varphi'$. In *separable* normed spaces, we do have a metric for any dense sequence.

Exercise 4.82 Let X be any separable normed space and $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ any *dense* sequence in X, that is, the subset $\{x_n : n \in \mathbb{N}\} \subseteq X$ is dense in X. (Such a sequence must exist, by definition of separable spaces.) Prove that $d_{\mathcal{S}}$ is a metric.

For metrics of this type, it turns out that the associated topology for norm-bounded subsets of the topological dual X^{td} is independent of the choice of the dense sequence.

Exercise 4.83 Let X be any separable normed space on $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Pick two dense sequences $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ and $\mathcal{S}' = (x_n')_{n\in\mathbb{N}}$ in X as well as a *norm-bounded* subset $\Omega \subseteq X^{\mathrm{td}}$. Prove that the metric spaces $(\Omega, d_{\mathcal{S}})$ and $(\Omega, d_{\mathcal{S}'})$ have exactly the same topology. Show, moreover, that a sequence $(\varphi_n)_{n\in\mathbb{N}}$ in Ω converges to $\varphi \in \Omega$ (with respect to $d_{\mathcal{S}}$ and $d_{\mathcal{S}'}$) iff, for all $x \in X$, one has

$$\lim_{n\to\infty} \varphi_n(x) = \varphi(x)\,.$$

In particular, for every $x \in X$, the mapping $\varphi \mapsto \varphi(x)$ from Ω to $\mathbb{K}$ is continuous.

The unique topology of norm-bounded subsets of the topological dual X^{td} that we refer to in the above exercise is called the "weak* topology." Functions that are continuous with respect to this topology, i.e., with respect to any of the metrics of the last exercise, are called weak*-continuous. In fact, this notion can be extended to any subset (not necessarily norm-bounded) of the topological dual of any normed space (not necessarily separable) in a way that the last part of the last exercise still hold true. More precisely, the weak* topology of X^{td} is, by definition, the smallest topology for which, for all $x \in X$, the mapping $\varphi \mapsto \varphi(x)$ from Ω to $\mathbb{K}$ is continuous. However, in this more general case, the weak* topology is not given by a metric anymore. For more details, see for instance [18, Part I, Section 3.14].

Recall that, for any C^*-algebra, the set of states $E(\mathcal{A}) \subseteq \mathcal{A}^{\mathrm{td}}$ is (by definition) norm-bounded. By Lemma 4.75 and Proposition 4.78, the Gelfand transform (Definition 4.79) is a bipositive injective linear transformation $\mathcal{A} \to C_{w^*}(E(\mathcal{A}); \mathbb{C})$, where $C_{w^*}(E(\mathcal{A});\mathbb{C})$ denotes the space of weak*-continuous functions $E(\mathcal{A}) \to \mathbb{C}$. Observe, moreover, that $C_{w^*}(E(\mathcal{A});\mathbb{C}) \subseteq C(E(\mathcal{A});\mathbb{C})$.

We prove in the following that, for any *unital* C^*-algebra $\mathcal{A}$, $E(\mathcal{A})$ is compact with respect to the weak* topology.

Proposition 4.84 *Let $\mathcal{A}$ be any separable unital C^*-algebra and $\mathcal{S} = (A_n)_{n\in\mathbb{N}}$ any dense sequence in $\mathcal{A}$. $(E(\mathcal{A}), d_{\mathcal{S}})$ is a compact metric space.*

Proof

1. Take any sequence $(\rho_m)_{m\in\mathbb{N}}$ in $E(\mathcal{A})$. As $E(\mathcal{A})$ is norm-bounded, for all $n \in \mathbb{N}$, the numerical sequence $(\rho_m(A_n))_{m\in\mathbb{N}}$ has a converging subsequence. By standard arguments involving so-called "diagonal subsequences," there is a subsequence $(\rho_{m_k})_{k\in\mathbb{N}}$ of $(\rho_m)_{m\in\mathbb{N}}$ such that, for all $n \in \mathbb{N}$, $(\rho_{m_k}(A_n))_{k\in\mathbb{N}}$ converges in $\mathbb{C}$.
2. As states are linear functionals, it follows that, for all $A \in \operatorname{span}(\{A_n : n \in \mathbb{N}\} \cup \{\mathbb{1}\})$ (i.e., A is a linear combination of elements of $\{A_n : n \in \mathbb{N}\} \cup \{\mathbb{1}\}$), the sequence $(\rho_{m_k}(A))_{k\in\mathbb{N}}$ converges in $\mathbb{C}$. Note that we used here that $(\rho_{m_k}(\mathbb{1}))_{k\in\mathbb{N}}$ trivially converges, being a constant (equal 1) sequence, by the normalization of states.
3. Define the mapping $\tilde{\rho} : \operatorname{span}(\{A_n : n \in \mathbb{N}\} \cup \{\mathbb{1}\}) \to \mathbb{C}$ by

$$\tilde{\rho}(A) \doteq \lim_{k\to\infty} \rho_{m_k}(A) , \qquad A \in \operatorname{span}(\{A_n : n \in \mathbb{N}\} \cup \{\mathbb{1}\}) .$$

   By linearity of states, $\tilde{\rho}$ is linear and $\tilde{\rho}(\mathbb{1}) = 1$. Moreover, the operator norm of $\tilde{\rho}$ is at most 1, because ρ_{m_k}, $k \in \mathbb{N}$, are all norm-one linear functionals. As $\tilde{\rho}(\mathbb{1}) = 1$, one thus has $\|\tilde{\rho}\|_{\mathrm{op}} = 1$.
4. By the Hahn-Banach extension theorem (Theorem 7.40), $\tilde{\rho}$ extends to a continuous linear functional $\rho \in \mathcal{A}^{\mathrm{td}}$ with $\|\rho\|_{\mathrm{op}} = 1 = \rho(\mathbb{1})$. In particular, $\rho \in E(\mathcal{A})$. By construction of this state, one finally has

$$\lim_{k\to\infty} d_{\mathcal{S}}(\rho_{m_k}, \rho) = \lim_{k\to\infty} \max_{n\in\mathbb{N}} \frac{2^{-n}|(\rho_{m_k} - \rho)(A_n)|}{1 + |(\rho_{m_k} - \rho)(A_n)|} = 0 .$$

   Note that the term 2^{-n} is important to obtain the last limit by reducing the above maximum to a finite set of natural numbers n, at the cost of some arbitrary $0 < \varepsilon \ll 1$.

□

In fact, if the C^*-algebra $\mathcal{A}$ is non-unital, then there is always a sequence of states $\rho_n \in E(\mathcal{A})$, $n \in \mathbb{N}$, such that, for any (dense) sequence $\mathcal{S}$ in $\mathcal{A}$, one has that

$$\lim_{n\to\infty} d_{\mathcal{S}}(\rho_n, 0) = 0 .$$

See Corollary 4.91. Clearly, this property prevents the set of states of any non-unital C^*-algebra from being weak*-compact. In other words, $E(\mathcal{A})$ is weak*-compact iff $\mathcal{A}$ is unital.

As explained after Definition 4.79, any C^*-algebra $\mathcal{A}$ is equivalent, as a $*$-ordered vector space, to the subspace of $C(E(\mathcal{A}); \mathbb{C}) \supseteq C_{w^*}(E(\mathcal{A}); \mathbb{C})$ formed by the Gelfand transforms of all elements of $\mathcal{A}$. We thus infer from Proposition 4.84 the following corollary.

Corollary 4.85 *Every* separable *unital C^*-algebra is equivalent, as a $*$-preordered vector space, to a subspace of the space of continuous functions on a compact metric space.*

Exercise 4.86 Let H be any separable Hilbert space. Show that $\mathcal{B}(H)$ is separable iff the dimension of H is finite.

Hint: To prove that $\mathcal{B}(H)$ is not separable when H is infinite dimensional, remark that, in this case, the family of all orthogonal projectors on H is not countable. To show this, construct an injective mapping from $2^{\mathbb{N}}$ to $\mathcal{P}(H)$, where $2^{\mathbb{N}}$ stands for the set of all subsets of $\mathbb{N}$.

This last exercise shows that the separable case is particular, in a sense, within the theory of C^*-algebras. It is nevertheless an important case, for it encompasses, among many other important cases, quantum spin and lattice-fermion systems in quantum statistical mechanics.

4.6 C^*-Algebra Homomorphisms and Representations

This section is devoted to the study of important general properties of $*$-homomorphisms between C^*-algebras, as well as representations of such algebras on Hilbert spaces. Note that general $*$-homomorphisms are defined in Definition 7.60 and that they are said to be unital when they preserve units.

We start by proving that the image of a $*$-homomorphism between C^*-algebras, whose domain is a unital C^*-algebra, is again a unital C^*-algebra.

Theorem 4.87 *Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two C^*-algebras with $\mathcal{A}_1$ being unital and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ a $*$-homomorphism. Then, $\Theta(\mathcal{A}_1) \subseteq \mathcal{A}_2$ is a C^*-subalgebra of $\mathcal{A}_2$ and $\Theta(\mathfrak{1})$ is a unit for this algebra.*

Proof Note that the norm closure $\overline{\Theta(\mathcal{A}_1)}$ of (the $*$-subalgebra) $\Theta(\mathcal{A}_1)$ in $\mathcal{A}_2$ is a C^*-subalgebra of $\mathcal{A}_2$, $\Theta(\mathfrak{1})$ being clearly a unit of the C^*-algebra $\overline{\Theta(\mathcal{A}_1)}$. Thus, we may assume without loss of generality that the $*$-homomorphism Θ is unital, as a mapping from $\mathcal{A}_1$ to $\overline{\Theta(\mathcal{A}_1)}$. If Θ is injective, then the theorem directly follows from Corollary 4.110. If Θ is not injective, note that $\ker(\Theta) \subseteq \mathcal{A}_1$, the kernel of Θ, is a closed ($*$-)ideal of $\mathcal{A}_1$, Θ being continuous, by Lemma 4.96 (ii). Thus, from Proposition 4.55, the quotient $\mathcal{A}_1/\ker(\Theta)$ is a (unital) C^*-algebra. Let the mapping $\bar{\Theta} : \mathcal{A}_1/\ker(\Theta) \to \mathcal{A}_2$ be defined as in Exercise 7.8. From Exercise 7.73, $\bar{\Theta}$ is an injective $*$-homomorphism, and the general case then follows from the injective one, because $\Theta(\mathcal{A}_1) = \bar{\Theta}(\mathcal{A}_1)$. □

We discuss now the notion of representation of $*$-algebras as spaces of bounded operators on complex Hilbert spaces.

Definition 4.88 (Representations of $*$-Algebras on Complex Hilbert Spaces) Let $\mathcal{A}$ be any (not necessarily unital) $*$-algebra:

(i) Pairs (H, π), where H is a complex Hilbert space and $\pi : \mathcal{A} \to \mathcal{B}(H)$ a $*$-homomorphism, are called "$*$-algebra representation" or simply "representation" of $\mathcal{A}$.
(ii) The representation (H, π) of $\mathcal{A}$ is said to be "faithful" whenever π is one-to-one.
(iii) Let (H, π) be a representation of $\mathcal{A}$. The vector $\Omega \in H$ is called "cyclic" for this representation if

$$\overline{\pi(\mathcal{A})\Omega} \doteq \overline{\{\pi(A)\Omega : A \in \mathcal{A}\}} = H \ .$$

(iv) Triples (H, π, Ω), where $\Omega \in H$, $\|\Omega\| = 1$, is cyclic for the representation (H, π), are called "cyclic representations" of $\mathcal{A}$.
(v) The representation (H, π) of $\mathcal{A}$ is called "irreducible" if $G = \{0\}$ and $G = H$ are the unique *closed* subspaces G of H such that

$$\pi(\mathcal{A})G \doteq \{\pi(A)x : x \in G,\ A \in \mathcal{A}\} \subseteq G \ .$$

In other words, $\{0\}$ and H are the unique closed "invariant spaces" of the representation.
(vi) The representation (H, π) of $\mathcal{A}$ is said to be "nondegenerate" if

$$H_0 \doteq \bigcap_{A \in \mathcal{A}} \ker(\pi(A)) = \{0\} \ .$$

(vii) Two representations (H_1, π_1) and (H_2, π_2) of $\mathcal{A}$ are equivalent if there is a unitary linear transformation $U : H_1 \to H_2$ such that, for all $A \in \mathcal{A}$, $\pi_1(A) = U^{-1} \circ \pi_2(A) \circ U$.
(viii) Two cyclic representations (H_1, π_1, Ω_1) and (H_2, π_2, Ω_2) of $\mathcal{A}$ are equivalent if there is a unitary linear transformation $U : H_1 \to H_2$ such that $\Omega_2 = U(\Omega_1)$ and, for all $A \in \mathcal{A}$, $\pi_1(A) = U^{-1} \circ \pi_2(A) \circ U$.

One simple example of representation of any $*$-algebra $\mathcal{A}$ are the trivial ones (H, π), i.e., those for which $\pi(A) = 0$ for all $A \in \mathcal{A}$, given some complex Hilbert space H. One example of a faithful representation of the C^*-algebra $\mathcal{B}(H)$ of bounded operators on an arbitrary complex Hilbert space H is the so-called identity representation $(H, \mathrm{id}_{\mathcal{B}(H)})$. Note additionally that, for all $\Omega \in H$, $\|\Omega\| = 1$, $(H, \mathrm{id}_{\mathcal{B}(H)}, \Omega)$ is a cyclic representation of $\mathcal{B}(H)$.

Observe from Theorem 4.87 that the range $\pi(\mathcal{A}) \subseteq \mathcal{B}(H)$ of a representation (H, π) of a C^*-algebra $\mathcal{A}$ is a C^*-subalgebra of $\mathcal{B}(H)$. Thus, if $\mathcal{A}$ has a faithful representation, then $\mathcal{A}$ is equivalent, as a $*$-algebra, to a concrete C^*-algebra, i.e., a C^*-subalgebra of $\mathcal{B}(H)$. Additionally, if $\mathcal{A}$ is unital, then one may choose the C^*-subalgebra of $\mathcal{B}(H)$ as being unital (i.e., containing $\mathrm{id}_{\mathcal{B}(H)}$, the unit of $\mathcal{B}(H)$). See Exercise 4.94 in this context. The existence of faithful representations is guaranteed by the celebrated Gelfand-Naimark theorem [56].

Theorem 4.89 (Gelfand-Naimark) *Every C^*-algebra possesses a faithful representation. In particular, any unital C^*-algebra $\mathcal{A}$ is equivalent, as a $*$-algebra, to some unital C^*-subalgebra of $\mathcal{B}(H)$, for some complex Hilbert space H.*

In fact, in the case of simple C^*-algebras, any nontrivial representation is faithful, as proven in the next lemma. In particular, in this special case, the Gelfand-Naimark theorem is a direct consequence of Theorem 4.113 proven below. Note, in this context, that in quantum statistical physics, many important C^*-algebras, like the fermion algebras or the quantum spin algebras, are simple C^*-algebras. See Lemma 4.153 and Corollary 4.167, as well as Sect. 5.1. Recall that this property refers here to Definition 7.48 (iii) for normed algebras: A normed algebra $\mathcal{A}$ is "simple" if $\{0\}$ and $\mathcal{A}$ are the only *closed* ideals of $\mathcal{A}$.

Lemma 4.90 *Any nontrivial representation (H, π) of a simple C^*-algebra $\mathcal{A}$ is faithful.*

Proof Let (H, π) be a non-trivial representation (H, π) of a simple C^*-algebra $\mathcal{A}$. In particular, $\ker(\pi) \neq \mathcal{A}$. By Exercise 7.26, $\ker(\pi)$ is ideal. By Proposition 4.97, π is continuous. $\ker(\pi)$ is thus a closed ideal of $\mathcal{A}$. Hence, as $\mathcal{A}$ is a simple normed algebra, $\ker(\pi) = \{0\}$ and, consequently, π is faithful. □

The following corollary on the existence of faithful representations for any C^*-algebra yields, in the separable case, that the state space of a non-unital C^*-algebra is not weak*-compact. See remarks after Proposition 4.84.

Corollary 4.91 *Let $\mathcal{A}$ be an arbitrary C^*-algebra that is separable and non-unital. There is a sequence of states of $\mathcal{A}$ tending to zero in the weak* topology.*

Proof Fix all parameters of the corollary. Take any sequence $\mathcal{S} = (A_n)_{n\in\mathbb{N}}$ that is dense in $\mathcal{A}$, this C^*-algebra being separable, and let the metric $d_{\mathcal{S}}$ be defined on the state space $E(\mathcal{A}) \subseteq \mathcal{A}^{\mathrm{td}}$ as in Definition 4.80. For all $n \in \mathbb{N}$, let

$$B_n \doteq \sum_{k=1}^{n} (\mathrm{Re}\{A_k\}^+ + \mathrm{Re}\{A_k\}^- + \mathrm{Im}\{A_k\}^+ + \mathrm{Im}\{A_k\}^-) \in \mathcal{A}^+ ,$$

where, for any $k \in \{1, \ldots, n\}$,

$$\mathrm{Re}\{A_k\} \doteq \frac{1}{2}\left(A_k + A_k^*\right) = \mathrm{Re}\{A_k\}^+ - \mathrm{Re}\{A_k\}^- ,$$

$$\mathrm{Im}\{A_k\} \doteq \frac{1}{2i}\left(A_k - A_k^*\right) = \mathrm{Im}\{A_k\}^+ - \mathrm{Im}\{A_k\}^- ,$$

are the orthogonal decompositions of $\mathrm{Re}\{A_k\}$ and $\mathrm{Im}\{A_k\}$, respectively, as given by Proposition 4.102 (ii). Take any faithful representation (H, π) of $\mathcal{A}$, which exists by the Gelfand-Naimark theorem (Theorem 4.89). By Exercise 4.20, for all $n \in \mathbb{N}$, $0 \in \sigma(B_n)$. Thus, by Lemma 4.96 (iii), $0 \in \sigma(\pi(B_n))$. By Proposition 7.242, for all $n \in \mathbb{N}$, there is a norm-one vector $x \in H$ such that $|\langle x, \pi(B_n)x\rangle| \leq n^{-1}$. By a simple adaptation of the proof of Lemma 4.111, one shows that, for all $n \in \mathbb{N}$, $\rho_n \doteq \langle x, \pi(\cdot)x\rangle$ is a state on $\mathcal{A}$. By construction, one has $|\rho_n(A_k)| \leq n^{-1}$ for all $n \in \mathbb{N}$ and $k \in \{1, \ldots, n\}$. Thus,

$$d_{\mathcal{S}}(0, \rho_n) \leq n^{-1} + 2^{-n}, \qquad n \in \mathbb{N},$$

and the sequence $(\rho_n)_{n\in\mathbb{N}}$ of states of $\mathcal{A}$ converges to zero in the weak* topology, by the definitions and results of Sect. 4.5.1. □

Uniqueness of representations of C^*-algebras is clearly wrong. Indeed, for any representation (H, π), we can construct another one by doubling the Hilbert space H and the mapping π, via a direct sum $H_1 \oplus H_2$ with H_1, H_2 being two copies of H. Therefore, one introduces a notion of "minimal" representations of C^*-algebras, which refers to the *irreducibility* of Definition 4.88 (v). This property has the following equivalent characterizations.

Theorem 4.92 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. Take a representation (H, π) of $\mathcal{A}$. The following three properties are equivalent:*

(i) *(H, π) is an irreducible representation of $\mathcal{A}$.*
(ii) *Every vector $x \in H\backslash\{0\}$ is cyclic for the representation (H, π).*
(iii) *$\pi(\mathcal{A})' = \mathbb{C}\mathrm{id}_H$, where $\pi(\mathcal{A})' \subseteq \mathcal{B}(H)$ is the commutant of $\pi(\mathcal{A}) \subseteq \mathcal{B}(H)$, defined by*

$$\pi(\mathcal{A})' \doteq \{B \in \mathcal{B}(H) : [B, \pi(A)] \doteq B\pi(A) - \pi(A)B = 0 \text{ for all } A \in \mathcal{A}\}$$

(see Definition 7.253).

Proof See [51, Proposition 2.3.8]. □

Exercise 4.93 Prove that the assertions (i) and (ii) in the above theorem are equivalent.

Recall that the commutant $\pi(\mathcal{A})' \subseteq \mathcal{B}(H)$ is a von Neumann algebra. In particular it is a unital C^*-subalgebra of $\mathcal{B}(H)$. Note that the identity representation of $\mathcal{B}(H)$ has all the three properties stated in the above theorem. See, for instance, Lemma 3.35, which independently proves the third property.

It is easy to see that if the representation (H, π) is not irreducible, then the orthogonal complement $G^\perp$ of any closed invariant subspace G satisfying $\{0\} \subsetneq G \subsetneq H$ is also a closed invariant space, with $\{0\} \subsetneq G^\perp \subsetneq H$. In particular, (H, π) splits in this case into two subrepresentations, $(G, \pi(\cdot)|_G)$ and $(G^\perp, \pi(\cdot)|_{G^\perp})$. So, irreducible representation can be see as minimal representations. In fact, infinite dimensional C^*-algebras possess, in general, infinitely many inequivalent irreducible representations. Restrictions of a representation (H, π) to closed invariant subspaces, like $(G, \pi(\cdot)|_G)$ and $(G^\perp, \pi(\cdot)|_{G^\perp})$, are called "subrepresentations" of (H, π). Observe that if the representation (H, π) of a $*$-algebra $\mathcal{A}$ is not trivial but degenerate, i.e.,

$$H_0 \doteq \bigcap_{A \in \mathcal{A}} \ker(\pi(A)) \neq \{0\}, H \ ,$$

then it is reducible, because H_0 is a nontrivial closed invariant subspace for the representation, in this case.

Remark, however, that irreducible representations are not the "natural building blocks" of an arbitrary representation of a C^*-algebra. Instead, one considers so-called factor representations (that are not necessarily irreducible, whereas any irreducible representation is a factor) and "von Neumann's factor decompositions of representations." A factor representation (H, π) of a C^*-algebra $\mathcal{A}$ is, by definition, a representation for which the bicommutant

$$\pi(\mathcal{A})'' \doteq \{A \in \mathcal{B}(H) : [A, B] = 0 \text{ for all } B \in \pi(\mathcal{A})'\} \subseteq \mathcal{B}(H)$$

of its range $\pi(\mathcal{A})$ is a von Neumann algebra whose center is trivial, i.e.,

$$\mathcal{Z}(\pi(\mathcal{A})'') \doteq \pi(\mathcal{A})'' \cap \pi(\mathcal{A})' = \mathbb{C}\mathrm{id}_H \subseteq \mathcal{B}(H) \ .$$

We will not discuss further on this decomposition theory and only mention it here for completeness. To see the key role of factors in the theory of operator algebras, we recommend for instance [60–62] or [63] and [49, Section 6] for a more concise presentation.

Exercise 4.94 Let $\mathcal{A}$ be a $*$-algebra and (H, π) a representation of $\mathcal{A}$. Show that:

(i) (H, π) is cyclic, that is, (H, π, x) is cyclic for all $x \in H$, only if it is nondegenerate. In particular, (H, π) is irreducible only if it is nondegenerate
(ii) If $\mathcal{A}$ is unital, then (H, π) is nondegenerate iff π is a unital $*$-homomorphism.
(iii) If (H, π) is a nontrivial representation, then it is the direct sum of a nondegenerate subrepresentation and a trivial one.

Hint: For the first and second parts, observe that the subspace

$$\{\pi(A)x \ : \ A \in \mathcal{A}, \ x \in H\} \subseteq H$$

is dense in H iff (H, π) is nondegenerate. For the third, note that the closure of the above subspace is an invariant subspace for any representation (H, π).

Because of the third part of the exercise, we assume, by default, that representations are neither trivial nor degenerate. If $\mathcal{A}$ is a non-unital C^*-algebra, then there is a version of the second part of the last exercise, referring to approximate units (Definition 4.49):

Lemma 4.95 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and (H, π) a representation of $\mathcal{A}$. (H, π) is nondegenerate iff, for any approximate unit $(E_i)_{i\in I}$ of $\mathcal{A}$, the net $(\pi(E_i))_{i\in I}$ in $\mathcal{B}(H)$ converges in the strong operator topology to the identity operator, i.e., for all $x \in H$, the net $(\pi(E_i)x)_{i\in I}$ converges in H to x.*

Proof Recall that any non-unital C^*-algebra $\mathcal{A}$ possesses an approximate unit, thanks to Corollary 4.51. Let (H, π) be a representation of $\mathcal{A}$ and $(E_i)_{i\in I} \subseteq \mathcal{A}$ an approximate unit. If for all $x \in H$, the net $(\pi(E_i)x)_{i\in I}$ converges in H to x, then the subspace

$$\mathcal{Y} = \{\pi(A)x \ : \ A \in \mathcal{A}, \ x \in H\} \subseteq H$$

is dense in H. By Definition 4.88 (vi), the representation (H, π) is thus nondegenerate, in this case. Conversely, if (H, π) is nondegenerate, then

$$H = H_0^{\perp} = \left(\bigcap_{A\in \mathrm{Re}\{\mathcal{A}\}} \pi(A)(H)^{\perp} \right)^{\perp} = \overline{\mathrm{span}\{\pi(A)x \ : \ A \in \mathrm{Re}\{\mathcal{A}\}, \ x \in H\}} = \overline{\mathcal{Y}},$$

using $X^{\perp\perp} = \bar{X}$, $X^{\perp} \cap Y^{\perp} = (X \cup Y)^{\perp}$, Lemma 7.237 (iii) and the equality

$$H_0 = \bigcap_{A\in \mathrm{Re}\{\mathcal{A}\}} \ker(\pi(A)) \ .$$

In other words, the subspace $\mathcal{Y}$ is dense in H. By Proposition 4.97, representations of C^*-algebras are contractions. Thus, for any approximate unit $(E_i)_{i\in I}$ of $\mathcal{A}$ and any $x \in H$,

$$\lim_{i\in I} \pi(E_i)(\pi(A)x) = \pi \left(\lim_{i\in I} E_i A \right) x = \pi(A)x \ .$$

Moreover, $\|\pi(E_i)\|_{\mathrm{op}} \leq 1$ for every $i \in I$ and, hence, by density of the subspace $\mathcal{Y}$, for all $x \in H$,

$$\lim_{i\in I} \pi(E_i)x = x \ .$$

□

In the proof of the last lemma, we use that representations of C^*-algebras are contractions. This is in fact a general property of $*$-homomorphisms. In the two next statements, we prove this fact together with other key properties of $*$-homomorphisms. We start with the unital case.

Lemma 4.96 *Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two unital C^*-algebras and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ a unital $*$-homomorphism:*

(i) $\Theta \in \mathcal{L}^+(\mathcal{A}_1; \mathcal{A}_2)$, *i.e., for all* $A \in \mathcal{A}_1^+$, $\Theta(A) \in \mathcal{A}_2^+$.
(ii) *For all* $A_1 \in \mathcal{A}_1$, $\sigma(\Theta(A_1)) \subseteq \sigma(A_1)$ *and* $\|\Theta(A_1)\| \leq \|A_1\|$. *In particular,* $\Theta \in \mathcal{B}(\mathcal{A}_1; \mathcal{A}_2)$.
(iii) *If* Θ *is faithful (i.e., one-to-one) then, for all* $A_1 \in \mathcal{A}_1$, $\sigma(\Theta(A_1)) = \sigma(A_1)$ *and* $\|\Theta(A_1)\| = \|A_1\|$.

Proof Exercise. Prove (i) and (ii) directly, i.e., without using Theorem 4.87. To prove (iii) use this theorem. □

The second part of the lemma can be extended to the non-unital case and sharpened as follows.

Proposition 4.97 *Let* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *be two arbitrary (not necessarily unital)* C^*-*algebras and* $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ *a* $*$-*homomorphism. For all* $A_1 \in \mathcal{A}_1$, *one has that*

$$\|\Theta(A_1)\| = \inf\{\|A_1 + B_1\| \ : \ B_1 \in \ker(\Theta)\} .$$

Proof Let $\widetilde{\mathcal{A}}_1$ and $\widetilde{\mathcal{A}}_2$ be respectively the unitizations of $\mathcal{A}_1$ and $\mathcal{A}_2$. By Exercise 7.64, there is a unique unital $*$-homomorphism $\tilde{\Theta} : \widetilde{\mathcal{A}}_1 \to \widetilde{\mathcal{A}}_2$ extending Θ. Note that Exercise 7.64 requires that $\mathcal{A}_2$ is unital, but it suffices to see Θ as a $*$-homomorphism from $\mathcal{A}_1$ to $\widetilde{\mathcal{A}}_2 \supseteq \mathcal{A}_2$. Thus, from Lemma 4.96 (ii), it follows that Θ is a contraction. Clearly, Θ is faithful only if $\tilde{\Theta}$ faithful. In particular, the assertion follows from Lemma 4.96 (iii) when Θ is faithful. If Θ is not faithful, consider the factorization $\Theta = \bar{\Theta} \circ \mathfrak{q}$ of Definition 7.7 with $\mathcal{I} = \ker(\Theta)$. By Exercise 7.73, $\bar{\Theta} : \mathcal{A}_1 \to \mathcal{A}_1 / \ker(\Theta)$ is a faithful $*$-homomorphism and, from Proposition 4.55, $\mathcal{A}_1 / \ker(\Theta)$ is a C^*-algebra, for the ideal $\ker(\Theta) \subseteq \mathcal{A}_1$ is closed, as Θ is continuous. Thus, the assertion for Θ follows from the first part of the proof applied to the faithful $*$-homomorphism $\bar{\Theta}$. □

Given two unital C^*-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, note from Lemma 4.96 that a $*$-isomorphism $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ is a unital (i.e., it maps the unit of $\mathcal{A}_1$ to the unit of $\mathcal{A}_2$) bipositive linear transformation. In other words, Θ is an equivalence of $*$-ordered vector spaces that preserves the units of the algebras. Interestingly, the converse of this fact holds also true.

Theorem 4.98 (Kadison) *Let* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *be two unital* C^*-*algebras and* $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ *a linear transformation.* Θ *is a* $*$-*isomorphism of algebras iff it is bijective and bipositive and maps the unit of* $\mathcal{A}_1$ *to the unit of* $\mathcal{A}_2$.

By the last theorem, the order structure of any unital C^*-algebra uniquely determines its $*$-algebra structure, up to an $*$-automorphism. Note, for instance, that spectra of C^*-algebra elements are invariant under $*$-isomorphisms. In particular, the spectrum of any element of a unital C^*-algebra is *uniquely determined by the order structure* of the C^*-algebra.

4.6.1 The Continuous Functional Calculus

In the following we discuss the so-called continuous functional calculus of an arbitrary C^*-algebra, which refers to a very important class of faithful (i.e., injective) $*$-homomorphisms of C^*-algebras. To begin with, for any self-conjugate element $A \in \mathrm{Re}\{\mathcal{A}\}$ of a C^*-algebra $\mathcal{A}$, we use the shorter notation $\tilde{C}_A$ for the (unital) C^*-algebra $C(\sigma(A); \mathbb{C})$ of continuous functions on the spectrum $\sigma(A)$ of A. Recall that spectra of self-conjugate C^*-algebra elements are compact subsets of $\mathbb{R}$. See Definitions 4.18 and 4.35, as well as Corollary 4.25. We further define $C_A \subseteq \tilde{C}_A$ as being the (generally non-unital) C^*-subalgebra of continuous functions vanishing at 0:

$$C_A \doteq \{f \in \tilde{C}_A \ : \ f(0) = 0\} .$$

In particular, $C_A = \tilde{C}_A$ when $0 \notin \sigma(A)$. Note that C_A can be canonically identified with the algebra $C_0(\sigma(A)\backslash\{0\}; \mathbb{C})$ of continuous function on $\sigma(A)\backslash\{0\}$ decaying at infinity, in the sense of Definition 7.166 (see also Exercise 7.189).

For any fixed $A \in \mathrm{Re}\{\mathcal{A}\}$, let the $*$-subalgebras $\widetilde{\mathrm{Pol}}_A \subseteq \tilde{C}_A$ and $\mathrm{Pol}_A \subseteq C_A$ be defined by

$$\widetilde{\mathrm{Pol}}_A \doteq \{f \in \tilde{C}_A \ : \ \text{there is a polynomial } \mathcal{P} \text{ such that } f = \mathcal{P} \text{ on } \sigma(A) \subseteq \mathbb{R}\} ,$$
$$\mathrm{Pol}_A \doteq C_A \cap \widetilde{\mathrm{Pol}}_A .$$

By the Stone-Weierstrass theorem (Theorem 7.191), observe that the subspace $\widetilde{\mathrm{Pol}}_A \subseteq \tilde{C}_A$ is dense in $(\tilde{C}_A, \|\cdot\|_\infty)$, where we recall that

$$\|f\|_\infty \doteq \sup\{|f(x)| : x \in \sigma(A)\}$$

for any continuous function $f \in \tilde{C}_A$ on the compact set $\sigma(A) \subseteq \mathbb{R}$. Similarly, Pol_A is a dense subspace of $(C_A, \|\cdot\|_\infty)$.

Assume that the C^*-algebra $\mathcal{A}$ is unital. Then, given $A \in \mathrm{Re}\{\mathcal{A}\}$ and an arbitrary $f \in \widetilde{\mathrm{Pol}}_A$, there are, in general, more than one polynomial $\mathcal{P}$ such that $f = \mathcal{P}$ on $\sigma(A) \subseteq \mathbb{R}$. (This is the case when the spectrum only contains a finite number of points $\sigma(A)$.) For a fixed $A \in \mathrm{Re}\{\mathcal{A}\}$ and $f \in \widetilde{\mathrm{Pol}}_A$, let $\mathcal{P}_1$ and $\mathcal{P}_2$ be any two polynomials such that $f = \mathcal{P}_1 = \mathcal{P}_2$ on $\sigma(A)$. In particular, the polynomial $\mathcal{P}_1 - \mathcal{P}_2$ vanishes on $\sigma(A)$. Note furthermore that, for any polynomial $\mathcal{P}$ and any $A \in \mathrm{Re}\{\mathcal{A}\}$, $\mathcal{P}(A) \in \mathcal{A}$ is a normal element of the unital C^*-algebra $\mathcal{A}$. See Definition 4.22. Hence, by Proposition 4.31,

$$\|\mathcal{P}_1(A) - \mathcal{P}_2(A)\| = \|(\mathcal{P}_1 - \mathcal{P}_2)(A)\| = \max\{|\sigma| \ : \ \sigma \in \sigma((\mathcal{P}_1 - \mathcal{P}_2)(A))\} .$$

We infer from Proposition 4.23 that

$$\sigma((\mathcal{P}_1 - \mathcal{P}_2)(A)) = (\mathcal{P}_1 - \mathcal{P}_2)(\sigma(A)) = \{0\}$$

and one thus has $\mathcal{P}_1(A) = \mathcal{P}_2(A)$. From this observation, for any $A \in \mathrm{Re}\{\mathcal{A}\}$, the condition $\tilde{\Phi}_A(f) \doteq \mathcal{P}(A)$, where $\mathcal{P}$ is any polynomial such that $f = \mathcal{P}$ on $\sigma(A)$, well defines a mapping $\tilde{\Phi}_A : \widetilde{\mathrm{Pol}}_A \to \mathcal{A}$, which has the following properties.

Lemma 4.99 *Let $\mathcal{A}$ be a unital C^*-algebra. For all $A \in \mathrm{Re}\{\mathcal{A}\}$, $\tilde{\Phi}_A : \widetilde{\mathrm{Pol}}_A \to \mathcal{A}$ is a unital $*$-homomorphism which is isometric, i.e., $\|\tilde{\Phi}_A(f)\| = \|f\|_\infty$ for all $f \in \widetilde{\mathrm{Pol}}_A$.*

Proof Exercise. □

By the above lemma, $\tilde{\Phi}_A \in \mathcal{B}(\widetilde{\mathrm{Pol}}_A, \mathcal{A})$ with $\|\tilde{\Phi}_A\|_{\mathrm{op}} = 1$. Hence, as $\mathcal{A}$ is a Banach space, it has a unique extension $\Phi_A \in \mathcal{B}(\tilde{C}_A, \mathcal{A})$. This extension is an injective unital $*$-homomorphism.

Proposition 4.100 (Unital Continuous Functional Calculus) *Let $\mathcal{A}$ be any unital C^*-algebra and take any $A \in \mathrm{Re}\{\mathcal{A}\}$. There is a unique unital $*$-homomorphism $\tilde{C}_A \to \mathcal{A}$, also denoted by $\tilde{\Phi}_A$, such that $\tilde{\Phi}_A(\mathrm{id}_{\sigma(A)}) = A$, where $\mathrm{id}_{\sigma(A)} \in \tilde{C}_A$ is the identity function, i.e., $\mathrm{id}_{\sigma(A)}(x) \doteq x$, $x \in \sigma(A)$. Additionally, $\tilde{\Phi}_A$ is norm-preserving, i.e., $\|\tilde{\Phi}_A(f)\|_{\mathcal{A}} = \|f\|_\infty$ for all $f \in \tilde{C}_A$, as well as positivity-preserving (i.e., order-preserving) and order-reflecting.*[15]

Proof

1. Fix $A \in \mathrm{Re}\{\mathcal{A}\}$. As already remarked above, by the Stone-Weierstrass theorem (Theorem 7.191), $\widetilde{\mathrm{Pol}}_A$ is a dense subspace of the normed space $(\tilde{C}_A, \|\cdot\|_\infty)$. For unital $*$-homomorphisms of C^*-algebras are continuous linear transformations (Lemma 4.96 (ii)), $\tilde{\Phi}_A$ is thus completely determined by its restriction on $\widetilde{\mathrm{Pol}}_A \subseteq \tilde{C}_A$. Since the condition $\tilde{\Phi}_A(\mathrm{id}_{\sigma(A)}) = A$ and the property of $\tilde{\Phi}_A$ being a unital $*$-homomorphism uniquely determine $\tilde{\Phi}_A(f)$ for all $f \in \widetilde{\mathrm{Pol}}_A$, $\tilde{\Phi}_A$ is unique, if it exists.
2. $\tilde{\Phi}_A \in \mathcal{L}(\widetilde{\mathrm{Pol}}_A; \mathcal{A})$ is bounded, thanks to Lemma 4.99. In particular, it has a unique extension $\tilde{\Phi}_A \in \mathcal{B}(\tilde{C}_A; \mathcal{A})$ (by Lemma 7.129). We now prove that

$$\tilde{\Phi}_A(ff') = \tilde{\Phi}_A(f)\tilde{\Phi}_A(f')$$

for all $f, f' \in \tilde{C}_A$. For any fixed $f, f' \in \tilde{C}_A$, let $(f_n)_{n\in\mathbb{N}}, (f'_n)_{n\in\mathbb{N}} \in \widetilde{\mathrm{Pol}}_A$ be two sequences such that $\lim_{n\to\infty} f_n = f$ and $\lim_{n\to\infty} f_n = f'$ in $(\tilde{C}_A, \|\cdot\|_\infty)$. These sequences exist, because $\widetilde{\mathrm{Pol}}_A$ is dense in $\tilde{C}_A$. For all $n \in \mathbb{N}$,

$$\begin{aligned}
&\tilde{\Phi}_A(ff') - \tilde{\Phi}_A(f)\tilde{\Phi}_A(f') \\
&\quad = \tilde{\Phi}_A(f_n f'_n) - \tilde{\Phi}_A(f_n)\tilde{\Phi}_A(f'_n) \\
&\qquad +(\tilde{\Phi}_A(ff'_n) - \tilde{\Phi}_A(f_n f'_n)) + (\tilde{\Phi}_A(ff') - \tilde{\Phi}_A(ff'_n)) \\
&\qquad -(\tilde{\Phi}_A(f)\tilde{\Phi}_A(f'_n) - \tilde{\Phi}_A(f_n)\tilde{\Phi}_A(f'_n))
\end{aligned}$$

[15] That is, for all $f, g \in \tilde{C}_A$, $\tilde{\Phi}_A(f) \geq \tilde{\Phi}_A(g)$ only if $f \geq g$.

$$
\begin{aligned}
&-(\tilde{\Phi}_A(f)\tilde{\Phi}_A(f') - \tilde{\Phi}_A(f)\tilde{\Phi}_A(f'_n)) \\
&= \tilde{\Phi}_A(f_n f'_n) - \tilde{\Phi}_A(f_n)\tilde{\Phi}_A(f'_n) \\
&\quad + \tilde{\Phi}_A((f - f_n)f'_n) + \tilde{\Phi}_A(f(f' - f'_n)) \\
&\quad - \tilde{\Phi}_A(f - f_n)\tilde{\Phi}_A(f'_n) - \tilde{\Phi}_A(f)\tilde{\Phi}_A(f' - f'_n) \\
&= \tilde{\Phi}_A((f - f_n)f'_n) + \tilde{\Phi}_A(f(f - f'_n)) \\
&\quad - \tilde{\Phi}_A(f - f_n)\tilde{\Phi}_A(f'_n) - \tilde{\Phi}_A(f)\tilde{\Phi}_A(f' - f'_n) .
\end{aligned}
$$

We used above that

$$\tilde{\Phi}_A(f_n f_n) - \tilde{\Phi}_A(f_n)\tilde{\Phi}_A(f'_n) = 0 ,$$

thanks to Lemma 4.99. Because $\tilde{\Phi}_A$ is a bounded linear operator and that $\tilde{C}_A$ and $\mathcal{A}$ are normed algebras,

$$\lim_{n\to\infty} \tilde{\Phi}_A((f - f_n)f'_n) = 0 , \qquad \lim_{n\to\infty} \tilde{\Phi}_A(f - f_n)\tilde{\Phi}_A(f'_n) = 0 ,$$

$$\lim_{n\to\infty} \tilde{\Phi}_A(f(f' - f'_n)) = 0 , \qquad \lim_{n\to\infty} \tilde{\Phi}_A(f)\tilde{\Phi}_A(f' - f'_n) = 0 .$$

3. With a similar argument one proves that, for all $f \in \tilde{C}_A$, $\tilde{\Phi}_A(f^*) = \tilde{\Phi}_A(f)^*$.
4. The fact that $\tilde{\Phi}_A$ is positivity-preserving and positivity-reflecting directly follows from Lemma 4.96 (iii). This concludes the proof.

□

For a fixed unital C^*-algebra $\mathcal{A}$, the family of unital $*$-homomorphisms $\tilde{\Phi}_A : \tilde{C}_A \to \mathcal{A}$, $A \in \mathrm{Re}\{\mathcal{A}\}$, is called the "continuous functional calculus" of $\mathcal{A}$. For all $A \in \mathrm{Re}\{\mathcal{A}\}$ and $f \in \tilde{C}_A$, we use the usual notation:

$$f(A) \doteq \tilde{\Phi}_A(f) .$$

The next corollary is a version of the continuous functional calculus for non-unital C^*-algebras:

Corollary 4.101 (Non-unital Continuous Functional Calculus) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and take any $A \in \mathrm{Re}\{\mathcal{A}\}$. There is a unique $*$-homomorphism $C_A \to \mathcal{A}$, denoted by Φ_A, such that $\Phi_A(\mathrm{id}_{\sigma(A)}) = A$, where $\mathrm{id}_{\sigma(A)} \in C_A$ is the identity function. The algebra homomorphism Φ_A is norm-preserving, as well as positivity-preserving (i.e., order-preserving) and order-reflecting.*

Proof Again, for any $A \in \mathrm{Re}\{\mathcal{A}\}$, the uniqueness of Φ_A follows from the fact that it is uniquely determined in the dense subspace $\mathrm{Pol}_A \subseteq C_A$, by the condition $\Phi_A(\mathrm{id}_{\sigma(A)}) = A$ and the fact that Φ_A is a $*$-homomorphism. If $\mathcal{A}$ is a unital C^*-algebra, then the existence of Φ_A directly follows from Proposition 4.100, by taking Φ_A as the restriction of $\tilde{\Phi}_A$ to $C_A \subseteq \tilde{C}_A$. If $\mathcal{A}$ is a non-unital, we then take $\Phi_A : C_A \to \tilde{\mathcal{A}}$ as being the restriction of $\tilde{\Phi}_A : \tilde{C}_A \to \tilde{\mathcal{A}}$ to $C_A \subseteq \tilde{C}_A$ and show that $\Phi_A(C_A) \subseteq \mathcal{A} \subseteq \tilde{\mathcal{A}}$. Recall that $\tilde{\mathcal{A}}$ denotes the unitization of $\mathcal{A}$. Clearly, $\Phi_A(\mathrm{Pol}_A) \subseteq \mathcal{A} \subseteq \tilde{\mathcal{A}}$. As $\mathcal{A}$ is a closed subspace of $\tilde{\mathcal{A}}$, Pol_A is dense in C_A and Φ_A is continuous, the desired property of Φ_A then follows. □

Now, Corollary 4.101 can be used to prove the following assertion.

Proposition 4.102 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra:*

(i) *For every $A \in \mathcal{A}^+$, there is a positive element $\sqrt{A} \in \mathcal{A}^+$ such that $(\sqrt{A})^2 = A$.*
(ii) *For every $A \in \mathrm{Re}\{\mathcal{A}\}$, there are positive elements $A^+, A^- \in \mathcal{A}^+$ such that $A = A^+ - A^-$, $A^-A^+ = A^+A = 0$ and $\|A^+\|, \|A^-\| \leq \|A\|$.*

Proof

(i) Let $A \in \mathcal{A}^+$. Then, $\sigma(A) \subseteq \mathbb{R}_0^+$ and thus $f(\sigma) \doteq \sqrt{\sigma}$, $\sigma \in \sigma(A)$, defines a positive continuous function on $\sigma(A)$ that vanishes at 0, i.e., $f \in C_A^+$. From Corollary 4.101 one has that $f(A) \geq 0$ and $(f(A))^2 = A$.
(ii) Let now $A \in \mathrm{Re}\{\mathcal{A}\}$ be arbitrary (i.e., not necessarily positive). Similar to the first part, note that $f^+(x) \doteq (|x| + x)/2$, $f^-(x) \doteq (|x| - x)/2$, $x \in \sigma(A)$, define two elements $f^+, f^- \in C_A^+$ such that $f^+ - f^- = \mathrm{id}_{\sigma(A)}$, $f^+ \cdot f^- = 0$, and $\|f^+\|_\infty, \|f^-\|_\infty \leq \|\mathrm{id}_{\sigma(A)}\|_\infty$. From the properties of the non-unital continuous functional calculus (Corollary 4.101), $f^+(A), f^-(A) \geq 0$, $f^+(A) - f^-(A) = A$, $\|f^+(A)\|, \|f^-(A)\| \leq \|A\|$ and

$$f^+(A)f^-(A) = f^-(A)f^+(A) = 0\,.$$

□

The decomposition in Proposition 4.102 (ii) is an improvement of Lemma 4.38 and is called the "orthogonal decomposition" of the self-conjugate element A. In fact, it can be additionally proven that such a decomposition with $A = A^+ - A^-$, $A^-A^+ = A^+A = 0$ and $A^+, A^- \in \mathcal{A}^+$ is unique for any $A \in \mathrm{Re}\{\mathcal{A}\}$, where $\mathcal{A}$ is an arbitrary C^*-algebra. See [51, Proposition 2.2.11]. Similarly, the "positive square root" $\sqrt{A} \in \mathcal{A}^+$ of any positive element $A \in \mathcal{A}^+$ is also unique. See [51, Theorem 2.2.10].

The last proposition implies the following equivalent characterizations of positive elements of an arbitrary C^*-algebra.

Corollary 4.103 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. For every $A \in \mathcal{A}$, the following properties are equivalent:*

(i) *$A \in \mathcal{A}^+$.*
(ii) *$A = B^2$ for some $B \in \mathrm{Re}\{\mathcal{A}\}$.*
(iii) *$A = B^*B$ for some $B \in \mathcal{A}$.*

Proof

1. (i)⇒(ii) follows from Proposition 4.102 (i), while (ii)⇒(iii) is trivial.
2. In order to prove that (iii)⇒(i), let $B \in \mathcal{A}$ and consider the (orthogonal) decomposition:

$$B^*B = (B^*B)^+ - (B^*B)^- .$$

We will prove that $(B^*B)^- = 0$. On the one hand, observe from Proposition 4.102 (ii) that

$$\begin{aligned}(B(B^*B)^-)^*(B(B^*B)^-) &= (B^*B)^-(B^*B)(B^*B)^- \\ &= (B^*B)^-((B^*B)^+ - (B^*B)^-)(B^*B)^- \\ &= -((B^*B)^-)^3 \in -\mathcal{A}^+ .\end{aligned}$$

3. On the other hand,

$$\begin{aligned}(B(B^*B)^-)(B(B^*B)^-)^* &= -(B(B^*B)^-)^*(B(B^*B)^-) \\ &\quad +2(\mathrm{Re}\{B(B^*B)^-\})^2 + 2(\mathrm{Im}\{B(B^*B)^-\})^2 \\ &= ((B^*B)^-)^3 + 2(\mathrm{Re}\{B(B^*B)^-\})^2 \\ &\quad +2(\mathrm{Im}\{B(B^*B)^-\})^2 .\end{aligned}$$

In particular, $(B(B^*B)^-)(B(B^*B)^-)^* \in \mathcal{A}^+$. (Use for instance Proposition 4.23, keeping in mind Definition 4.36.) Thus, by Lemma 4.21 (iii)

$$\begin{aligned}\sigma(-((B^*B)^-)^3) \cup \{0\} &= \sigma((B(B^*B)^-)^*(B(B^*B)^-)) \cup \{0\} \\ &= \sigma((B(B^*B)^-)(B(B^*B)^-)^*) \cup \{0\} \subseteq \mathbb{R}_0^+ .\end{aligned}$$

Hence, as $((B^*B)^-)^3 \in \mathcal{A}^+$

$$\sigma(((B^*B)^-)^3) = \{0\} .$$

This implies, in turn, by the spectral mapping property (Proposition 4.23), that $\sigma((B^*B)^-) = \{0\}$ and one has $(B^*B)^- = 0$.

□

Corollary 4.103 is a key argument in the construction of the celebrated *GNS* (Gelfand-Naimark-Segal) representation of states. This is a pivotal result in the whole theory of operator algebras, obtained in the 1940s. For the sake of simplicity, we start with a version of this construction for the special case of unital C^*-algebras and faithful states. We then explained how to arrive at the most general case afterward.

Corollary 4.104 (GNS Representation for Faithful States of Unital C^*-Algebras) *Let $\mathcal{A}$ be a unital C^*-algebra. Take any faithful state $\rho \in E(\mathcal{A})$. There is a faithful cyclic representation $(H_\rho, \pi_\rho, \Omega_\rho)$ of $\mathcal{A}$ such that, for all $A \in \mathcal{A}$,*

$$\left\langle \Omega_\rho, \pi_\rho(A)(\Omega_\rho)\right\rangle_\rho = \rho(A)$$

with $\langle\cdot,\cdot\rangle_\rho$ being the scalar product in H_ρ, named the GNS scalar semiproduct associated with ρ.

Proof

1. Define the sesquilinear form $\langle\cdot,\cdot\rangle_\rho : \mathcal{A} \times \mathcal{A} \to \mathbb{C}$ by

$$\left\langle A, A'\right\rangle_\rho \doteq \rho(A^*A') , \qquad A, A' \in \mathcal{A} . \tag{4.3}$$

   (See Definition 7.200.) Similar to the special case of the algebra of bounded operators on a complex Hilbert considered previously (see Lemma 2.36 and Proposition 4.114), one shows from Corollary 4.103 that this sesquilinear form is a scalar semiproduct. See Definition 7.198. (It is called, as before, the GNS scalar semiproduct associated with the positive functional ρ.)
2. As ρ is a faithful state, $\langle\cdot,\cdot\rangle_\rho$ is even a scalar product: For all $A \in \mathcal{A}$, note that $A = 0$ iff $A^*A = 0$. Observing that A^*A is positive, $A^*A = 0$ iff $\langle A, A\rangle_\rho = \rho(A^*A) = 0$. Thus, consider the pre-Hilbert space $(\mathcal{A}, \|\cdot\|_\rho)$, where $\|\cdot\|_\rho$ is the norm associated with the scalar product $\langle\cdot,\cdot\rangle_\rho$, and let (H_ρ, ι) be its completion, in the sense of Definition 7.88. Recall that any completion of a pre-Hilbert space is a Hilbert space. To be more pedagogical, we use the notation ι for the canonical mapping $\iota : \mathcal{A} \to \iota(\mathcal{A}) \subseteq H_\rho$, that is, $\iota(A)$ is the element $A \in \mathcal{A}$ seen as a vector of H_ρ. See Definition 7.88.
3. For all $A \in \mathcal{A}$, we define the linear mapping $\tilde{A} \in \mathcal{L}(\iota(\mathcal{A}); H_\rho)$ by

$$\tilde{A}(\iota(B)) \doteq \iota(AB) , \qquad B \in \mathcal{A} .$$

   Note that

$$\|\tilde{A}(\iota(B))\|^2_{H_\rho} = \rho(B^*A^*AB) \leq \|A\|^2 \rho(B^*B) = \|A\|^2 \|\iota(B)\|^2_{H_\rho} .$$

   To obtain this inequality, observe from Corollary 4.103 that $\rho(B^*(\cdot)B) \in \mathcal{L}^+(\mathcal{A}; \mathbb{C})$ and $\|A\|^2 \mathbb{1} \geq A^*A$. Hence, $\|\tilde{A}\|_{\mathrm{op}} \leq \|A\|$.
4. As $\iota(\mathcal{A}) \subseteq H_\rho$ is a dense subset and $\|\tilde{A}\|_{\mathrm{op}} \leq \|A\|$, there is a unique $\pi_\rho(A) \in \mathcal{B}(H_\rho)$ extending $\tilde{A}$. It is easy to check that, for all $A, A', B \in \mathcal{A}$ and $\alpha \in \mathbb{C}$,

$$\pi_\rho(AA')(\iota(B)) = \pi_\rho(A) \circ \pi_\rho(A')(\iota(B)) ,$$
$$\pi_\rho(A + A')(\iota(B)) = (\pi_\rho(A) + \pi_\rho(A'))(\iota(B)) ,$$
$$\pi_\rho(\alpha A)(\iota(B)) = (\alpha\pi_\rho(A))(\iota(B)) .$$

By density of $\iota(\mathcal{A}) \subseteq H_\rho$ and continuity of $\pi_\rho(A)$ and $\pi_\rho(A')$,

$$\pi_\rho(AA') = \pi_\rho(A) \circ \pi_\rho(A') ,$$
$$\pi_\rho(A + A') = \pi_\rho(A) + \pi_\rho(A') ,$$
$$\pi_\rho(\alpha A) = \alpha\pi_\rho(A) .$$

5. For all $A, B, B' \in \mathcal{A}$,

$$\begin{aligned}\left\langle \iota(B), \pi_\rho(A)(\iota(B')) \right\rangle_\rho &= \rho(B^* A B') \\ &= \rho((A^* B)^* B') \\ &= \left\langle \pi_\rho(A^*)(\iota(B)), \iota(B') \right\rangle_\rho .\end{aligned}$$

   Again by density of $\iota(\mathcal{A}) \subseteq H_\rho$, this implies that $\pi_\rho(A)^* = \pi_\rho(A^*)$. Clearly, $\pi_\rho(\mathbb{1}) = \mathrm{id}_{H_\rho}$. Hence, $\pi_\rho : A \mapsto \pi_\rho(A)$ is a unital $*$-homomorphism $\mathcal{A} \to \mathcal{B}(H_\rho)$ and, consequently, (H_ρ, π_ρ) is a representation of $\mathcal{A}$ (Definition 4.88 (i)).
6. Let $\Omega_\rho \doteq \iota(\mathbb{1}) \in H_\rho$. Clearly,

$$\pi_\rho(\mathcal{A})(\Omega_\rho) = \iota(\mathcal{A}) \subseteq H_\rho$$

   is a dense subspace and, thus, Ω_ρ is a cyclic vector of the representation (H_ρ, π_ρ) (Definition 4.88 (iii)). Additionally,

$$\left\| \Omega_\rho \right\|_{H_\rho}^2 = \rho(\mathbb{1}^* \mathbb{1}) = \rho(\mathbb{1}) = 1 .$$

   Hence, $(H_\rho, \pi_\rho, \Omega_\rho)$ is a cyclic representation of $\mathcal{A}$ (Definition 4.88 (iv)).
7. For all $A \in \mathcal{A}$,

$$\left\langle \Omega_\rho, \pi_\rho(A)(\Omega_\rho)) \right\rangle_\rho = \rho(\mathbb{1}^* A \mathbb{1}) = \rho(A) .$$

   Finally, note that, for all $A \in \mathcal{A}$,

$$\left\| \pi_\rho(A)(\Omega_\rho) \right\|_{H_\rho}^2 = \rho(A^* A) .$$

   Therefore, $\left\| \pi_\rho(A) \right\| > 0$ whenever $A \neq 0$. Hence, (H_ρ, π_ρ) is a faithful representation (Definition 4.88 (ii)).

□

The same construction can be performed for all (possibly non-faithful) states, using quotients of vector spaces to remove the degeneracy of the sesquilinear form (4.3), in order to define the Hilbert space H_ρ. We summarize the new steps, as well as a more general version of the last corollary, in the following proposition.

Proposition 4.105 (GNS Representation for States of Unital C^*-Algebras) *Let $\mathcal{A}$ be any unital C^*-algebra and $\rho \in E(\mathcal{A})$ an arbitrary (not necessarily faithful) state. Define the seminorm $\|\cdot\|_\rho$ in $\mathcal{A}$ as in the proof of Corollary 4.104. Let $\tilde{H}_\rho$ be the seminormed space $(\mathcal{A}, \|\cdot\|_\rho)$ and*

$$\mathcal{I}_\rho \doteq \{A \in \tilde{H}_\rho \ : \ \|A\|_\rho = 0\} \subseteq \tilde{H}_\rho \ .$$

For all $A \in \mathcal{A}$, let $\tilde{A} \in \mathcal{L}(\tilde{H}_\rho, \tilde{H}_\rho/\mathcal{I}_\rho)$ be defined by

$$\tilde{A}(B) \doteq [AB] \ , \qquad B \in \tilde{H}_\rho \equiv \mathcal{A} \ .$$

(i) *For all $A \in \mathcal{A}$, $\tilde{A} \in \mathcal{B}(\tilde{H}_\rho, \tilde{H}_\rho/\mathcal{I}_\rho)$ with $\|\tilde{A}\|_{\mathrm{op}} \le \|A\|$.*
(ii) *For every $A \in \mathcal{A}$, $\mathcal{I}_\rho \subseteq \ker(\tilde{A})$.*
(iii) *$\tilde{H}_\rho/\mathcal{I}_\rho$ is a pre-Hilbert space. (In particular, it can be completed to obtain a Hilbert space.)*
(iv) *For any completion $(H_\rho, \mathfrak{i})$ of $\tilde{H}_\rho/\mathcal{I}_\rho$ (Definition 7.88) and all $A \in \mathcal{A}$, there is a unique $\pi(A) \in \mathcal{B}(H_\rho)$ such that*

$$\pi(A)(\mathfrak{i}([B])) = \mathfrak{i}(\tilde{A}(B)) \ , \qquad B \in \tilde{H}_\rho \ .$$

(v) *For any completion $(H_\rho, \mathfrak{i})$ of $\tilde{H}_\rho/\mathcal{I}_\rho$, $(H_\rho, \pi(\cdot), \mathfrak{i}([\mathbb{1}]))$ is a cyclic representation of $\mathcal{A}$.*

Proof With Exercise 7.34 and Definition 7.36 in mind, (i) is proven exactly in the same way as done in the proof of Corollary 4.104 for the special case of faithful states. (iii) follows from Exercises 7.34 (iii), 7.35, and 7.196, keeping in mind Definition 7.194. Observe from (iii) that $\tilde{H}_\rho/\mathcal{I}_\rho$ is a normed space and, hence, (ii) follows from (i). From (i) combined with Exercises 7.8 and 7.38, for all $A \in \mathcal{A}$, there is a bounded operator $\tilde{\pi}(A) \in \mathcal{B}(\tilde{H}_\rho/\mathcal{I}_\rho)$ with $\|\tilde{A}\|_{\mathrm{op}} \le \|A\|$ and

$$\tilde{\pi}(A)([B]) = \tilde{A}(B) \ , \qquad B \in \tilde{H}_\rho \ .$$

Taking $\pi(A) \in \mathcal{B}(H_\rho)$ as being the unique continuous extension of $\mathfrak{i} \circ \tilde{\pi}(A)$ from the dense subspace $\mathfrak{i}(\tilde{H}_\rho/\mathcal{I}_\rho)$ to the whole H_ρ, (iv) follows. To prove that $(H_\rho, \pi(\cdot))$ is a representation of $\mathcal{A}$ (Definition 4.88 (i)), one uses an obvious adaptation of the argument used in the proof of Corollary 4.104 for the special case of faithful states. Again as in the proof of Corollary 4.104, one meanwhile shows that $\Omega_\rho \doteq \iota([\mathbb{1}]) \in H_\rho$ is a cyclic vector (Definition 4.88 (iii)), and we conclude that (H_ρ, π_ρ) is a cyclic representation (Definition 4.88 (iv). Finally, noting that, for all $A, B \in \tilde{H}_\rho \equiv \mathcal{A}$,

$$\langle \mathfrak{i}([A]), \mathfrak{i}([B]) \rangle_\rho = \rho(A^* B) \ ,$$

where $\langle\cdot,\cdot\rangle_\rho$ is the scalar product of the Hilbert space H_ρ, for all $A \in \mathcal{A}$, one has

$$\langle \mathfrak{i}([\mathbb{1}]), \pi(A)\mathfrak{i}([\mathbb{1}])\rangle = \langle \mathfrak{i}([\mathbb{1}]), \mathfrak{i}([A])\rangle = \rho(A) \ .$$

□

By applying Proposition 4.105 (v) to the unitization of an arbitrary (not necessarily unital) C^*-algebra), it follows that any C^*-algebra has a nontrivial cyclic representation. This fact is stated and proved in Theorem 4.113. For more details on the GNS construction, see also [51].

The GNS representation has led to very important applications of the Tomita-Takesaki theory (see, e.g., [51, 60–62]), developed in the 1970s, to quantum field theory and statistical mechanics. These developments mark the beginning of the algebraic approach to quantum mechanics and quantum field theory. For more details, see, e.g., [65].

The continuous functional calculus of C^*-algebras (Proposition 4.100 and Corollary 4.101) can also be used to characterize the smallest C^*-subalgebra, as well as the smallest unital C^*-subalgebra, containing one particular self-conjugate element, which is reminiscent of Examples 4.5 and 4.7.

Lemma 4.106 *Let $\mathcal{A}$ be an arbitrary (not necessarily unital) C^*-algebra and $A \in$* $\mathrm{Re}\{\mathcal{A}\}$. *Then,*

$$\mathcal{A}(A) \doteq \{f(A) \ : \ f \in C_A \equiv C_0(\sigma(A)\backslash\{0\};\mathbb{C})\} \subseteq \mathcal{A}$$

is the smallest C^*-subalgebra of $\mathcal{A}$ containing the element A (also called the "C^*-subalgebra generated by $A \in$ $\mathrm{Re}\{\mathcal{A}\}$*"). If $\mathcal{A}$ is unital, then*

$$\tilde{\mathcal{A}}(A) \doteq \{f(A) \ : \ f \in \tilde{C}_A \doteq C(\sigma(A);\mathbb{C})\} \subseteq \mathcal{A}$$

is the smallest unital C^*-subalgebra of $\mathcal{A}$ containing the element A (also called the unital C^*-subalgebra generated by $A \in$ $\mathrm{Re}\{\mathcal{A}\}$*). Additionally, both $\mathcal{A}(A)$ and $\tilde{\mathcal{A}}(A) \supseteq \mathcal{A}(A)$ (in the unital case) are commutative.*

Proof Exercise. □

Compare this lemma with Examples 4.5 and 4.7, keeping in mind that spectra of self-conjugate C^*-algebra elements are compact subsets of $\mathbb{R}$; see Definition 4.18 and Corollary 4.25. In fact, note that the continuous functional calculus defines a $*$-isomorphism $C_A \rightarrow \mathcal{A}(A)$, as well as a $*$-isomorphism $\tilde{C}_A \rightarrow \tilde{\mathcal{A}}(A)$ in the unital case. This fact implies the following properties of the continuous functional calculus.

Lemma 4.107 *Let $\mathcal{A}_1, \mathcal{A}_2$ be two (not necessarily unital) C^*-algebras and Θ : $\mathcal{A}_1 \rightarrow \mathcal{A}_2$ a $*$-homomorphism. Let $A_1 \in$* $\mathrm{Re}\{\mathcal{A}_1\}$ *and $f \in C_A$, or $f \in \tilde{C}_A$ if $\mathcal{A}_1$ and $\mathcal{A}_2$ are both unital:*

(i) $\sigma(f(A_1)) = f(\sigma(A_1)) \doteq \{f(x) \ : \ x \in \sigma(A_1)\}$. *(Spectral mapping property for continuous functions.)*
(ii) $\Theta(f(A_1)) = f(\Theta(A_1))$. *Note that* $\sigma(\Theta(A_1)) \subseteq \sigma(A_1)$ *in the unital case (Lemma* 4.96 (ii)*).*

Proof Exercise. □

Observe that Lemma 4.107 (i) is a generalization of Proposition 4.23. Additionally, also as a consequence of the fact $\mathcal{A}(A)$ and C_A are equivalent C^*-algebras for any $A \in \mathrm{Re}\{\mathcal{A}\}$, we obtain that the smallest C^*-subalgebra containing one particular self-conjugate element is a Banach lattice (a particular case of a Riesz space), as defined by Definition 7.315:

Proposition 4.108 *Let* $\mathcal{A}$ *be a (not necessarily unital)* C^**-algebra. The* C^**-subalgebra* $\mathcal{A}(A)$ *generated by some fixed* $A \in \mathrm{Re}\{\mathcal{A}\}$ *is a Banach lattice and* $\max\{0, A\} = A^+ \in \mathcal{A}(A)$*, where* A^+ *is the positive part of the orthogonal decomposition of* A*, as defined in Proposition* 4.102 (ii)*. If* $\mathcal{A}$ *is unital, then the same is true for the unital* C^**-subalgebra* $\tilde{\mathcal{A}}(A)$ *generated by* $A \in \mathrm{Re}\{\mathcal{A}\}$ *and* $\mathcal{A}(A)$ *is an order ideal of* $\tilde{\mathcal{A}}(A)$ *(Definition* 1.4*).*

Proof Exercise. *Hint*: See Exercise 1.5 (ii) for the last part of the proposition. □

In fact, recall that, by the Gelfand theorem (see, e.g., [51, Theorem 2.1.11A] or Proposition 4.124 for the separable case), any *commutative unital* C^*-algebra $\mathcal{A}$ is equivalent, as a $*$-algebra, to the algebra of continuous functions on a compact space. Hence, any such a C^*-algebra is a Banach lattice. Additionally, it is not difficult to see that, for any $A \in \mathrm{Re}\{\mathcal{A}\}$, one has again that $\max\{0, A\} = A^+ \in \mathcal{A}(A) \subseteq \mathcal{A}$. So, the above proposition refers only to a very special case of this fact. Thus, if $\mathcal{B} \subseteq \mathcal{A}$ is a C^*-subalgebra of a commutative unital C^*-algebra $\mathcal{A}$, then

$$\max\{0, A\} = A^+ \in \mathcal{A}(A) \subseteq \mathcal{B} \subseteq \mathcal{A} .$$

From this fact we conclude that any C^*-subalgebra of a commutative unital C^*-algebra is a Riesz subspace (Definition 7.267 (ii)) of this algebra. Observing that the unitization of a commutative algebra is again commutative, it follows that any C^*-subalgebra of an arbitrary (not necessarily unital) commutative C^*-algebra is a Riesz subspace of this algebra.

To close the present paragraph, we prove a lemma, which is the main argument to prove Theorem 4.87 (as Corollary 4.110 below) for the special case of injective $*$-homomorphisms.

Lemma 4.109 *Let* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *be two unital* C^**-algebras and* $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ *a faithful (i.e., one-to-one) unital* $*$*-homomorphism. Then,* Θ *preserves the norm, i.e., for all* $A_1 \in \mathcal{A}_1$, $\|\Theta(A_1)\| = \|A_1\|$.

Proof

1. Let $A_1 \in \mathrm{Re}\{\mathcal{A}_1\}$ (i.e., A is self-conjugate). We already know that $\|\Theta(A_1)\| \leq \|A_1\|$, by Lemma 4.96 (ii). Assume that $\|\Theta(A_1)\| < \|A_1\|$. Note that $\Theta(A_1)$ is

self-conjugate, for A_1 is self-conjugate. In particular, $\|\Theta(A_1)\|$ and $\|A_1\|$ are, respectively, the spectral radii of $\Theta(A_1)$ and A_1, thanks to Proposition 4.31 (i).

2. As $\sigma(\Theta(A_1)) \subseteq \sigma(A_1)$ (Lemma 4.96 (ii)), there is a function $f \in \tilde{C}_A$, $f \neq 0$, vanishing on $\sigma(\Theta(A_1)) \subsetneq \sigma(A)$. In particular,
$$\sigma(\Theta(f(A_1))) = \sigma(f(\Theta(A_1))) = \{0\} .$$
For $\Theta(f(A_1))$ is a normal element of $\mathcal{A}_2$ (because $f(A_1)$ is a normal element of $\mathcal{A}_1$), it follows that $\Theta(f(A_1)) = 0$. But $f(A_1) \neq 0$, for $f \neq 0$ and the functional calculus is injective, by Proposition 4.100. This contradicts the fact that Θ is faithful. Therefore, $\|\Theta(A_1)\| = \|A_1\|$ for all $A_1 \in \mathrm{Re}\{\mathcal{A}_1\}$.
3. If $A_1 \in \mathcal{A}_1$ is not necessarily self-conjugate, note that
$$\|\Theta(A_1)\| = \sqrt{\|\Theta(A_1)^*\Theta(A_1)\|} = \sqrt{\left\|\Theta(A_1^*A_1)\right\|} = \sqrt{\left\|A_1^*A_1\right\|} = \|A_1\| ,$$
$A_1^*A_1$ being self-conjugate.

□

Note that this lemma is nothing else than Lemma 4.96 (iii). Observe however that Lemma 4.96 (iii) is proven as an exercise by using Theorem 4.87, whereas we would like to use this result to prove Theorem 4.87. Therefore, to avoid a circle argument, we state and prove Lemma 4.109. In fact, the latter yields the following result for the special case of injective unital $*$-homomorphisms.

Corollary 4.110 *Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two unital C^*-algebras and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ a faithful unital $*$-homomorphism. Then, $\Theta(\mathcal{A}_1) \subseteq \mathcal{A}_2$ is a unital C^*-subalgebra of $\mathcal{A}_2$.*

Proof Exercise. *Hint:* Use the norm preservation property of Θ to show that $\Theta(\mathcal{A}_1)$ is closed in $\mathcal{A}_2$. □

4.6.2 States as Equivalence Classes of Cyclic Representations

In the present paragraph, we show that the states of any C^*-algebra are in one-to-one relation to the cyclic representations of the algebra. We start with the following simple result.

Lemma 4.111 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $R \doteq (H, \pi, \Omega)$ a cyclic representation of $\mathcal{A}$ (Definition 4.88). The linear functional $\rho_R : \mathcal{A} \to \mathbb{C}$ defined by*
$$\rho_R(A) \doteq \langle \Omega, \pi(A)\Omega \rangle , \qquad A \in \mathcal{A} , \tag{4.4}$$
is a state on $\mathcal{A}$.

Proof By Corollary 4.103, any positive element of $\mathcal{A}$ is of the form A^*A for some $A \in \mathcal{A}$. For any such an element one then has

$$\rho_R(A^*A) = \langle \Omega, \pi(A^*A)\Omega \rangle = \langle \pi(A)\Omega, \pi(A)\Omega \rangle = \|\pi(A)\Omega\|^2 \geq 0 .$$

Thus, ρ_R is a positive linear function, and in order to prove that it is a state, we have to show that $\|\rho_R\|_{\text{op}} = 1$, by the definition of states (Definition 4.61 (i)). To this end, take any approximate unit $(E_i)_{i \in I}$ of $\mathcal{A}$ (see Definition 4.49 and Corollary 4.51) and, by Proposition 4.58, note that

$$\|\rho_R\|_{\text{op}} = \lim_{i \in I} \rho_R(E_i^2) = \lim_{i \in I} \|\pi(E_i)\Omega\|^2 .$$

Observing that any cyclic representation is nondegenerate (Exercise 4.94 (i)), the assertion then follows from Lemma 4.95. □

Recall that two cyclic representations (H_1, π_1, Ω_1) and (H_2, π_2, Ω_2) of a C^*-algebra $\mathcal{A}$ are equivalent if there is a unitary linear transformation $U : H_1 \to H_2$ such that $\Omega_2 = U(\Omega_1)$ and, for all $A \in \mathcal{A}$, $\pi_1(A) = U^{-1} \circ \pi_2(A) \circ U$. See Definition 4.88 (viii). It turns out that two equivalent cyclic representations correspond to the same state (4.4) of $\mathcal{A}$:

Lemma 4.112 *Let $R_1 \doteq (\pi_1, H_1, \Omega_1)$ and $R_2 \doteq (\pi_2, H_2, \Omega_2)$ be two cyclic representations of a (not necessarily unital) C^*-algebra $\mathcal{A}$. Then, $\rho_{R_1} = \rho_{R_2}$ iff R_1 and R_2 are equivalent cyclic representations.*

Proof

1. Assume that R_1 and R_2 are equivalent. By definition, there is a unitary linear transformation $U : H_1 \to H_2$ such that

   $$U(\Omega_1) = \Omega_2 \qquad \text{and} \qquad \pi_1(A) = U^{-1} \circ \pi_2(A) \circ U \text{ for all } A \in \mathcal{A} .$$

   Hence, for all $A \in \mathcal{A}$,

   $$\begin{aligned} \rho_{R_2}(A) &= \langle \Omega_2, \pi_2(A)\Omega_2 \rangle \\ &= \langle U(\Omega_1), U \circ \pi_1(A) \circ U^{-1}(U(\Omega_1)) \rangle \\ &= \langle U(\Omega_1), U \circ \pi_1(A)(\Omega_1) \rangle \\ &= \langle \Omega_1, \pi_1(A)\Omega_1 \rangle = \rho_{R_1}(A) . \end{aligned}$$

2. Assume now that $\rho_{R_1} = \rho_{R_2}$ and define the (dense) subspace

   $$\widetilde{H}_1 \doteq \{\pi_1(A)(\Omega_1) \ : \ A \in \mathcal{A}\} \subseteq H_1 .$$

Observe that, for all $A \in \mathcal{A}$,

$$\begin{aligned}
\|\pi_1(A)(\Omega_1)\|^2 &= \langle \pi_1(A)(\Omega_1), \pi_1(A)(\Omega_1)\rangle \\
&= \left\langle \Omega_1, \pi_1(A)^*\pi_1(A)(\Omega_1)\right\rangle \\
&= \left\langle \Omega_1, \pi_1(A^*A)(\Omega_1)\right\rangle \\
&= \rho_{R_1}(A^*A) \\
&= \rho_{R_2}(A^*A) \\
&= \langle \pi_2(A)(\Omega_2), \pi_2(A)(\Omega_2)\rangle = \|\pi_2(A)(\Omega_2)\|^2 \ .
\end{aligned}$$

In particular, for all $A, A' \in \mathcal{A}$, one has $\pi_1(A)(\Omega_1) = \pi_1(A')(\Omega_1)$ (i.e., $\left\|\pi_1(A - A')(\Omega_1)\right\| = 0$) iff $\pi_2(A)(\Omega_2) = \pi_2(A')(\Omega_2)$ (i.e., $\left\|\pi_2(A - A')(\Omega_2)\right\| = 0$). Thus, the condition

$$\widetilde{U}(\pi_1(A)(\Omega_1)) = \pi_2(A)(\Omega_2) \ , \qquad A \in \mathcal{A} \ ,$$

well defines a mapping $\widetilde{U} : \widetilde{H}_1 \to H_2$. This mapping is linear and norm-preserving, by the above equality. As the subspace $\widetilde{H}_1$ is dense in H_1 (because Ω_1 is cyclic for R_1), $\widetilde{U}$ has a unique linear bounded extension $U \in \mathcal{B}(H_1; H_2)$, which is also norm-preserving.

3. For the subspace

$$\widetilde{H}_2 \doteq \{\pi_2(A)(\Omega_2) \ : \ A \in \mathcal{A}\} \subseteq H_2$$

is dense (Ω_2 is cyclic for R_2), we conclude furthermore that U is surjective. U is a so-called isometry and in fact a unitary linear transformation $H_1 \to H_2$. Observe for instance that, for all $A, A' \in \mathcal{A}$,

$$\begin{aligned}
\langle U\pi_1(A)\Omega_1, U\pi_1(A')\Omega_1\rangle &= \langle \pi_2(A)\Omega_2, \pi_2(A')\Omega_2\rangle \\
&= \langle \Omega_2, \pi_2(A^*A')\Omega_2\rangle \\
&= \rho_{R_2}(A^*A') \\
&= \rho_{R_1}(A^*A') \\
&= \langle \Omega_1, \pi_1(A^*A')\Omega_1\rangle \\
&= \langle \pi_1(A)\Omega_1, \pi_1(A')\Omega_1\rangle \ .
\end{aligned}$$

4. Let $(E_i)_{i\in I}$ be any approximate unit of $\mathcal{A}$; see Definition 4.49 and Corollary 4.51. By Lemma 4.95,

$$U(\Omega_1) = \lim_{i\in I} U(\pi_1(E_i)(\Omega_1)) = \lim_{i\in I} \widetilde{U}(\pi_1(E_i)(\Omega_1)) = \lim_{i\in I} \pi_2(E_i)(\Omega_2) = \Omega_2 \ .$$

5. Finally, for all $A, A' \in \mathcal{A}$,

$$\begin{aligned} U^{-1} \circ \pi_2(A) \circ U(\pi_1(A')(\Omega_1)) &= U^{-1} \circ \pi_2(A)(\widetilde{U}(\pi_1(A')(\Omega_1))) \\ &= U^{-1} \circ \pi_2(A)(\pi_2(A')(\Omega_2)) \\ &= U^{-1}(\pi_2(AA')(\Omega_2)) \\ &= \pi_1(AA')(\Omega_1) \\ &= \pi_1(A)(\pi_1(A')(\Omega_1)) \, . \end{aligned}$$

As the operators $U^{-1} \circ \pi_2(A) \circ U$ and $\pi_1(A)$ are bounded and the subspace $\widetilde{H}_1 \subseteq H_1$ is dense, one has

$$\pi_1(A) = U^{-1} \circ \pi_2(A) \circ U$$

for every $A \in \mathcal{A}$.

□

Observe, however, that if two cyclic representations (H_1, π_1, Ω_1) and (H_2, π_2, Ω_2) of a C^*-algebra $\mathcal{A}$ do not define the same state (being thus inequivalent cyclic representations), then we can *not* conclude that the representations (H_1, π_1) and (H_2, π_2) are not equivalent (Definition 4.88 (vii)). In other words, this situation does not prevent the representations (H_1, π_1) and (H_2, π_2) from being equivalent.

By the last lemma, equivalence classes of cyclic representations are one-to-one related to states. The following theorem says that this relation is a one-to-one correspondence and states of any C^*-algebra can thus be seen as equivalence classes of cyclic representations of this algebra.

Theorem 4.113 (GNS Representation) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra and $\rho \in E(\mathcal{A})$. There is a cyclic representation $(H_\rho, \pi_\rho, \Omega_\rho)$ of $\mathcal{A}$ such that*

$$\rho(A) = \langle \Omega_\rho, \pi_\rho(A)\Omega_\rho \rangle$$

for all $A \in \mathcal{A}$. The representation π_ρ is faithful whenever the state ρ is faithful.

Proof Recall that in the unital case, this theorem refers to Proposition 4.105 (v). If $\mathcal{A}$ is not unital, consider a cyclic representation $(H_{\tilde{\rho}}, \pi_{\tilde{\rho}}, \Omega_{\tilde{\rho}})$ for the unique extension $\tilde{\rho} \in E(\tilde{\mathcal{A}})$ of $\rho \in E(\mathcal{A})$, which exists by Exercise 4.63. Let $(E_i)_{i \in I}$ be any approximate unit for $\mathcal{A} \subseteq \tilde{\mathcal{A}}$. See Definition 4.49 and Corollary 4.51. From Proposition 4.58,

$$1 = \|\rho\|_{\mathrm{op}} = \lim_{i \in I} \rho(E_i^2) = \lim_{i \in I} \left\langle \Omega_{\tilde{\rho}}, \pi_{\tilde{\rho}}(E_i^2)(\Omega_{\tilde{\rho}}) \right\rangle \, .$$

As $\pi_{\tilde{\rho}}$ is a contraction (by Proposition 4.97) and $\|E_i\| \leq 1$ for all $i \in I$, it follows that

$$\lim_{i \in I} \pi_{\tilde{\rho}}(E_i^2)(\Omega_{\tilde{\rho}}) = \Omega_{\tilde{\rho}} .$$

This fact implies that the subspace $\pi_{\tilde{\rho}}(\mathcal{A})\Omega_{\tilde{\rho}} \subseteq \pi_{\tilde{\rho}}(\tilde{\mathcal{A}})\Omega_{\tilde{\rho}} \subseteq H_{\tilde{\rho}}$ is dense in $H_{\tilde{\rho}}$. Thus, $(H_\rho, \pi_\rho, \Omega_\rho)$, where $H_\rho \doteq H_{\tilde{\rho}}$, $\Omega_\rho \doteq \Omega_{\tilde{\rho}}$ and π_ρ is the restriction of $\pi_{\tilde{\rho}}$ to (the subalgebra) $\mathcal{A} \subseteq \tilde{\mathcal{A}}$, is a cyclic representation of $\mathcal{A}$ with

$$\rho(A) = \langle \Omega_\rho, \pi_\rho(A)\Omega_\rho \rangle , \qquad A \in \mathcal{A} .$$

□

For more details on the GNS representation, see also [51].

To conclude this paragraph, we show that the purity (i.e., extremality, by Proposition 4.71) of states is equivalent to the irreducibility of the associated cyclic representations (see Definition 4.88). With this aim, we first prove the following technical result.

Proposition 4.114 *Let $\mathcal{A}$ be a (not necessarily unital) C^*-algebra and take any state $\rho \in E(\mathcal{A})$. Let $(H_\rho, \pi_\rho, \Omega_\rho)$ be any cyclic representation of $\mathcal{A}$ associated with ρ. The linear transformation $T \mapsto \rho_T$ from $\mathcal{B}(H_\rho)$ to $\mathcal{A}^{\mathrm{td}}$ defined by*

$$\rho_T(A) \doteq \left\langle \Omega_\rho, T \circ \pi_\rho(A)(\Omega_\rho) \right\rangle , \qquad A \in \mathcal{A} , \tag{4.5}$$

is bipositive on $\pi_\rho(\mathcal{A})' \subseteq \mathcal{B}(H_\rho)$ and establishes a one-to-one correspondence between

$$\{T \in \mathcal{B}(H_\rho)^+ \cap \pi_\rho(\mathcal{A})' : \|T\|_{\mathrm{op}} \leq 1\} ,$$

which is noting else that the interval $[0, \mathbb{1}]$ in $\pi_\rho(\mathcal{A})' \subseteq \mathcal{B}(H_\rho)$ seen as a unital C^*-algebra, and the interval

$$[0, \rho] \doteq \mathcal{A}'^+ \cap (\rho - \mathcal{A}'^+) = \{\rho' \in \mathcal{A}' \ : \ \rho \succeq \rho' \succeq 0\} \subseteq \mathcal{A}'^+ .$$

By the first part of the proposition, this correspondence is order-preserving and order-reflecting,[16] *that is, it is an equivalence of partially ordered sets.*

Proof

1. Take any $T \in \mathcal{B}(H_\rho)$. Clearly, $\rho_T \in \mathcal{A}^{\mathrm{td}}$, i.e., it is a continuous linear functional on $\mathcal{A}$. Recall that $\pi_\rho(\mathcal{A})'$ is, by definition, the commutator of $\pi_\rho(\mathcal{A}) \subseteq \mathcal{B}(H_\rho)$, that is,

$$\pi_\rho(\mathcal{A})' \doteq \{B \in \mathcal{B}(H_\rho) : [B, \pi_\rho(A)] \doteq B\pi_\rho(A) - \pi_\rho(A)B = 0 \text{ for all } A \in \mathcal{A}\} .$$

16 That is, for all $T_1, T_2 \in \mathcal{B}(H_\rho)$, $\rho_{T_1} \succeq \rho_{T_2}$ only if $T_1 \geq T_2$.

If $T \in \mathcal{B}(H_\rho)^+ \cap \pi_\rho(\mathcal{A})'$, then, for all $A \in \mathcal{A}$,

$$\begin{aligned}\rho_T(A^*A) &= \langle \Omega_\rho, T \circ \pi_\rho(A^*A)(\Omega_\rho)\rangle \\ &= \langle \Omega_\rho, T \circ \pi_\rho(A)^* \circ \pi_\rho(A)(\Omega_\rho)\rangle \\ &= \langle \Omega_\rho, \pi_\rho(A)^* \circ T \circ \pi_\rho(A)(\Omega_\rho)\rangle \\ &= \langle \pi_\rho(A)(\Omega_\rho), T \circ \pi_\rho(A)(\Omega_\rho)\rangle \geq 0 \, .\end{aligned}$$

As any positive element of $\mathcal{A}$ equals A^*A for some $A \in \mathcal{A}$ (Corollary 4.103), we conclude that $\rho_T \in \mathcal{A}'^+$, in this case. In other words, the mapping $T \mapsto \rho_T$ from $\pi_\rho(\mathcal{A})' \subseteq \mathcal{B}(H_\rho)$ to $\mathcal{A}^{\mathrm{td}}$ is positive.

2. Take $T, T' \in \pi_\rho(\mathcal{A})' \subseteq \mathcal{B}(H_\rho)$. If $\rho_T = \rho_{T'}$ then, for all $A \in \mathcal{A}$,

$$\langle \pi_\rho(A)(\Omega_\rho), (T - T') \circ \pi_\rho(A)(\Omega_\rho)\rangle = 0 \, .$$

By cyclicity of Ω_ρ and continuity of scalar products in Hilbert spaces, one has that

$$\langle x, (T - T')(x)\rangle = 0$$

for every $x \in H_\rho$. But this implies that $T - T' = 0$. Hence, $T \mapsto \rho_T$ defines an injective mapping from $\pi_\rho(\mathcal{A})'$ to $\mathcal{A}^{\mathrm{td}}$.

3. Take any $T \in \pi'_\rho(\mathcal{A})$ for which $\rho_T \in \mathcal{A}'^+$. In particular, one has $\rho_T(A^*A) \geq 0$ for every $A \in \mathcal{A}$. By a similar argument as the one in the last point, we conclude that $\langle x, T(x)\rangle \geq 0$ for every $x \in H_\rho$. But this implies that $T \in \mathcal{B}(H_\rho)^+$. Together with the first point, we deduce that the mapping $T \mapsto \rho_T$ from $\pi_\rho(\mathcal{A})'$ to $\mathcal{A}^{\mathrm{td}}$ is bipositive (see Definition 1.7).
4. For $T \in \mathcal{B}(H_\rho)^+ \cap \pi_\rho(\mathcal{A})'$, $\|T\|_{\mathrm{op}} \leq 1$, one has that for all $A \in \mathcal{A}$,

$$\begin{aligned}\rho(A^*A) - \rho_T(A^*A) &= \langle \pi_\rho(A)(\Omega_\rho), \pi_\rho(A)(\Omega_\rho)\rangle \\ &\quad - \langle \pi_\rho(A)(\Omega_\rho), T \circ \pi_\rho(A)(\Omega_\rho)\rangle \\ &\geq (1 - \|T\|_{\mathrm{op}}) \left\|\pi_\rho(A)(\Omega_\rho)\right\|^2 \geq 0 \, .\end{aligned}$$

Hence, $\rho \succeq \rho_T$ and $T \mapsto \rho_T$ thus define an (injective, order-preserving, and order-reflecting) mapping from $\{T \in \mathcal{B}(H_\rho)^+ \cap \pi_\rho(\mathcal{A})' : \|T\|_{\mathrm{op}} \leq 1\}$ to $[0, \rho]$.

5. It remains to prove the surjectivity of this mapping. Take any $\varphi \in [0, \rho]$ and note that, for all $A, A' \in \mathcal{A}$, by the Cauchy-Schwarz inequality (for the GNS scalar semiproduct associated with the positive functional φ),

$$\begin{aligned}|\varphi(A^*A')| &\leq \sqrt{\varphi(A^*A)}\sqrt{\varphi(A'^*A')} \leq \sqrt{\rho(A^*A)}\sqrt{\rho(A'^*A')} \\ &= \left\|\pi_\rho(A)(\Omega_\rho)\right\| \left\|\pi_\rho(A')(\Omega_\rho)\right\| \, .\end{aligned}$$

In particular, for $B, B' \in \mathcal{A}$ with $\pi_\rho(A)(\Omega_\rho) = \pi_\rho(B)(\Omega_\rho)$ and $\pi_\rho(A')(\Omega_\rho) = \pi_\rho(B')(\Omega_\rho)$ (i.e., $\left\|\pi_\rho(A - B)(\Omega_\rho)\right\| = 0$ and $\left\|\pi_\rho(A' - B')(\Omega_\rho)\right\| = 0$), one has

$$\varphi(B^* B') = \varphi(A^* A') \ .$$

6. Let $\tilde{H}_\rho \doteq \pi_\rho(\mathcal{A})(\Omega_\rho)$. Note that $\tilde{H}_\rho$ is a dense subspace of H_ρ, thanks to the cyclicity of Ω_ρ. By the last point, the condition

$$\tilde{S}(\pi_\rho(A)(\Omega_\rho), \pi_\rho(A')(\Omega_\rho)) = \varphi(A^* A') \ , \qquad A, A' \in \mathcal{A} \ ,$$

well defines a sesquilinear form $\tilde{S} : \tilde{H}_\rho \times \tilde{H}_\rho \to \mathbb{C}$ (see Definition 7.200). Moreover, by the last inequality, for all $x, x' \in \tilde{H}_\rho$,

$$|\tilde{S}(x, x')| \leq \|x\| \, \|x'\| \ .$$

From this and the density of the subspace $\tilde{H}_\rho$ in H_ρ, there is a unique sesquilinear form $S : H_\rho \times H_\rho \to \mathbb{C}$ satisfying

$$|S(x, x')| \leq \|x\| \, \|x'\|$$

for every $x, x' \in H_\rho$ and

$$S(x, x') = \tilde{S}(x, x')$$

for every $x, x' \in \tilde{H}_\rho$. From the Riesz representation theorem (Corollary 7.216, or Riesz lemma [66, Theorem 11.4]), there is a unique $T \in \mathcal{B}(H_\rho)$ such that

$$S(x, x') = \left\langle x, T(x') \right\rangle \ , \qquad x, x' \in H_\rho \ .$$

Moreover, $\|T\|_{\mathrm{op}} \leq 1$.
7. Note that

$$\left\langle \pi_\rho(A)(\Omega_\rho), T \circ \pi_\rho(A)(\Omega_\rho) \right\rangle = \varphi(A^* A) \geq 0$$

for all $A \in \mathcal{A}$. By cyclicity of Ω_ρ and boundedness of T, $\langle x, T(x) \rangle \geq 0$ for all $x \in H_\rho$ and, hence, T is positive.
8. For all $A, A' \in \mathcal{A}$,

$$\begin{aligned}
&\left\langle \pi_\rho(A')(\Omega_\rho), T \circ \pi_\rho(A) \circ \pi_\rho(A')(\Omega_\rho) \right\rangle \\
&= \varphi(A'^*(AA')) \\
&= \varphi((A^* A')^* A')
\end{aligned}$$

$$= \left\langle \pi_\rho(A)^* \circ \pi_\rho(A')(\Omega_\rho), T \circ \pi_\rho(A')(\Omega_\rho) \right\rangle$$
$$= \left\langle \pi_\rho(A')(\Omega_\rho), \pi_\rho(A) \circ T \circ \pi_\rho(A')(\Omega_\rho) \right\rangle ,$$

that is,

$$\left\langle \pi_\rho(A')(\Omega_\rho), (\pi_\rho(A) \circ T - T \circ \pi_\rho(A)) \circ \pi_\rho(A')(\Omega_\rho) \right\rangle = 0 .$$

Again by cyclicity of Ω_ρ (and boundedness of T), for all $x \in H_\rho$,

$$\left\langle x, (\pi_\rho(A) \circ T - T \circ \pi_\rho(A))(x) \right\rangle = 0 .$$

Hence, $\pi_\rho(A) \circ T = T \circ \pi_\rho(A)$ for all $A \in \mathcal{A}$, that is, $T \in \pi'_\rho(\mathcal{A})$.

9. Finally, take any approximate unit $(E_i)_{i\in I}$ of $\mathcal{A}$ (see Definition 4.49 and Corollary 4.51) and note from Exercise 4.94 and Lemma 4.95 that, for all $A \in \mathcal{A}$,

$$\rho_T(A) = \left\langle \Omega_\rho, T \circ \pi_\rho(A)(\Omega_\rho) \right\rangle = \lim_{i\in I} \left\langle \pi_\rho(E_i)\Omega_\rho, T \circ \pi_\rho(A)(\Omega_\rho) \right\rangle$$
$$= \lim_{i\in I} \tilde{S}(\pi_\rho(E_i)(\Omega_\rho), \pi_\rho(A)(\Omega_\rho)) = \lim_{i\in I} \varphi(E_i A) = \varphi(A) .$$

□

In the commutative case, recall that states of unital C^*-algebras refer to probability measures, thanks to the Riesz-Markov theorem (Theorem 4.68). If ρ is any state and $\varphi \in [0, \rho]$, both seen as classical measures, then φ is absolutely continuous with respect to ρ. The above proposition says in this case that the measure φ is obtained from ρ by multiplication, in a sense, with an algebra element, which can be seen as the "density" of φ with respect to ρ. Thus, the proposition can be seen as a noncommutative version of the classical Radon-Nikodym theorem [66, Theorem I.19]. There are other important results of this kind that will not be discussed here.

Recall now that a representation (H, π) of a unital C^*-algebra is irreducible if $G = \{0\}$ and $G = H$ are the unique *closed* subspaces of H such that $\pi(\mathcal{A})G \subseteq G$, by Definition 4.88 (v).

Corollary 4.115 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. The state $\rho \in E(\mathcal{A})$ is extreme iff its associated cyclic representation $(H_\rho, \pi_\rho, \Omega_\rho)$ is irreducible.*

Proof

1. Note that, for all $\alpha \in [0, 1]$, $\alpha\rho = \rho_{\alpha \mathrm{id}_{H_\rho}}$, as defined by (4.5). If (H_ρ, π_ρ) is irreducible, then, by Theorem 4.92 and Proposition 4.114, for any $\varphi \in [0, \rho]$, there is

$$T \in \mathcal{B}(H_\rho)^+ \cap \pi_\rho(\mathcal{A})' = \mathcal{B}(H_\rho)^+ \cap \mathbb{C}\mathrm{id}_{H_\rho}$$

such that $\|T\|_{\text{op}} \leq 1$ and $\varphi = \rho_T$. But, in this case, $T = \alpha \text{id}_{H_\rho}$, that is, $\varphi = \alpha\rho$, for some $\alpha \in [0, 1]$. As a consequence, ρ is pure and equivalently extreme, thanks to Proposition 4.71.

2. If (H_ρ, π_ρ) is not irreducible, then there is some $A \in \pi'_\rho(\mathcal{A})$ with $A \notin \mathbb{C}\text{id}_{H_\rho}$. As $\pi_\rho(\mathcal{A})'$ is a unital C^*-subalgebra of $\mathcal{B}(H_\rho)$, by Lemma 4.38, there are

$$A^+_{\text{Re}}, A^-_{\text{Re}}, A^+_{\text{Im}}, A^-_{\text{Im}} \in \pi_\rho(\mathcal{A})' \cap \mathcal{B}(H_\rho)^+$$

such that

$$A = (A^+_{\text{Re}} - A^-_{\text{Re}}) + i(A^+_{\text{Im}} - A^-_{\text{Im}}) .$$

Clearly, at least one of these four positive elements is not in $\mathbb{C}\text{id}_{H_\rho}$. This means that $\pi_\rho(\mathcal{A})'$ contains some positive $T \notin \mathbb{C}\text{id}_{H_\rho}$. By multiplying it with a well-chosen positive constant, we can assume that $\|T\|_{\text{op}} \leq 1$. By Proposition 4.114, $\rho_T \in [0, \rho]$ and $\rho_T \neq \alpha\rho$ for every $\alpha \in [0, 1]$. In other words, ρ is not pure and, hence, not an extreme state (Proposition 4.71).

□

Corollary 4.116 *Let H be any complex Hilbert space. Any vector state on $\mathcal{B}(H)$ is an extreme state.*

Proof Exercise. □

4.7 Commutative C^*-Algebras

We already discussed commutative unital C^*-algebras via the example of complex-valued continuous functions on a compact Hausdorff space, for instance in Examples 4.65, 4.67 as well as in Corollaries 4.72 and 4.73. Here, we show that this case is in fact already the most general one, up to a $*$-isomorphism. This is done via the concept of characters, as defined below. Then, we will deduce from this result on unital commutative C^*-algebras that algebras of functions decaying at infinity on a locally compact space constitute the most general example of non-unital commutative C^*-algebras, up to $*$-isomorphisms. See Corollary 4.77 in this context.

Definition 4.117 (Character) Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. A self-conjugate linear functional $\varphi : \mathcal{A} \to \mathbb{C}$ is said to be a "character" of $\mathcal{A}$ if it is non-zero and multiplicative, i.e., for all $A, A' \in \mathcal{A}$, $\varphi(AA') = \varphi(A)\varphi(A')$. The (possibly empty) set of all characters of $\mathcal{A}$ is denoted by $X(\mathcal{A})$.

By seeing $\mathbb{C}$ as a C^*-algebra, note that a character of $\mathcal{A}$ is, by definition, a $*$-homomorphism $\mathcal{A} \to \mathbb{C}$. It turns out that any character is an extreme state.

Lemma 4.118 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. $X(\mathcal{A})$ is a set of extreme states of $\mathcal{A}$.*

Proof

1. If $X(\mathcal{A})$ is empty, then there is nothing to be proven. Thus, suppose that $X(\mathcal{A})$ is nonempty and take any $\varphi \in X(\mathcal{A})$. Then, for all $A \in \mathrm{Re}\{\mathcal{A}\}$, $\varphi(A^2) = \varphi(A)^2 \geq 0$, by multiplicativity of φ. Note here that $\varphi(A) \in \mathbb{R}$, for φ and A are self-conjugate. Hence, φ is a positive linear functional. In particular, by Proposition 4.58, it is continuous and, for any approximate unity $(E_i)_{i\in I}$ for $\mathcal{A}$, one has that $\|\varphi\|_{\mathrm{op}} = \lim_{i\in I} \varphi(E_i^2) < \infty$. Recall that every non-unital C^*-algebra possesses an approximate unit; see Definition 4.49 and Corollary 4.51.
2. Again by multiplicativity of φ and Proposition 4.58 for some approximate unit $(E_i)_{i\in I}$ of $\mathcal{A}$,

$$\|\varphi\|_{\mathrm{op}}^2 = \left(\lim_{i\in I} \varphi(E_i^2)\right)\left(\lim_{j\in I} \varphi(E_j^2)\right) = \lim_{i\in I}\lim_{j\in I} \varphi(E_i^2 E_j^2) = \lim_{j\in I} \varphi(E_j^2) = \|\varphi\|_{\mathrm{op}} .$$

   Note that we used above that, for all $j \in I$,

$$\lim_{i\in I}(E_i - 1)E_j^2 = 0$$

   and, thus,

$$\lim_{i\in I} E_i^2 E_j^2 = \lim_{i\in I} E_i((E_i - 1) + 1)E_j^2 = \lim_{i\in I} E_i E_j^2 = E_j^2 .$$

   As $\varphi \neq 0$, it follows that $\|\varphi\|_{\mathrm{op}} = 1$. Therefore, any character $\varphi \in X(\mathcal{A})$ is a state.
3. Pick $\rho \in X(\mathcal{A}) \subseteq E(\mathcal{A})$ and let $(H_\rho, \pi_\rho, \Omega_\rho)$ be a cyclic representation associated with the state ρ. Take any $A, A' \in \mathcal{A}$ such that

$$\left\|\pi_\rho(A)(\Omega_\rho)\right\|^2 = \rho(A^*A) \neq 0 \qquad \text{and} \qquad \left\|\pi_\rho(A')(\Omega_\rho)\right\|^2 = \rho(A'^*A') \neq 0 .$$

   In particular, by multiplicativity of ρ, one has that $\rho(A), \rho(A') \neq 0$. Again by multiplicativity of ρ, for all $c, c' \in \mathbb{C}$,

$$\rho((cA + c'A')^*(cA + c'A')) = \overline{(c\rho(A) + c'\rho(A'))}(c\rho(A) + c'\rho(A')) .$$

4. Hence, choosing $c = \rho(A') \neq 0$ and $c' = -\rho(A) \neq 0$,

$$\left\|c\pi_\rho(A)(\Omega_\rho) + c'\pi_\rho(A')(\Omega_\rho)\right\|^2 = \rho((cA + c'A')^*(cA + c'A')) = 0 .$$

   This implies that the Hilbert space H_ρ is one-dimensional, by cyclicity of Ω_ρ. In particular, $\pi_\rho'(\mathcal{A}) = \mathbb{C}\mathrm{id}_{H_\rho}$. By Theorem 4.92 and Corollary 4.115, it then follows that ρ is an extreme state.

□

Exercise 4.119 Let $\mathcal{A}$ be any separable *unital* C^*-algebra. Prove that $X(\mathcal{A}) \subseteq E(\mathcal{A})$ is compact with respect to the weak* topology.

Beyond the fact that characters are extreme states of unital C^*-algebra, in case they exist of course, their expectation values are always spectral values of the corresponding algebra elements.

Lemma 4.120 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. For all $\rho \in X(\mathcal{A})$ and all $A \in \mathcal{A}$, $\rho(A)$ belongs to the spectrum of A.*

Proof We first consider the case of $\mathcal{A}$ being unital. Fix $A \in \mathcal{A}$ and take any $z \in \mathbb{C}$ in the resolvent set $R_{\mathcal{A}}(A)$ of A (Definition 4.18). Then, by the multiplicativity of characters, for all $\rho \in X(\mathcal{A})$,

$$1 = \rho(\mathbb{1}) = \rho((z\mathbb{1} - A)(z\mathbb{1} - A)^{-1}) = (z - \rho(A))\rho((z\mathbb{1} - A)^{-1}) .$$

In particular, $\rho(A) \neq z$ for any $z \in R_{\mathcal{A}}(A)$, i.e., $\rho(A) \in \sigma(A)$. If $\mathcal{A}$ is non-unital, for any given character $\rho \in X(\mathcal{A})$, consider the unique unital $*$-homomorphism $\tilde{\rho} : \tilde{\mathcal{A}} \to \mathbb{C}$ extending ρ to the unitization $\tilde{\mathcal{A}}$ of $\mathcal{A}$. See Exercise 7.64. Observe that $\tilde{\rho}$ is again a character. Hence, by the first part of the proof, $\rho(A) = \tilde{\rho}(A) \in \sigma(A)$. □

Conversely, given any self-conjugate element $A \in \mathrm{Re}\{\mathcal{A}\}$ of a C^*-algebra $\mathcal{A}$, we may use characters to construct states $\rho \in E(\mathcal{A})$ with $\rho(A) = a$ for any chosen spectral value $a \in \sigma(A)\backslash\{0\}$:

Lemma 4.121 *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. For any $A \in \mathrm{Re}\{\mathcal{A}\}$ and all $a \in \sigma(A)\backslash\{0\}$, there is a state $\rho \in E(\mathcal{A})$ with $\rho(A) = a$. The state ρ may be chosen as being a pure (i.e., extreme) state. Additionally, there is a sequence of states $\rho_n \in E(\mathcal{A})$, $n \in \mathbb{N}$, such that $\lim_{n\to\infty} \rho_n(A) = 0 \in \sigma(A)$. (Recall by Exercise 4.20 that 0 must be a spectral value of A, $\mathcal{A}$ being non-unital.)*

Proof Take any $A \in \mathrm{Re}\{\mathcal{A}\}$ and $a \in \sigma(A)$. Recall from Proposition 4.100 that the continuous functional calculus defines a $*$-isomorphism $C(\sigma(A); \mathbb{C}) \to \tilde{\mathcal{A}}(A) \subseteq \tilde{\mathcal{A}}$, where $\tilde{\mathcal{A}}$ denotes the unitization of $\mathcal{A}$ and

$$\tilde{\mathcal{A}}(A) \doteq \{f(A) \ : \ f \in C(\sigma(A); \mathbb{C})\} \subseteq \mathcal{A}$$

is the smallest unital C^*-subalgebra of $\tilde{\mathcal{A}}$ containing the element A, as in Lemma 4.106. Thus, there is a unique linear functional $\tilde{\varphi} : \tilde{\mathcal{A}}(A) \to \mathbb{C}$ satisfying

$$\tilde{\varphi}(f(A)) = f(a) , \qquad f \in C(\sigma(A); \mathbb{C}) .$$

It is clearly a character. Let φ be the restriction of $\tilde{\varphi}$ to the smallest C^*-subalgebra $\mathcal{A}(A) \subseteq \tilde{\mathcal{A}}(A)$ of $\mathcal{A}$ containing the element A (see again Lemma 4.106). If $a \neq 0$ then φ is non-zero, as $\varphi(A) = a$. In particular, it is a character of $\mathcal{A}(A) \subseteq \mathcal{A}$. By Lemma 4.118 it is an extreme state on $\mathcal{A}(A)$. By Lemma 4.75 (i) it extends to a state $\rho \in E(\mathcal{A})$ with $\rho(A) = a$. By Theorem 4.76 (i) ρ may be chosen as being an extreme state. Finally, to prove the last part of the lemma, one proceeds as in the proof of Corollary 4.91. Note that this proof does not require anymore the separability of $\mathcal{A}$, which is assumed in the corollary, by other reasons (in fact to ensure the weak* convergence in Corollary 4.91). □

We now improve Lemma 4.118 in the special case of commutative unital C^*-algebras, which is our main concern here.

Lemma 4.122 *If $\mathcal{A}$ is a* commutative *(not necessarily unital) C^*-algebra, then $X(\mathcal{A})$ is exactly the set $\mathcal{E}(\mathcal{A})$ of all extreme states of $\mathcal{A}$. In particular, $X(\mathcal{A})$ is nonempty.*

Proof From Lemma 4.118, one has $X(\mathcal{A}) \subseteq \mathcal{E}(\mathcal{A})$. So, we have to prove that any extreme state on $\mathcal{A}$ is a character. Let $\rho \in \mathcal{E}(\mathcal{A})$ and take any cyclic representation $(\mathcal{H}_\rho, \pi_\rho, \Omega_\rho)$ for this state. Then, by Theorem 4.92 and Corollary 4.115, $\pi_\rho(\mathcal{A})' = \mathbb{C}\mathrm{id}_H$. For $\mathcal{A}$ is commutative, $\pi_\rho(\mathcal{A}) \subseteq \pi_\rho(\mathcal{A})'$ and, consequently, $\pi_\rho(\mathcal{A}) = \mathbb{C}\mathrm{id}_H$. From this it follows that ρ is multiplicative and, hence, a character of $\mathcal{A}$. □

Therefore, the Gelfand transform of Definition 4.79 can be refined in the case of commutative C^*-algebras as follows.

Definition 4.123 (Gelfand Transform for Commutative C^*-Algebra) Let $\mathcal{A}$ be any *commutative* (not necessarily unital) C^*-algebra. We define a linear transformation $\Xi : \mathcal{A} \to \mathcal{F}(X(\mathcal{A}); \mathbb{C})$ by

$$\Xi(A)(\rho) \doteq \rho(A) , \qquad A \in \mathcal{A},\ \rho \in E(\mathcal{A}) .$$

The function $\Xi(A) : X(\mathcal{A}) \to \mathbb{C}$ is well-known in the literature as the "Gelfand transform" of $A \in \mathcal{A}$.

Given any separable commutative C^*-algebra $\mathcal{A}$, from the results of Sect. 4.5.1, note that, for every $A \in \mathcal{A}$, the Gelfand transform $\Xi(A) : X(\mathcal{A}) \to \mathbb{C}$ is weak* continuous. If $\mathcal{A}$ is unital, then Ξ is a mapping from $\mathcal{A}$ to the C^*-algebra $C_{w^*}(X(\mathcal{A}); \mathbb{C})$, for $X(\mathcal{A})$ is weak*-compact (Exercise 4.119). In fact, the Gelfand transform defined above is even a $*$-isomorphism of C^*-algebras.

Proposition 4.124 *For any separable commutative unital C^*-algebra $\mathcal{A}$, the Gelfand transform $\Xi : \mathcal{A} \to C_{w^*}(X(\mathcal{A}); \mathbb{C})$ is a $*$-isomorphism. In particular, it is a bipositive bijective linear transformation, i.e., an equivalence of ordered vector spaces.*

Idea of Proof

1. As $\mathcal{A}$ is commutative, its character set equals the set of extreme states of $\mathcal{A}$ (Lemma 4.122), and it is well-known that the extreme states of any C^*-algebra separate its elements. Thus, the Gelfand transform on $\mathcal{A}$ is an injective linear transformation.

2. As characters are states, $\Xi \in \mathcal{L}^+(\mathcal{A}; C_{w^*}(X(\mathcal{A}); \mathbb{C}))$. In particular, it is a self-conjugate linear transformation. Moreover, by the multiplicativity of characters, for all $A, A' \in \mathcal{A}$, one has

$$\Xi\left(AA'\right) = \Xi(A)\,\Xi\left(A'\right) .$$

Hence, Ξ is an injective $*$-homomorphism.
3. By using Theorem 4.76 (ii) and the fact that all extreme states of a commutative unital C^*-algebra are characters (Lemma 4.122), we show that Ξ is bipositive.
4. As algebra elements separate states and the Gelfand transform of the unit of the algebra is the function which is constant equal 1, by the Stone-Weierstrass theorem (Theorem 7.191), the Gelfand transform has a dense image in $C_{w^*}(X(\mathcal{A}); \mathbb{C})$.
5. Now, recalling from Theorem 4.87 that the image under a $*$-homomorphism of C^*-algebra is a C^*-subalgebra (and is, hence, closed), it follows that for any commutative unital C^*-algebra $\mathcal{A}$, Ξ is surjective, and, thus, the Gelfand transform defines a $*$-isomorphism $\mathcal{A} \to C_{w^*}(X(\mathcal{A}); \mathbb{C})$.

□

Observe that the above proposition can be extended to the non-separable case. See, for instance, [51, Theorem 2.1.11A]. Recall, however, that, in this case, the weak* topology is not metrizable and thus the approach of Sect. 4.5.1 is not adequate anymore.

To close the present section, we prove a version of Proposition 4.124 for the non-unital case. In fact, it is a simple corollary of the proposition applied to unitizations of non-unital C^*-algebras together with the following lemma.

Lemma 4.125 *Let $\mathcal{A}$ be any separable commutative non-unital C^*-algebra and $\tilde{\mathcal{A}}$ its unitization. The Gelfand transform $\Xi : \tilde{\mathcal{A}} \to C_{w^*}(X(\tilde{\mathcal{A}}); \mathbb{C})$ maps $\mathcal{A} \subseteq \tilde{\mathcal{A}}$ onto the subspace of $C_{w^*}(X(\tilde{\mathcal{A}}); \mathbb{C})$ vanishing at the character $\rho_0 \in X(\tilde{\mathcal{A}})$ defined by Example 4.74.*

Proof Given any C^*-algebra $\mathcal{A}$ and its unitization $\tilde{\mathcal{A}} \doteq \mathbb{C} \times \mathcal{A}$, the character $\rho_0 \in X(\tilde{\mathcal{A}})$, as defined by Example 4.74, is the mapping $(\alpha, A) \mapsto \alpha$ from $\tilde{\mathcal{A}}$ to $\tilde{\mathcal{A}}/\mathcal{A} \equiv \mathbb{C}$. In other words, for all $(\alpha, A) \in \tilde{\mathcal{A}}$, $\rho_0((\alpha, A)) = \alpha$. Thus, for any $A \equiv (0, A) \in \mathcal{A} \subseteq \tilde{\mathcal{A}}$,

$$[\Xi(A)](\rho_0) = \rho_0(A) = 0 .$$

Hence,

$$\Xi(\mathcal{A}) \subseteq \{f \in C_{w^*}(X(\tilde{\mathcal{A}}); \mathbb{C}) \;:\; f(\rho_0) = 0\} .$$

On the other hand, if $[\Xi((\alpha, A))](\rho_0) = \alpha = 0$ then one has that $A \equiv (\alpha, A) \in \mathcal{A}$ and, observing that $\Xi(\tilde{\mathcal{A}}) = C_{w^*}(X(\tilde{\mathcal{A}}); \mathbb{C})$ (Proposition 4.124), the above inclusion is in fact an identity of subspaces of $C_{w^*}(X(\tilde{\mathcal{A}}); \mathbb{C})$. □

Corollary 4.126 *Let $\mathcal{A}$ be any (not necessarily unital) separable commutative C^*-algebra. The set of characters $X(\mathcal{A})$ is locally compact with respect to the weak* topology, and the range $\Xi(\mathcal{A}) \subseteq \mathcal{F}(X(\mathcal{A}); \mathbb{C})$ of the Gelfand transform Ξ is the space $C_{w^*,0}(X(\mathcal{A}); \mathbb{C})$ of weak*-continuous functions $X(\mathcal{A}) \to \mathbb{C}$ decaying at infinity (see Definition 7.166). Additionally, $\Xi : \mathcal{A} \to C_{w^*,0}(X(\mathcal{A}); \mathbb{C})$ is a $*$-isomorphism of C^*-algebras.*

Proof If $\mathcal{A}$ is unital, then $X(\mathcal{A})$ is weak*-compact, by Exercise 4.119. Hence, by Definition 7.166,

$$C_{w^*,0}(X(\mathcal{A}); \mathbb{C}) = C_{w^*}(X(\mathcal{A}); \mathbb{C})$$

and the corollary trivially follows from Proposition 4.124 in this special case. Thus, assume that $\mathcal{A}$ is non-unital. Recall that any character (state) of $\mathcal{A}$ uniquely extends to a character (state) of $\tilde{\mathcal{A}}$, the unitization of $\mathcal{A}$. See Exercise 7.64. Thus, $X(\mathcal{A})$ $(E(\mathcal{A}))$ can be canonically identified with a subset of $X(\tilde{\mathcal{A}})$ $(E(\tilde{\mathcal{A}}))$. Additionally, this identification is clearly compatible with the corresponding weak* topologies. Note that the character $\rho_0 \in X(\tilde{\mathcal{A}})$, defined by Example 4.74, is not the extension of a character of $\mathcal{A}$, as $\rho_0(\mathcal{A}) = \{0\}$. Thus, $\rho_0 \notin X(\mathcal{A}) \subseteq X(\tilde{\mathcal{A}})$. On the other hand, as $\rho_0((\alpha, A)) = \alpha$ for $(\alpha, A) \in \tilde{\mathcal{A}}$, if $\rho \in X(\tilde{\mathcal{A}})$ is different from ρ_0, then $\rho(A) \neq 0$ for some $A \in \mathcal{A}$. Hence, in this case, ρ is the unique extension of some character of $\mathcal{A}$, as the restriction of ρ to $\mathcal{A}$ is non-zero, linear, self-conjugate, and multiplicative. See Definition 4.117. From this we conclude that $X(\mathcal{A}) = X(\tilde{\mathcal{A}})\backslash\{\rho_0\}$. As $X(\tilde{\mathcal{A}})$ is a metrizable compact space, it follows that $X(\mathcal{A})$ is locally compact, as any closed ball of $X(\tilde{\mathcal{A}})\backslash\{\rho_0\}$ is compact for any metric associated with the weak* topology of $X(\tilde{\mathcal{A}})$. With this observation the corollary then follows from Proposition 4.124 combined with Lemma 4.125. □

Note that the last corollary is reminiscent of Corollary 4.77. Remark also the analogy between Propositions 4.100 and 4.124, on the one hand, and Corollaries 4.101 and 4.126, on the other hand.

Exactly as in the unital case, Corollary 4.126 can be extended to the non-separable case. See, for instance, [51, Theorem 2.1.11A]. As before, in this case, the weak* topology is not metrizable and, as a consequence, the approach of Sect. 4.5.1 is again not adequate.

4.8 The Universal C^*-Algebra of a Family of Polynomial Relations

Let $\mathcal{A}$ be a complex unital algebra. For an arbitrary complex polynomial

$$\mathcal{P}(x_1, \ldots, x_n) = c_0 + \sum_{M=1}^{N} \sum_{\pi:\{1,\ldots,M\}\to\{1,\ldots,n\}} c_\pi x_{\pi(1)} \cdots x_{\pi(M)} \, ,$$

$$c_0, c_\pi \in \mathbb{C} \, , \ \ N \in \mathbb{N} \, ,$$

of $n \in \mathbb{N}$ *non-commuting* variables, we define, for any $A_1, \ldots, A_n \in \mathcal{A}$, the element

$$\mathcal{P}(A_1, \ldots, A_n) \doteq c_0 \mathfrak{1} + \sum_{M=1}^{N} \sum_{\pi:\{1,\ldots,M\}\to\{1,\ldots,n\}} c_\pi A_{\pi(1)} \cdots A_{\pi(M)} \in \mathcal{A} .$$

If $\mathcal{A}$ is a *non-unital* complex algebra, then we restrict the above definition of $\mathcal{P}(A_1, \ldots, A_n)$ to the special case of polynomials $\mathcal{P}$ for which the constant term c_0 is zero. This notation is used to define polynomial relations and its C^*-representations.

Definition 4.127 (C^*-Representation of a Family of Polynomial Relations) Let I be a nonempty (index) set and $\mathcal{A}$ any (not necessarily unital) C^*-algebra:

(i) "Polynomial relations" for I are four-tuples $(\mathcal{P}, \bar{J}, J, \eta)$, where $\eta \in \mathbb{R}_0^+$, J and $\bar{J}$ are respectively mappings $\{1, \ldots, n_J\} \to I$ and $\{1, \ldots, n_{\bar{J}}\} \to I$, $n_J, n_{\bar{J}} \in \mathbb{N}_0$, $n_J + n_{\bar{J}} \geq 1$, and $\mathcal{P}$ is a complex polynomial in $n_J + n_{\bar{J}}$ non-commuting variables.

(ii) Given a mapping $a : I \to \mathcal{A}$, the pair $(\mathcal{A}, a)$ is a " C^*-representation" of a family $\mathfrak{R} = \{(\mathcal{P}_\omega, \bar{J}_\omega, J_\omega, \eta_\omega)\}_{\omega\in\Omega}$ of polynomial relations for I if, for all $\omega \in \Omega$,

$$\left\| \mathcal{P}_\omega(a \circ \bar{J}_\omega(1)^*, \ldots, a \circ \bar{J}_\omega(n_{\bar{J}_\omega})^*, a \circ J_\omega(1), \ldots, a \circ J_\omega(n_{J_\omega})) \right\| \leq \eta_\omega .$$

For every subset $I' \subseteq I$, $\mathfrak{R}_{I'}$ denotes the (sub)family

$$\mathfrak{R}_{I'} \doteq \{(\mathcal{P}_\omega, \bar{J}_\omega, J_\omega, \eta_\omega) \ : \omega \in \Omega,\ J_\omega(\{1, \ldots, n_{J_\omega}\}),$$
$$\bar{J}_\omega(\{1, \ldots, n_{\bar{J}_\omega}\}) \subseteq I'\} \subseteq \mathfrak{R}$$

of polynomial relations. In other words, $\mathfrak{R}_{I'}$ is the subfamily of polynomial relations only referring to points in the subset $I' \subseteq I$. In particular, $\mathfrak{R}_I = \mathfrak{R}$. Clearly, if $(\mathcal{A}, a)$ is a C^*-representation of $\mathfrak{R}$, then $(\mathcal{A}, a|_{I'})$ is a C^*-representation of $\mathfrak{R}_{I'}$. Conversely, if $I' \supseteq I$ is a bigger index set than the original one I, then $\mathfrak{R}= \mathfrak{R}_I$ is canonically seen as a family of polynomial relations for I'. (Note, in this case, that $a|_{I'\backslash I}$ is arbitrary, for there is no condition imposed on the elements $a(i) \in \mathcal{A}$, $i \in I'\backslash I$, by some polynomial relation.)

(iii) We say that the above family $\mathfrak{R}$ of relations is "admissible" if it has at least one C^*-representation.

(iv) Let $(\mathcal{A}, a)$ be a C^*-representation of $\mathfrak{R}$. Let $(\mathcal{A}', a')$ be a C^*-representation of $\mathfrak{R}'$, a second family of polynomial relations for a set I'. We say that the $*$-homomorphism $\Theta' : \mathcal{A} \to \mathcal{A}'$ is a "morphism of C^*-representations" from $(\mathcal{A}, a)$ to $(\mathcal{A}', a')$, if $I \subseteq I'$ and Θ' satisfies $\Theta' \circ a = a'|_I$. Note in particular that, by Proposition 4.97, if there is such a morphism, then $(\mathcal{A}', a'|_I)$ is necessarily a C^*-representation of $\mathfrak{R}$.

(v) The C^*-representation $(\mathcal{A}, a)$ of $\mathfrak{R}$ is called "universal" if, for any other C^*-representation $(\mathcal{A}', a')$ of $\mathfrak{R}$, there is a *unique* morphism $\Theta' : \mathcal{A} \to \mathcal{A}'$ of C^*-representations.

(vi) We say that the family $\mathfrak{R}$ of polynomial relations is "unital" if, for any C^*-representation $(\mathcal{A}, a)$ of $\mathfrak{R}$, $\mathcal{A}$ is a unital C^*-algebra and every morphism of C^*-representations of $\mathfrak{R}$ is a unital $*$-homomorphism.

(vii) We call here a family $\mathfrak{R}$ of polynomial relations "simple" if it has a universal C^*-representation corresponding to a simple[17] C^*-algebra.

Frequently, one writes polynomial relations in a less formal way than explicitly giving four-tuples $(\mathcal{P}, \bar{J}, J, \eta)$. For example, instead of writing $(\mathcal{P}, \bar{J}, J, \eta)$ with $\mathcal{P}(\bar{x}, x) = \bar{x} - x$, $\bar{J}, J : \{1\} \to \{i_0\}$, $\eta = 0$, one simply writes the following more intuitive expression: $a(i_0)^* = a(i_0)$. Sometimes, for simplicity, we will say "C^*-representation" instead of (the full term) "C^*-representation of polynomial relations."

Observe that the C^*-representations of polynomial relations can be seen as the objects of a category[18] denoted here by $C^*{-}\mathrm{Rep}$, the arrows of which are the morphisms of C^*-representations. For a given family $\mathfrak{R}$ of polynomial relations, considering the subcategory $C^*{-}\mathrm{Rep}_{\mathfrak{R}}$ of all C^*-representations of $\mathfrak{R}$ along with all the corresponding morphisms of C^*-representations, universal C^*-representations are exactly the initial objects[19] of this (sub)category. It is well-known that two initial objects of any category are necessarily isomorphic, in the sense of the given category. For a proof of this fact, as well as a comprehensive and fairly complete introduction to the theory of categories, we recommend [10]. In the category $C^*{-}\mathrm{Rep}_{\mathfrak{R}}$ of C^*-representations of a given family $\mathfrak{R}$ of polynomial relations, these categorial isomorphisms correspond to usual $*$-isomorphisms of C^*-algebras.

Proposition 4.128 *Let $\mathfrak{R}$ be an admissible family of polynomial relations for the index set I. Let $(\mathcal{A}, a)$ be a universal C^*-representation of $\mathfrak{R}$.*

(i) *For all $i \in I$,*

$$\begin{aligned}
&\|a(i)\|_{\mathcal{A}} \\
&\quad = \sup\{\|a'(i)\|_{\mathcal{A}'} : (\mathcal{A}', a') \text{ a } C^*\text{-representation of } \mathfrak{R}\} \\
&\quad = \sup\{\|a'(i)\|_{\mathcal{A}'} : (\mathcal{A}', a') \text{ a } C^*\text{-repr. of } \mathfrak{R} \\
&\qquad \text{for some unital } C^*\text{-algebra } \mathcal{A}'\} \\
&\quad = \sup\{\|a'(i)\|_{\mathrm{op}} : (\mathcal{B}(H), a') \text{ a } C^*\text{-repr. of } \mathfrak{R} \text{ for some Hilbert space } H\}\,.
\end{aligned}$$

[17] Recall that a normed algebra $\mathcal{A}$ is "simple" if $\{0\}$ and $\mathcal{A}$ are the only closed ideals of $\mathcal{A}$.

[18] A category $\mathcal{C}$ consists of (i) a collection of objects A, B, C,..., and (ii) a collection of arrows, or morphisms, $f, g, h : A \to B, \ldots$, including an identity $1_A : A \to A$ for every object A, for which, for objects A, B, C and every $f : A \to B$, $g : B \to C$ and $h : C \to D$, we have $h \circ (g \circ f) = (h \circ g) \circ f$ (associative law) and $f \circ 1_A = f = 1_B \circ f$ (unit).

[19] An initial object of a category $\mathcal{C}$ is an object I such that, for every object O in $\mathcal{C}$, there is exactly one arrow $I \to O$.

(ii) *If $(\mathcal{A}', a')$ is a second universal C^*-representation of $\mathfrak{R}$, there is a $*$-isomorphism $\Theta' : \mathcal{A} \to \mathcal{A}'$ such that $a' = \Theta' \circ a$. This $*$-isomorphism is the unique $*$-homomorphism $\Theta' : \mathcal{A} \to \mathcal{A}'$ satisfying $a' = \Theta' \circ a$. In particular, if $\mathfrak{R}$ has a universal C^*-representation on a simple C^*-algebra, i.e., $\mathfrak{R}$ is simple, then all universal C^*-representations of $\mathfrak{R}$ correspond to simple C^*-algebras.*

Proof

1. Note that the bound $\|a'(i)\|_{\mathcal{A}'} \leq \|a(i)\|_{\mathcal{A}}$, $i \in I$, is a consequence of Proposition 4.97. From this inequality it immediately follows that

$$\|a(i)\|_{\mathcal{A}} = \sup\{\|a'(i)\|_{\mathcal{A}'} \ : \ (\mathcal{A}', a') \text{ a } C^*\text{-representation of } \mathfrak{R}\} .$$

The second equality for $\|a(i)\|_{\mathcal{A}}$ follows from Proposition 4.16, and the third one from Theorem 4.89.
2. Assume that $(\mathcal{A}_1, a_1)$ and $(\mathcal{A}_2, a_2)$ are two universal C^*-representations of the family $\mathfrak{R}$ of relations. By definition of universal C^*-representations, there are unique $*$-homomorphisms $\Theta_{21} : \mathcal{A}_1 \to \mathcal{A}_2$ and $\Theta_{12} : \mathcal{A}_2 \to \mathcal{A}_1$ such that $a_1 = \Theta_{12} \circ a_2$ and $a_2 = \Theta_{21} \circ a_1$. In particular, $(\Theta_{21} \circ \Theta_{12}) \circ a_2 = a_2$ and $(\Theta_{12} \circ \Theta_{21}) \circ a_1 = a_1$.
3. On the other hand, obviously, $\mathrm{id}_{\mathcal{A}_1} \circ a_1 = a_1$ and $\mathrm{id}_{\mathcal{A}_2} \circ a_2 = a_2$. Again by the definition of a universal C^*-representation, one has $\Theta_{12} \circ \Theta_{21} = \mathrm{id}_{\mathcal{A}_1}$ and $\Theta_{21} \circ \Theta_{12} = \mathrm{id}_{\mathcal{A}_2}$. In other words, Θ_{12} and Θ_{21} are $*$-isomorphisms of the C^*-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$.

□

An arbitrary universal C^*-representation of a family $\mathfrak{R}$ of polynomial relations for the index set I is denoted here by $(C^*(I, \mathfrak{R}), a)$. Because of the properties of universal C^*-representations given in the above proposition, $C^*(I, \mathfrak{R})$, which is unique up to a $*$-isomorphism, is called the "universal C^*-algebra" associated with $(I, \mathfrak{R})$.

Exercise 4.129 Show that $\mathfrak{R} = \{a(i_0) = a(i_0)^*, \|a(i_0)\| \leq 1\}$ is an admissible family of polynomial relations for $I = \{i_0\}$ and that $(C_0([-1, 0) \cup (0, 1]; \mathbb{C}), a)$, with

$$a(i_0) = \mathrm{id}_{[-1,0)\cup(0,1]} \in C_0([-1, 0) \cup (0, 1]; \mathbb{C}) ,$$

is a universal C^*-representation of $\mathfrak{R}$. (In particular, $C^*(I, \mathfrak{R})$ is $*$-isomorphic to the non-unital C^*-algebra $C_0([-1, 0) \cup (0, 1]; \mathbb{C})$ of Example 4.7.)

In the context of universal algebras, we now discuss the notion of generation of C^*-algebras.

Definition 4.130 Let $\mathcal{A}$ be any C^*-algebra and $\mathcal{G} \subseteq \mathcal{A}$ a nonempty subset. The smallest C^*-subalgebra of $\mathcal{A}$ which contains $\mathcal{G}$ is denoted by $C^*(\mathcal{G}) \subseteq \mathcal{A}$. We say that $\mathcal{G}$ "generates" $\mathcal{A}$ if $C^*(\mathcal{G}) = \mathcal{A}$.

Note that, for any fixed C^*-algebra, the intersection of any family of C^*-subalgebra of $\mathcal{A}$ is itself a C^*-subalgebra of $\mathcal{A}$, since any intersection of closed sets in a normed space is closed. Thus, for any nonempty $\mathcal{G} \subseteq \mathcal{A}$, $C^*(\mathcal{G}) \subseteq \mathcal{A}$ exists. In fact, it is the intersection of all C^*-subalgebras of $\mathcal{A}$ containing $\mathcal{G}$.

Exercise 4.131 Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be C^*-algebras and $\mathcal{G}_1 \subseteq \mathcal{A}_1$ a set of generators of $\mathcal{A}_1$, i.e., a set $\mathcal{G}_1 \subseteq \mathcal{A}_1$ such that $\mathcal{A}_1 = C^*(\mathcal{G}_1)$. Let $\Theta, \Theta' : \mathcal{A}_1 \to \mathcal{A}_2$ be two $*$-homomorphisms. Show that $\Theta = \Theta'$ iff $\Theta(A_1) = \Theta'(A_1)$ for all $A_1 \in \mathcal{G}_1$.

The following corollary of Proposition 4.128 plays an important role in various constructions later on.

Corollary 4.132 *Let $(C^*(I, \mathfrak{R}), a)$ be a universal C^*-representation of the family $\mathfrak{R}$ of polynomial relations, as in Definition 4.127. Then*

$$C^*(I, \mathfrak{R}) = C^*(a(I)) .$$

Proof By definition of the generated C^*-algebra $C^*(a(I))$, we trivially have $a(I) \subseteq C^*(a(I))$ and thus observe that the pair $(C^*(a(I)), a)$ is a C^*-representation of $\mathfrak{R}$. Let $(\mathcal{A}', a')$ by any C^*-representation of $\mathfrak{R}$. By the universal property (Definition 4.127 (v)), there is a unique $*$-homomorphism $\Theta : C^*(I, \mathfrak{R}) \to \mathcal{A}'$ such that $a' = \Theta \circ a$. Let the $*$-homomorphism $\Theta' : C^*(a(I)) \to \mathcal{A}'$ be the restriction of Θ to $C^*(a(I))$. Then, by construction, $a' = \Theta' \circ a$. But, if $\Theta'' : C^*(a(I)) \to \mathcal{A}'$ is a second $*$-homomorphism such that $a' = \Theta' \circ a$, then one has $\Theta'' = \Theta'$, thanks to Exercise 4.131. Hence, Θ' is the unique $*$-homomorphism $C^*(a(I)) \to \mathcal{A}'$ satisfying $a' = \Theta' \circ a$ and thus $(C^*(a(I)), a)$ is a universal C^*-representation of $\mathfrak{R}$. By Proposition 4.128, the inclusion mapping $\mathfrak{i} : C^*(a(I)) \to C^*(I, \mathfrak{R})$, which is trivially a $*$-homomorphism satisfying $a' = \mathfrak{i} \circ a$, must be a $*$-isomorphism. In particular, $\mathfrak{i}$ is surjective and, thus, $C^*(a(I)) = C^*(I, \mathfrak{R})$. □

Corollary 4.133 *Let $(C^*(I, \mathfrak{R}), a)$ be a universal C^*-representation of the family $\mathfrak{R}$ of polynomial relations and $(\mathcal{A}', a')$ an arbitrary C^*-representation of a second family $\mathfrak{R}'$ of polynomial relations. There is at most one morphism $\Theta' : C^*(I, \mathfrak{R}) \to \mathcal{A}'$ of C^*-representations. Such a morphism exists iff $(\mathcal{A}', a'|_I)$ is a C^*-representation of $\mathfrak{R}$.*

Proof Let $\Theta'_1, \Theta'_2 : C^*(I, \mathfrak{R}) \to \mathcal{A}'$ be two morphisms of C^*-representations. By definition of such morphisms (Definition 4.127 (iv)), $I \subseteq I'$ and, for all $i \in I$,

$$\Theta'_1(a(i)) = \Theta'_2(a(i)) = a'(i) .$$

Thus, by Exercise 4.131 and Corollary 4.132, it follows that $\Theta'_1 = \Theta'_2$. The second part of the corollary directly follows from the definition of universal C^*-representations (Definition 4.127 (v)) and Proposition 4.97. □

The next theorem gives a necessary and sufficient condition for the existence of a universal C^*-representation of a given family of relations.

Theorem 4.134 *Let $\mathfrak{R}$ be a family of polynomial relations for the index set I. There is a universal C^*-representation of $\mathfrak{R}$ iff the following properties hold true:*

(i) *$\mathfrak{R}$ is admissible;*
(ii) *For all $i \in I$, $\sup\{\|a(i)\|_{\mathcal{A}} : (\mathcal{A}, a) \text{ a } C^*\text{-representation of } \mathfrak{R}\} < \infty$.*

Proof

1. Observe that we have already shown the necessity of the conditions (i) and (ii) for the existence of a universal C^*-representation of $\mathfrak{R}$. See Proposition 4.128 (i). Suppose now that $\mathfrak{R}$ is admissible and (ii) holds true.
2. Let $(\mathcal{A}, a)$ be any C^*-representation of $\mathfrak{R}$. By Exercise 7.67, the mapping $a : I \to \mathcal{A}$ uniquely extends to a unital $*$-homomorphism $\tilde{\mathfrak{F}}^*(I) \to \tilde{\mathcal{A}}$, also denoted by a, where $\tilde{\mathcal{A}}$ is the unitization of $\mathcal{A}$ while $\tilde{\mathfrak{F}}^*(I)$ is the unital free $*$-algebra generated by I. See Definition 7.65. By (ii) and the fact that $\mathcal{A}$ is a normed $*$-algebra, for all $F \in \tilde{\mathfrak{F}}^*(I)$,

$$\|F\| \doteq \sup\{\|a(F)\|_{\mathcal{A}} : (\mathcal{A}, a) \text{ a } C^*\text{-representation of } \mathfrak{R}\} < \infty \,.$$

It is easy to check that $(\tilde{\mathfrak{F}}^*(I), \|\cdot\|)$ is a seminormed $*$-algebra for which

$$\left\|F^* F\right\| = \|F\|^2 \,, \qquad F \in \tilde{\mathfrak{F}}^*(I) \,.$$

3. Let

$$\tilde{\mathfrak{F}}^*_0(I) \doteq \{F \in \tilde{\mathfrak{F}}^*(I) \;:\; \|F\| = 0\} \subseteq \tilde{\mathfrak{F}}^*(I) \,.$$

By Exercise 7.79, $\tilde{\mathfrak{F}}^*_0(I)$ is a $*$-ideal of $\tilde{\mathfrak{F}}^*(I)$ and the quotient $\tilde{\mathfrak{F}}^*(I)/\tilde{\mathfrak{F}}^*_0(I)$ is a normed $*$-algebra. By the corresponding property of the seminorm of $\tilde{\mathfrak{F}}^*(I)$,

$$\left\|[F]^*[F]\right\| = \left\|[F^* F]\right\| = \left\|F^* F\right\| = \|F\|^2 = \|[F]\|^2 \,, \qquad F \in \tilde{\mathfrak{F}}^*(I) \,,$$

and, thus, by Exercise 7.90, any completion of the quotient $\tilde{\mathfrak{F}}^*(I)/\tilde{\mathfrak{F}}^*_0(I)$ is a C^*-algebra (Exercise 7.90). We will show that any such a completion $(\mathcal{X}, \mathfrak{i})$ has a C^*-subalgebra which is a universal C^*-algebra associated with $(I, \mathfrak{R})$.
4. With this aim, let $(\mathcal{A}, a)$ be a C^*-representation associated with $\mathfrak{R}$. Then, as already explained, by Exercise 7.67, the mapping $a : I \to \mathcal{A}$ uniquely extends to a unital $*$-homomorphism $\tilde{\mathfrak{F}}^*(I) \to \tilde{\mathcal{A}}$, where $\tilde{\mathcal{A}}$ is the unitization of $\mathcal{A}$, also denoted by a. Note that, by construction, a is a contraction, i.e., $\|a\|_{\text{op}} \leq 1$, where $\tilde{\mathfrak{F}}^*(I)$ is endowed with the seminorm defined in 2. while $\tilde{\mathcal{A}}$ keeps of course its natural norm as a C^*-algebra. See Definition 7.36. Again by construction, one has that $\tilde{\mathfrak{F}}^*_0(I) \subseteq \ker(a)$. Thus, by Exercise 7.38, there is a unique contraction $\tilde{\mathfrak{F}}^*(I)/\tilde{\mathfrak{F}}^*_0(I) \to \tilde{\mathcal{A}}$, also denoted by a, for which

$$a([F]) = a(F) \,, \qquad F \in \tilde{\mathfrak{F}}^*(I) \,.$$

See Exercise 7.8. By Exercise 7.73 (ii), this new mapping is a $*$-homomorphism from $\tilde{\mathfrak{F}}^*(I)/\tilde{\mathfrak{F}}_0^*(I)$ to $\tilde{\mathcal{A}}$. Its unique continuous extension to some fixed completion $(\mathcal{X}, \mathfrak{i})$ of $\tilde{\mathfrak{F}}^*(I)/\tilde{\mathfrak{F}}_0^*(I)$ is denoted by Θ. Note that it is a $*$-homomorphism $\mathcal{X} \to \tilde{\mathcal{A}}$ satisfying

$$a(i) = \Theta \circ \mathfrak{i}([i])) , \qquad i \in I .$$

By Exercise 4.131, the restriction of Θ to $C^*(\mathfrak{i}([I]))$ is the unique $*$-homomorphism $C^*(\mathfrak{i}([I])) \to \mathcal{A} \subseteq \tilde{\mathcal{A}}$ satisfying this equation. This proves that

$$(C^*(\mathfrak{i}([I])), \mathfrak{i}([\cdot]))$$

is a universal C^*-representation associated with $\mathfrak{R}$.

□

For more details on universal C^*-representations of polynomial relations, see [67, Chapter II.8.3] and [68]. It is possible to show that any unital C^*-algebra is the universal C^*-algebra of some family of relations. However, the definition of a C^*-algebra via a family of relations is generally only useful if the relations have a simple presentation.

In some situations, one may construct C^*-representations, also universal ones, of families $\mathfrak{R}$ of polynomial relations as limits of C^*-representations of families of polynomial relations, which are smaller than $\mathfrak{R}$. This kind of construction will be used later on in specific important cases, and we introduce here some general notions and facts related to it.

Definition 4.135 (Directed Systems of C^*-Representations) Let $\mathfrak{D}$ be any directed (with respect to the inclusion) family of nonempty sets. (See Definition 1.13.) For all $\Lambda \in \mathfrak{D}$, let $\mathfrak{R}_\Lambda$ be a family of polynomial relations for the (index) set Λ and $(\mathcal{A}_\Lambda, a_\Lambda)$ a C^*-representation of $\mathfrak{R}_\Lambda$. Further, for all $\Lambda, \Lambda' \in \mathfrak{D}$ with $\Lambda \supseteq \Lambda'$, let $i_{\Lambda\Lambda'}$ be a morphism of C^*-representations, from $(\mathcal{A}_{\Lambda'}, a_{\Lambda'})$ to $(\mathcal{A}_\Lambda, a_\Lambda)$ (in particular, $(\mathcal{A}_\Lambda, i_{\Lambda\Lambda'} \circ a_\Lambda)$ is a C^*-representation of $\mathfrak{R}_{\Lambda'}$):

(i) The family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ of C^*-representations, along with the family $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\Lambda\supseteq\Lambda'}$ of morphisms of C^*-representations, is called a "directed system of C^*-representations" if, for all $\Lambda, \Lambda', \Lambda'' \in \mathfrak{D}$ with $\Lambda \supseteq \Lambda' \supseteq \Lambda''$ one has $i_{\Lambda\Lambda'} \circ i_{\Lambda'\Lambda''} = i_{\Lambda\Lambda''}$, $i_{\Lambda\Lambda} = \mathrm{id}_{\mathcal{A}_\Lambda}$.
(ii) The directed system is "faithful" if the $*$-homomorphisms $i_{\Lambda\Lambda'}$, $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$ are all faithful (i.e., injective).
(iii) The directed system is "unital" if, for every $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, $\mathcal{A}_\Lambda$ is a unital C^*-algebra and $i_{\Lambda\Lambda'}$ is a unital $*$-homomorphism.

Note that, in categorial terms, seeing $\mathfrak{D}$ as a so-called poset category (i.e., the objects of the category are the elements of $\mathfrak{D}$ and, for arbitrary $\Lambda, \Lambda' \in \mathfrak{D}$, there is (at most) one arrow $\Lambda' \to \Lambda$ if $\Lambda \supseteq \Lambda'$), directed systems of C^*-representations are exactly the covariant functors[20] $\mathfrak{D} \to C^*\text{−Rep}$, mapping $\Lambda \in \mathfrak{D}$ to $(\mathcal{A}_\Lambda, a_\Lambda)$ and the (unique) arrow $\Lambda' \to \Lambda$ to $i_{\Lambda\Lambda'}$. By virtue of Corollary 4.133, in the special case of universal C^*-representations, we may equivalently define directed systems of C^*-representations in a simpler way.

Definition 4.136 (Directed Systems of Universal C^*-Representations) Let $\mathfrak{D}$ be any directed family of nonempty sets. For all $\Lambda \in \mathfrak{D}$ let $\mathfrak{R}_\Lambda$ be a family of polynomial relations for the set Λ. The family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$, where, for every $\Lambda \in \mathfrak{D}$, $(\mathcal{A}_\Lambda, a_\Lambda)$ is a *universal* C^*-representation of $\mathfrak{R}_\Lambda$, is called a "directed system of universal C^*-representations" if, for all $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, $(\mathcal{A}_\Lambda, a_\Lambda|_{\Lambda'})$ is a C^*-representation of $\mathfrak{R}_{\Lambda'}$. For all $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, the unique morphism $\mathcal{A}_{\Lambda'} \to \mathcal{A}_\Lambda$ of C^*-representations is denoted by $i_{\Lambda\Lambda'}$.

Clearly, if the family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ defines a directed system of universal C^*-representations, then this family, along with the family $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\, \Lambda\supseteq\Lambda'}$ of unique morphisms $i_{\Lambda\Lambda'} : \mathcal{A}_{\Lambda'} \to \mathcal{A}_\Lambda$ of C^*-representations, corresponds to a directed system of C^*-representations in the primary sense.

Recall that if $\mathfrak{R}_\Lambda$ is a family of polynomial relations for a given (index) set Λ and Λ' is a second set with $\Lambda' \supseteq \Lambda$, then $\mathfrak{R}_\Lambda$ is canonically seen as a family of polynomial relations for the bigger set Λ'. Here, for any directed family $\mathfrak{D}$ of nonempty sets and a collection $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, of families of polynomial relations, $\mathfrak{R}_{\cup\mathfrak{D}}$ denotes the family of polynomial relations which is the union of all families $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, seen as families of polynomial relations for the index set $\cup\mathfrak{D}$.

Definition 4.137 (Cocones Associated with C^*-Representations) Let $\mathfrak{D}$ be any directed family of nonempty sets, and $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ together with $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\, \Lambda\supseteq\Lambda'}$ be a directed system of C^*-representations referring to a collection $\{\mathfrak{R}_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of families of polynomial relations:

(i) The C^*-representation $(\mathcal{A}, a)$ of $\mathfrak{R}_{\cup\mathfrak{D}}$, along with a family $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of morphisms $\mathcal{A}_\Lambda \to \mathcal{A}$ of C^*-representations, is a "cocone" associated with the given directed system of C^*-representations, if, for all $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, one has that $\Theta_\Lambda \circ i_{\Lambda\Lambda'} = \Theta_{\Lambda'}$. The C^*-representation $(\mathcal{A}, a)$ is called the "vertex" of the cocone, in this case.

(ii) For two of such cocones, $(\mathcal{A}, a)$ together with $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ and $(\mathcal{A}', a')$ together with $\{\Theta'_\Lambda\}_{\Lambda\in\mathfrak{D}}$, we say that the morphism of C^*-representations $\Theta' : \mathcal{A} \to \mathcal{A}'$ is a "morphism of cocones" if, for every $\Lambda \in \mathfrak{D}$, $\Theta'_\Lambda = \Theta' \circ \Theta_\Lambda$.

(iii) The cocone given by $(\mathcal{A}, a)$ together with $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ is the "inductive limit" of the corresponding directed system of C^*-representations, if, for any second cocone, given by the vertex $(\mathcal{A}', a')$ together with the family $\{\Theta'_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of morphisms of C^*-representations, there is a *unique* morphism of cocones $\Theta' : \mathcal{A} \to \mathcal{A}'$.

[20] In the category theory, a functor F is a transformation between two categories $\mathcal{C}_1$ and $\mathcal{C}_2$ mapping objects of $\mathcal{C}_1$ to objects of $\mathcal{C}_2$ and morphisms in $\mathcal{C}_1$ to morphisms in $\mathcal{C}_2$. It is covariant when the directions of arrows are preserved, i.e., $A \to B$ in $\mathcal{C}_1$ yields $F(A) \to F(B)$ in $\mathcal{C}_2$.

(iv) The cocone is “faithful” if every $*$-homomorphism Θ_Λ, $\Lambda \in \mathfrak{D}$, defining it is faithful.
(v) The cocone is “unital” if $\mathcal{A}$ is a unital C^*-algebra, and, for every $\Lambda \in \mathfrak{D}$, $\mathcal{A}_\Lambda$ and Θ_Λ are respectively a unital C^*-algebra and a unital $*$-homomorphism.

The following proposition is the analogue of Proposition 4.128 (ii) for cocones of directed systems of C^*-representations.

Proposition 4.138 *Let $(\mathcal{A}, a)$ together with $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ and $(\mathcal{A}', a')$ together with $\{\Theta'_\Lambda\}_{\Lambda\in\mathfrak{D}}$ be two cocones for the same direct system of C^*-representations. If both cocones are inductive limits, then the unique morphism $\mathcal{A} \to \mathcal{A}'$ of cocones (which is by definition a $*$-homomorphism) is a $*$-isomorphism. Thus, up to a $*$-isomorphism, a direct system of C^*-representations has at most one directed limit.*

Proof The proposition is proven in a similar way as Proposition 4.128 (ii). We omit the details. □

In fact, similar to universal C^*-representations, inductive limits of a direct system of C^*-representations are initial objects of a category of cocones associated with this system, and, again, the above proposition refers to the general categorial fact that initial objects are unique up to isomorphism, already mentioned above in the context of (the category of) C^*-representations. The arrows of this category are exactly the morphisms of cocones, as defined in Definition 4.137 (ii).

Similar to Definition 4.136, in the case of *universal* directed systems of C^*-representations, the associated cocones have a simpler definition. In fact, they can be identified with C^*-representations of $\mathfrak{R}_{\cup\mathfrak{D}}$.

Definition 4.139 (Cocones Associated with Universal C^*-Representations) Given a directed family $\mathfrak{D}$ of nonempty sets, let $\mathfrak{R}_\Lambda$ be a family of polynomial relations for every (index) set $\Lambda \in \mathfrak{D}$. Let the family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ be a directed system of *universal* C^*-representations of the polynomial relations $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$. We call a C^*-representation $(\mathcal{A}, a)$ a “cocone associated with the given directed system of universal C^*-representations” if, for all $\Lambda \in \mathfrak{D}$, $(\mathcal{A}, a|_\Lambda)$ is a C^*-representation of $\mathfrak{R}_\Lambda$. For all $\Lambda \in \mathfrak{D}$, Θ_Λ denotes the unique morphism $\mathcal{A}_\Lambda \to \mathcal{A}$ of C^*-representations.

Given a family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ representing a directed system of *universal* C^*-representations and any cocone $(\mathcal{A}, a)$ in the above sense, notice that the C^*-representation $(\mathcal{A}, a)$, along with the family $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of morphisms of C^*-representations as in the last definition, corresponds to a cocone associated with a directed system of C^*-representations, in the primary sense (Definition 4.137 (i)). In fact, it is the unique cocone whose vertex is $(\mathcal{A}, a)$.

Exercise 4.140 Given a directed family $\mathfrak{D}$ of nonempty sets, let $\mathfrak{R}_\Lambda$ be a family of polynomial relations for every (index) set $\Lambda \in \mathfrak{D}$. Let the family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ be a directed system of *universal* C^*-representations for the polynomial relations $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$. Show that the C^*-representation $(\mathcal{A}, a)$ for $\mathfrak{R}_{\cup\mathfrak{D}}$ is universal iff it is a inductive limit associated with the directed system of universal C^*-representations.

The following corollary is the analogue of Corollary 4.132 for C^*-representations associated with cocones.

Corollary 4.141 *Given a directed family $\mathfrak{D}$ of nonempty sets, let $\mathfrak{R}_\Lambda$ be a family of polynomial relations for every (index) set $\Lambda \in \mathfrak{D}$. Let $(\mathcal{A}, a)$ be the the vertex of a cocone associated with a directed system C^*-representations of $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$. If the cocone is an inductive limit, then*

$$\mathcal{A} = C^*(a(\cup\mathfrak{D})) .$$

Conversely, the vertex $(\mathcal{A}, a)$ is a universal C^*-representation of the family $\mathfrak{R}_{\cup\mathfrak{D}}$ of polynomial relations only if the corresponding cocone is an inductive limit.

Proof The first part of the corollary is proven, mutatis mutandis, in the same way as Corollary 4.132, using Proposition 4.138 instead of Proposition 4.128. The second part is a direct consequence of the definitions of universal C^*-representations and inductive limits of cocones for directed systems of C^*-representations. We omit the details. □

From the second part of the corollary, if one wishes to construct a universal C^*-representation by means of a cocone associated with some directed system of C^*-representations, then the cocone has to be an inductive limit. Later on we will construct some important universal C^*-representations exactly in this way. Observe, however, that this does not mean that the vertices of directed limits are generally universal C^*-representations. By Exercise 4.140, the vertex of any inductive limit of directed system of *universal* C^*-representations is a universal C^*-representation.

An important application of the first part of Corollary 4.141 is the following proposition saying that families of "compatible" states of the C^*-algebras of a directed system of C^*-representations uniquely define states of the C^*-algebra of any faithful inductive limit of the directed system.

Definition 4.142 (Compatible Families of States of a Directed System of C^*-Representations) Let $\mathfrak{D}$ be any directed family of nonempty sets, and $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ together with $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\, \Lambda\supseteq\Lambda'}$ be a directed system of C^*-representations. Let $\rho_\Lambda \in E(\mathcal{A}_\Lambda)$ be some state on the C^*-algebra $\mathcal{A}_\Lambda$ for each $\Lambda \in \mathfrak{D}$. We say that the family $\{\rho_\Lambda\}_{\Lambda\in\mathfrak{D}}$ is a "compatible family of states for the directed system" if

$$\rho_\Lambda \circ i_{\Lambda\Lambda'} = \rho_{\Lambda'} , \qquad \Lambda, \Lambda' \in \mathfrak{D}, \Lambda \supseteq \Lambda' .$$

Proposition 4.143 (States of Inductive Limits via Compatible Families of States) *Let $\mathfrak{D}$ be any directed family of nonempty sets and $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ together with $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\, \Lambda\supseteq\Lambda'}$ be a directed system of C^*-representations. Let $\{\rho_\Lambda\}_{\Lambda\in\mathfrak{D}}$ be any family of states $\rho_\Lambda \in E(\mathcal{A}_\Lambda)$ that is compatible for the directed system of C^*-representations. Consider an arbitrary inductive limit of the directed system of C^*-representations, which is given by a C^*-representation $(\mathcal{A}, a)$ together with a family $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of morphisms $\mathcal{A}_\Lambda \to \mathcal{A}$ of C^*-representations. If the inductive* *limit (as a cocone) is faithful, then there is a unique state $\rho \in E(\mathcal{A})$ of the C^*-algebra satisfying $\rho \circ \Theta_\Lambda = \rho_\Lambda$ for all $\Lambda \in \mathfrak{D}$.*

Proof Note first that, for all $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, $\Theta_\Lambda(\mathcal{A}_\Lambda)$ and $\Theta_{\Lambda'}(\mathcal{A}_{\Lambda'})$ are $*$-subalgebras of $\mathcal{A}$, and $\Theta_\Lambda(\mathcal{A}_\Lambda) \supseteq \Theta_{\Lambda'}(\mathcal{A}_{\Lambda'})$. This is so, because

$$\Theta_{\Lambda'}(\mathcal{A}_{\Lambda'}) = \Theta_\Lambda \circ i_{\Lambda\Lambda'}(\mathcal{A}_{\Lambda'}) \qquad \text{and} \qquad i_{\Lambda\Lambda'}(\mathcal{A}_{\Lambda'}) \subseteq \mathcal{A}_\Lambda \ .$$

See Definitions 4.135 (i) and 4.137 (i). Let

$$\mathcal{A}_0 \doteq \bigcup_{\Lambda \in \mathfrak{D}} \Theta_\Lambda(\mathcal{A}) \subseteq \mathcal{A} \ .$$

Observe that $\mathcal{A}_0$ is again a $*$-subalgebra of $\mathcal{A}$. Moreover, if the cocone is an inductive limit, then $\mathcal{A}_0$ is dense in $\mathcal{A}$, by Corollary 4.141. There is a unique (complex) linear functional $\tilde{\rho}$ on $\mathcal{A}_0$ such that

$$\tilde{\rho}(\Theta_\Lambda(A)) = \rho_\Lambda(A) \ , \qquad \Lambda \in \mathfrak{D} \ , \ A \in \mathcal{A}_\Lambda \ . \tag{4.6}$$

The uniqueness and linearity of such a functional is trivial, provided it is well-defined. We have in fact to keep in mind that, for all $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, $\Theta_\Lambda(\mathcal{A}_\Lambda) \supseteq \Theta_{\Lambda'}(\mathcal{A}_{\Lambda'})$, and we thus need to ensure that

$$\tilde{\rho}(\Theta_\Lambda(A)) = \tilde{\rho}(\Theta_{\Lambda'}(A'))$$

for any $A' \in \mathcal{A}_{\Lambda'}$ and $A \in \mathcal{A}_\Lambda$ so that $\Theta_{\Lambda'}(A') = \Theta_\Lambda(A)$. To this end, note first that, if the cocone is faithful, then, for any $\Lambda \in \mathfrak{D}$ and $A_1, A_2 \in \mathcal{A}_\Lambda$ with $\Theta_\Lambda(A_1) = \Theta_\Lambda(A_2)$ one has that $A_1 = A_2$ and, hence, $\rho_\Lambda(A_1) = \rho_\Lambda(A_2)$. Now, for any $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, and any $A' \in \mathcal{A}_{\Lambda'}$ and $A \in \mathcal{A}_\Lambda$ with $\Theta_{\Lambda'}(A') = \Theta_\Lambda(A)$, one has that

$$\Theta_{\Lambda'}(A') = \Theta_\Lambda(i_{\Lambda\Lambda'}(A')) = \Theta_\Lambda(A) \ .$$

By compatibility of the family of states with the directed system, it follows that

$$\tilde{\rho}(\Theta_\Lambda(A)) \doteq \rho_\Lambda(A) = \rho_\Lambda(i_{\Lambda\Lambda'}(A')) = \rho_{\Lambda'}(A') \doteq \tilde{\rho}(\Theta_{\Lambda'}(A')) \ .$$

As a consequence, Eq. (4.6) yields a unique (complex) linear functional $\tilde{\rho}$ on $\mathcal{A}_0$. By Proposition 4.97, the (operator) norm of $\tilde{\rho}$ as a linear functional on the subspace $\mathcal{A}_0$ of the normed space $\mathcal{A}$ is exactly 1. Additionally, $\tilde{\rho}$ is positive. By density of $\mathcal{A}_0$, and recalling that any positive element of $\mathcal{A}$ has the form A^*A for some $A \in \mathcal{A}$ (Corollary 4.103), it follows that $\tilde{\rho}$ uniquely extends to a state on $\mathcal{A}$. ☐

In the next proposition we show that *faithful* directed systems of C^*-representations always have an inductive limit, which is again faithful:

Proposition 4.144 *Any faithful directed system of C^*-representations has an inductive limit that is faithful (as a cocone).*

Proof

1. Let $\mathfrak{D}$ be any directed family of nonempty sets and $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ together with $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\, \Lambda\supseteq\Lambda'}$ be a directed system of C^*-representations of families $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, of polynomial relations. Define the set

$$\mathfrak{X} \doteq \{(\Lambda, A_\Lambda) \ : \ \Lambda \in \mathfrak{D} \ , \ A_\Lambda \in \mathcal{A}_\Lambda \ \} \ ,$$

i.e., $\mathfrak{X}$ is the disjoint union $\sqcup_{\Lambda\in\mathfrak{D}}\mathcal{A}_\Lambda$ of all the C^*-algebras $\mathcal{A}_\Lambda$, $\Lambda \in \mathfrak{D}$. We say that two elements (Λ, A_Λ) and $(\Lambda', A_{\Lambda'})$ of this set are equivalent if for some $\Lambda'' \in \mathfrak{D}$, $\Lambda'' \subseteq \Lambda, \Lambda'$ and $A_{\Lambda''} \in \mathcal{A}_{\Lambda''}$, one has that

$$A_\Lambda = i_{\Lambda\Lambda''}(A_{\Lambda''}) \qquad \text{and} \qquad A_{\Lambda'} = i_{\Lambda'\Lambda''}(A_{\Lambda''}) \ .$$

This clearly defines an equivalence relation in $\mathfrak{X}$, by the definition of directed systems of C^*-representations (Definition 4.135). Let $\mathcal{B}$ be the corresponding set of equivalence classes. For all $\Lambda \in \mathfrak{D}$, we define the mapping $\Theta_\Lambda : \mathcal{A}_\Lambda \to \mathcal{B}$ by

$$\Theta_\Lambda(A) \doteq [(\Lambda, A)] \ , \qquad A \in \mathcal{A}_\Lambda \ .$$

Additionally, we define a mapping $b : \cup\mathfrak{D} \to \mathcal{B}$ by the condition

$$b(x) = [(\Lambda, a_\Lambda(x))] \ , \qquad \Lambda \in \mathfrak{D}, \ x \in \Lambda \ .$$

Observe that this mapping is well-defined, as for any $\Lambda, \Lambda' \in \mathfrak{D}$ and $x \in \Lambda \cap \Lambda' \doteq \Lambda''$,

$$a_\Lambda(x) = i_{\Lambda\Lambda''}(a_{\Lambda''}(x)) \qquad \text{and} \qquad a_{\Lambda'}(x) = i_{\Lambda'\Lambda''}(a_{\Lambda''}(x))) \ ,$$

by definition of directed system of C^*-representations.
2. Again by the definition of directed systems of C^*-representations, there are mappings (operations) $+ : \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ (sum), $\circ : \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ (product), scalar multiplication $\cdot : \mathbb{C} \times \mathcal{B} \to \mathcal{B}$ and complex conjugation $(\cdot)^* : \mathcal{B} \to \mathcal{B}$ that are uniquely defined by the following conditions:

$$[(\Lambda, A)] + [(\Lambda, A')] = [(\Lambda, A + A')] \ ,$$
$$[(\Lambda, A)] \circ [(\Lambda, A')] = [(\Lambda, AA')] \ ,$$
$$\alpha \cdot [(\Lambda, A)] = [(\Lambda, \alpha A)] \ ,$$
$$[(\Lambda, A)]^* = [(\Lambda, A^*)]$$

for any $\Lambda \in \mathfrak{D}$, all $A, A' \in \mathcal{A}_\Lambda$ and $\alpha \in \mathbb{C}$. $(\mathcal{B}, +, \cdot, \circ, ^*)$ is clearly a $*$-algebra and the mappings $\Theta_\Lambda : \mathcal{A}_\Lambda \to \mathcal{B}$, $\Lambda \in \mathfrak{D}$, are $*$-homomorphisms.

3. If the directed system of C^*-representations is faithful (Definition 4.135 (ii)), then

$$\|[(\Lambda, A)]\| \doteq \inf\{\|B\| \ : \ B \in [(\Lambda, A)]\}$$

define the (unique) norm on $\mathcal{B}$ such that

$$\|[(\Lambda, A)]\| = \|A\|$$

for any $\Lambda \in \mathfrak{D}$ and all $A \in \mathcal{A}_\Lambda$, thanks to Proposition 4.97. $(\mathcal{B}, +, \cdot, \circ, ^*, \|\cdot\|)$ is clearly a normed $*$-algebra for which $\|A^*A\| = \|A\|^2$ for all $A \in \mathcal{B}$. Additionally, all the mappings $\Theta_\Lambda : \mathcal{A}_\Lambda \to \mathcal{B}$, $\Lambda \in \mathfrak{D}$, are norm-preserving. Thus, for any completion $(\mathcal{A}, \mathfrak{i})$ of this normed $*$-algebra, $\mathcal{A}$ is a C^*-algebra (Exercise 7.90) and the mappings $\mathfrak{i} \circ \Theta_\Lambda : \mathcal{A}_\Lambda \to \mathcal{A}$, $\Lambda \in \mathfrak{D}$, are faithful $*$-homomorphisms; see Definition 7.88. Let $a \doteq \mathfrak{i} \circ b$. It is easy to check that, for all $\Lambda \in \mathfrak{D}$, $(\mathcal{A}, a|_\Lambda)$ is a C^*-representation of the family $\mathfrak{R}_\Lambda$ of polynomial relations and $(\mathcal{A}, a)$ is the C^*-representation of $\mathfrak{R}_{\cup\mathfrak{D}}$. Recall that $\mathfrak{R}_{\cup\mathfrak{D}}$ denotes the family of polynomial relations which is the union of all families $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, seen as families of polynomial relations for the index set $\cup\mathfrak{D}$. Clearly, the C^*-representation $(\mathcal{A}, a)$ of $\mathfrak{R}_{\cup\mathfrak{D}}$, along with the family $\{\mathfrak{i} \circ \Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of $*$-homomorphisms, defines a faithful cocone (Definition 4.137) associated with the directed system of C^*-representations.
4. Take any cocone associated with the directed system of C^*-representations, which is given by a C^*-representation $(\mathcal{A}', a')$ together with a family $\{\Theta'_\Lambda\}_{\Lambda\in\mathfrak{D}}$ of morphisms $\mathcal{A}_\Lambda \to \mathcal{A}'$ of C^*-representations. Notice that there is a unique mapping $\Theta' : \mathfrak{i}(\mathcal{B}) \to \mathcal{A}'$ satisfying

$$\Theta'(\mathfrak{i} \circ \Theta_\Lambda(A)) = \Theta'_\Lambda(A) \ , \qquad \Lambda \in \mathfrak{D}, \ A \in \mathcal{A}_\Lambda \ .$$

The uniqueness of Θ' is trivial. Its existence follows from the fact that, $\mathfrak{i} \circ \Theta_\Lambda$ being faithful, $\mathfrak{i} \circ \Theta_\Lambda(A)$ uniquely determines $A \in \mathcal{A}_\Lambda$ and, for any $\Lambda, \Lambda' \in \mathfrak{D}$ such that $\Lambda' \supseteq \Lambda$,

$$\Theta'_{\Lambda'}(i_{\Lambda'\Lambda}(A)) = \Theta'_\Lambda(A) \ , \qquad A \in \mathcal{A}_\Lambda \ .$$

In other words, the definition of Θ' does not depend on the representatives of equivalence classes, which are the elements of $\mathcal{B}$, and Θ' is thus well-defined. Additionally, this mapping is clearly a norm-preserving $*$-homomorphism. Thus, it uniquely extends to a $*$-homomorphism $\Theta' : \mathcal{A} \to \mathcal{A}'$. It is easy to check that it is a morphism of cocones, i.e., for every $\Lambda \in \mathfrak{D}$, $\Theta'_\Lambda = \Theta' \circ \mathfrak{i} \circ \Theta_\Lambda$ (Definition 4.137 (ii)).
5. On the other hand, if $\tilde{\Theta} : \mathcal{A} \to \mathcal{A}'$ is a morphism of cocones, then, for all $\Lambda \in \mathfrak{D}$ and all $A \in \mathcal{A}_\Lambda$, one has $\tilde{\Theta}(\Theta_\Lambda(A)) = \Theta'_\Lambda(A)$. But this implies that $\tilde{\Theta} = \Theta'$, because these two $*$-homomorphisms are equal to each other in the

dense subspace $\mathfrak{i}(\mathcal{B}) \subseteq \mathcal{A}$. Thus, Θ' defines an inductive limit for the directed system of C^*-representations, by virtue of Definition 4.137 (iii).

□

From Proposition 4.144 we therefore arrive at the following conclusion about the admissibility of the union of families of polynomial relations.

Corollary 4.145 *Let $\mathfrak{D}$ be any directed family of nonempty sets and $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, a collection of families of polynomial relations. If there is a* faithful *directed system of C^*-representations of $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, then the family $\mathfrak{R}_{\cup\mathfrak{D}}$ of polynomial relations (see discussions prior to Definition* 4.137*) is admissible.*

For collections of simple families of polynomial relations, one has the following stronger version of the above corollary, which is a kind of "compactness" result for simple unital families of polynomial relations.

Proposition 4.146 (Compactness of Simple Unital Families of Polynomal Relations) *Let I be any nonempty index set and $\mathfrak{R}$ a family of polynomial relations for I. Let $\mathfrak{D}$ be the directed set of all nonempty* finite *subsets of I. If, for every $\Lambda \in \mathfrak{D}$, $\mathfrak{R}_\Lambda$ is a family of polynomial relations that is simple and unital (see Definition* 4.127 (vi)–(vii)*, then (the full family) $\mathfrak{R}=\mathfrak{R}_{\cup\mathfrak{D}}$ is simple and unital.*

Proof

1. Assume that, for all $\Lambda \in \mathfrak{D}$, $\mathfrak{R}_\Lambda$ is a family of polynomial relations that is simple and unital. Take, for every $\Lambda \in \mathfrak{D}$, a universal C^*-representation $(\mathcal{A}_\Lambda, a_\Lambda)$ of $\mathfrak{R}_\Lambda$. Note that, by assumption and Proposition 4.128, $\mathcal{A}_\Lambda$ is a simple unital C^*-algebra. As $\mathfrak{R}_\Lambda$, $\Lambda \in \mathfrak{D}$, are restrictions of the same bigger family $\mathfrak{R}$ of polynomial relations (see Definition 4.127 (ii)), one has that, for all $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda' \subseteq \Lambda$, $(\mathcal{A}_\Lambda, a_\Lambda|_{\Lambda'})$ is a C^*-representation of $\mathfrak{R}_{\Lambda'}$. Thus, by universality, the family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ of C^*-representations defines a unique directed system of C^*-representations. See Definition 4.136 and discussions afterward.
2. Recalling that the kernels of $*$-homomorphisms of C^*-algebras are closed ideals of these algebras (see, e.g., Exercise 7.70), by simplicity of $\mathcal{A}_\Lambda$, $\Lambda \in \mathfrak{D}$, the morphisms of C^*-representations $i_{\Lambda\Lambda'} : \mathcal{A}_{\Lambda'} \to \mathcal{A}_\Lambda$ for $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$, are either trivial (i.e., $i_{\Lambda\Lambda'} = 0$) or faithful. As the family $\mathfrak{R}_\Lambda$ of polynomial relations is unital (see Definition 4.127 (vi)), $i_{\Lambda\Lambda'}$ is thus faithful and unital. By Proposition 4.144, the directed system of C^*-representation has an inductive limit. Let $(\mathcal{A}, a)$ denote its vertex as a cocone. By Exercise 4.140, $(\mathcal{A}, a)$ is a universal C^*-representation of $\mathfrak{R}=\mathfrak{R}_{\cup\mathfrak{D}}$.
3. For all $\Lambda \in \mathfrak{D}$, let $\Theta_\Lambda : \mathcal{A}_\Lambda \to \mathcal{A}$ be the unique morphism of C^*-representations. Note that these $*$-homomorphisms of C^*-algebras are faithful, again by simplicity of $\mathcal{A}_\Lambda$ and the fact that the family $\mathfrak{R}_\Lambda$ of polynomial relations is unital. Thus, for all $\Lambda \in \mathfrak{D}$, $\mathcal{B}_\Lambda \doteq \Theta_\Lambda(\mathcal{A}_\Lambda) \subseteq \mathcal{A}$ are simple unital C^*-subalgebras of $\mathcal{A}$. Note that the family $\{\mathcal{B}_\Lambda\}_{\Lambda\in\mathfrak{D}}$ is an increasing net of (simple) C^*-subalgebras of $\mathcal{A}$, as a direct consequence of the definition of cocones associated with directed systems of C^*-representations (Definition 4.137 (i)). By

Corollary 4.141, the union $\cup\{\mathcal{B}_\Lambda : \Lambda \in \mathfrak{D}\}$ is a dense $*$-subalgebra of $\mathcal{A}$. Thus, by Proposition 4.57, $\mathcal{A}$ is a simple.

4. $\mathcal{A}$ is a unital C^*-algebra, because, for any $\Lambda \in \mathfrak{D}$, $(\mathcal{A}, a|_\Lambda)$ is a C^*-representation of $\mathfrak{R}_\Lambda$, this family of polynomial relations being unital. Let $(\mathcal{A}', a')$ and $(\mathcal{A}'', a'')$ be any two C^*-representations of $\mathfrak{R}=\mathfrak{R}_{\cup\mathfrak{D}}$, and $\Theta' : \mathcal{A}' \to \mathcal{A}''$ any morphism of C^*-representations of $\mathfrak{R}$. Then, for any $\Lambda \in \mathfrak{D}$, $(\mathcal{A}', a'|_\Lambda)$ and $(\mathcal{A}'', a''|_\Lambda)$ are C^*-representations of $\mathfrak{R}_\Lambda$, and Θ' is a morphism of C^*-representations of $\mathfrak{R}_\Lambda$. As $\mathfrak{R}_\Lambda$ is a family of polynomial relations that is unital, Θ' is a unital $*$-homomorphism. Thus, by Definition 4.127 (vi), the family $\mathfrak{R}=\mathfrak{R}_{\cup\mathfrak{D}}$ of polynomial relations is unital.

□

In the following paragraphs, we introduce classes of C^*-algebras, which are central in quantum statistical mechanics, as classes of universal algebras of families of polynomial relations. The last corollary and last proposition turn out to be particularly useful to construct these algebras.

4.8.1 Universal Tensor Products of Unital C-Algebras*

We start with a class of universal C^*-algebras named here "universal tensor products" of unital C^*-algebras. In quantum statistical mechanics, important examples of such algebras are the "spin algebras," which refer to the observables of quantum spin systems.

Definition 4.147 (Universal Tensor Products of Unital C^*-Algebras) Let Ω be a nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be some arbitrary unital C^*-algebra. Define the index set:

$$I_\Omega \doteq \{(\omega, A_\omega) \ : \ \omega \in \Omega \ , \ A_\omega \in \mathcal{A}_\omega\}$$

(i.e., I_Ω is the disjoint union $\sqcup_{\omega\in\Omega}\mathcal{A}_\omega$ of all the C^*-algebras $\mathcal{A}_\omega$, $\omega \in \Omega$), as well as the following family of polynomial relations for I_Ω:

$$\begin{aligned}\mathfrak{R} \doteq \{&a(\omega, \alpha A_\omega + A'_\omega) = \alpha a(A_\omega, x) + a(A'_\omega, x) \ : \ A_\omega, \\ &A'_\omega \in \mathcal{A}_\omega, \ \alpha \in \mathbb{C}, \ \omega \in \Omega\} \\ \cup\{&a(\omega, A_\omega)a(\omega', A_{\omega'}) = a(\omega', A_{\omega'})a(\omega, A_\omega) \ : \ A_\omega \in \mathcal{A}_\omega, \ A_{\omega'} \in \mathcal{A}_{\omega'}, \ \omega, \\ &\omega' \in \Omega, \ \omega \neq \omega'\} \\ \cup\{&a(\omega, A_\omega A'_\omega) = a(\omega, A_\omega)a(\omega, A'_\omega) \ : \ A_\omega, A'_\omega \in \mathcal{A}_\omega, \ \omega \in \Omega\} \\ \cup\{&a(\omega, \mathfrak{1}_\omega) = \mathfrak{1} \ : \ \omega \in \Omega\} \cup \{a(\omega, A^*_\omega) \\ &= a(\omega, A_\omega)^* \ : \ A_\omega \in \mathcal{A}_\omega, \ \omega \in \Omega\} \ ,\end{aligned}$$

a being a mapping from I_Ω to some C^*-algebra $\mathcal{A}$. Here, for any $\omega \in \Omega$, $\mathfrak{1}_\omega$ denotes the unit of $\mathcal{A}_\omega$ while $\mathfrak{1}$ is the one of $\mathcal{A}$. Then, we define the "universal tensor product" of the family $\mathcal{A}_\omega$, $\omega \in \Omega$, of unital C^*-algebras by

$$\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega \doteq (C^*(I_\Omega, \mathfrak{R}), a) ,$$

where $(C^*(I_\Omega, \mathfrak{R}), a)$ is any fixed universal C^*-representation of $\mathfrak{R}$. By a slight abuse of notation, sometimes $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ also denotes only the universal C^*-algebra $C^*(I_\Omega, \mathfrak{R})$.

Note that the above polynomial relations are just saying that, for all $\omega \in \Omega$, $a(\omega, \cdot)$ is a unital $*$-homomorphism $\mathcal{A}_\omega \to \mathcal{A}$ and that $a(\omega, A_\omega), a(\omega', A_{\omega'}) \in \mathcal{A}$ commute, whenever $\omega \neq \omega'$.

Exercise 4.148 Let Ω be a nonempty countable set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ some arbitrary separable unital C^*-algebra. Prove that the universal tensor product $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ is also a separable unital C^*-algebra.

Definition 4.147 assumes the existence of a universal C^*-representation for $\mathfrak{R}$. This requires some arguments: By Theorem 4.134, $\mathfrak{R}$ must be admissible, i.e., a C^*-representation $(\mathcal{A}, a)$ must exist, and, for all $\omega \in \Omega$ and $A \in \mathcal{A}_\omega$, one must have

$$\sup\{\|a(A, x)\|_{\mathcal{A}} \ : \ (\mathcal{A}, a) \text{ a } C^*\text{-representation of } \mathfrak{R}\} < \infty . \tag{4.7}$$

The last property can be easily verified: If $(\mathcal{A}, a)$ is any C^*-representation of the above-defined family $\mathfrak{R}$ of polynomial relations, then, for every fixed $\omega \in \Omega$, define the mapping $\Theta_\omega : \mathcal{A}_\omega \to \mathcal{A}$ by $\Theta_\omega(A) \doteq a(\omega, A)$ for all $A \in \mathcal{A}_\omega$. By the polynomial relations in $\mathfrak{R}$, Θ_ω is a unital $*$-homomorphism, and Lemma 4.96 (ii) implies that $\|a(\omega, A)\|_{\mathcal{A}} \leq \|A\|_{\mathcal{A}_\omega}$ for all $A \in \mathcal{A}_\omega$ and $\omega \in \Omega$. Hence, Eq. (4.7) holds true.

It remains to prove that the family of polynomial relations defining $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ is admissible:

Proposition 4.149 (The Tensor Product Relations Are Admissible and Unital) *Let Ω be any nonempty set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be some arbitrary unital C^*-algebra. Then, the family $\mathfrak{R}$ of polynomial relations defined in Definition 4.147 is admissible and unital.*

Proof The fact that $\mathfrak{R}$ is a unital family of polynomial relations is clear, because, for any C^*-representation $(\mathcal{A}, a)$ of $\mathfrak{R}$ and all $\omega \in \Omega$, one has that $a(\omega, \mathfrak{1}_\omega) = \mathfrak{1} \in \mathcal{A}$. See Definition 4.127 (vi). For all $\omega \in \Omega$, let (H_ω, π_ω) be any representation of the unital C^*-algebra $\mathcal{A}_\omega$. See Definition 4.88 (i). By Theorem 4.89 or Theorem 4.113, recall that any C^*-algebra always has a representation. Let be $\mathcal{P}_f(\Omega)$ be the set of *finite* subsets of Ω and define the directed set:

$$\mathfrak{D} \doteq \left\{ \{(\omega, A_\omega) \ : \ \omega \in \tilde{\Lambda} \ , \ A_\omega \in \mathcal{A}_\omega\} \ : \ \tilde{\Lambda} \in \mathcal{P}_f(\Omega)\backslash\{\emptyset\} \right\} .$$

That is, the elements of $\mathfrak{D}$ are the disjoint unions $\sqcup_{\omega\in\tilde{\Lambda}}\mathcal{A}_\omega$ of the C^*-algebras $\mathcal{A}_\omega$, $\omega \in \tilde{\Lambda}$, for any finite nonempty subset $\tilde{\Lambda} \subseteq \Omega$. From now the set $\tilde{\Lambda}$ stands for the projection of the set $\Lambda \in \mathfrak{D}$ over Ω, i.e.,

$$\tilde{\Lambda} \doteq \{\omega \in \Omega \; : \; (\omega, A_\omega) \in \Lambda \text{ for some } A_\omega \in \mathcal{A}_\omega\} \in \mathcal{P}_f(\Omega)\backslash\{\emptyset\} \; .$$

Assume that Ω is endowed with a total order. In fact, recall that Zermelo's theorem, also known as the well-ordering theorem, states that every set has even a well-ordering, which is, by definition, a particular case of a total order. For every $\Lambda \in \mathfrak{D}$, we define the unital C^*-algebra:

$$\mathcal{A}_\Lambda \doteq \mathcal{B}\left(\bigotimes_{\omega\in\tilde{\Lambda}} H_\omega\right) \; ,$$

where $\bigotimes_{\omega\in\tilde{\Lambda}} H_\omega$ denotes the Hilbert space that is the (Hilbert-Schmidt) tensor product ($\otimes_2$) of the Hilbert spaces H_ω, $\omega \in \tilde{\Lambda}$, ordered by the total order of Ω. See Sect. 7.3.8 for a self-contained exposition on (Hilbert-Schmidt) tensor products of Hilbert spaces. For any $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda' \subsetneq \Lambda$, we canonically identify the Hilbert spaces:

$$\bigotimes_{\omega\in\tilde{\Lambda}} H_\omega \qquad \text{and} \qquad \left(\bigotimes_{\omega\in\tilde{\Lambda}'} H_\omega\right) \otimes_2 \left(\bigotimes_{\omega\in\tilde{\Lambda}\backslash\tilde{\Lambda}'} H_\omega\right)$$

(see Corollary 7.261) and define the mapping $\mathfrak{i}_{\Lambda\Lambda'} : \mathcal{A}_{\Lambda'} \to \mathcal{A}_\Lambda$ by

$$\mathfrak{i}_{\Lambda\Lambda'}(A_\Lambda) \doteq A_\Lambda \otimes \mathrm{id}_{\bigotimes_{\omega\in\tilde{\Lambda}\backslash\tilde{\Lambda}'} H_\omega}$$

for any $\Lambda, \Lambda' \in \mathfrak{D}$, $\Lambda \supseteq \Lambda'$. Clearly, $\mathfrak{i}_{\Lambda\Lambda'}$ is a faithful unital $*$-homomorphism. Further, for all $\Lambda \in \mathfrak{D}$, we define a mapping $a_\Lambda : \Lambda \to \mathcal{A}_\Lambda$ by

$$a_\Lambda(\omega, A_\omega) \doteq \mathfrak{i}_{\Lambda(\{\omega\}\times\mathcal{A}_\omega)} \circ \pi_\omega(A_\omega) \; , \qquad (\omega, A_\omega) \in \Lambda \; .$$

It is each to check that, for all $\Lambda \in \mathfrak{D}$, $(\mathcal{A}_\Lambda, a_\Lambda)$ is a C^*-representation of $\mathfrak{R}_\Lambda$ and that the family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ of C^*-representations together with the family $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\,\Lambda\supseteq\Lambda'}$ of $*$-homomorphisms defines a faithful directed system of C^*-representations (Definition 4.135). By Corollary 4.145, the family $\mathfrak{R}{=}\mathfrak{R}_{\cup\mathfrak{D}}$ is admissible. □

From the universal property of universal tensor products of unital C^*-algebras, we define natural $*$-homomorphisms:

$$\bigotimes_{\omega\in\Omega'} \mathcal{A}_\omega \to \bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$$

for every nonempty subset $\Omega' \subseteq \Omega$:

Definition 4.150 (Canonical Inclusions of Universal Tensor Products) Let Ω be a nonempty set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be a unital C^*-algebra. For any nonempty subset $\Omega' \subseteq \Omega$, define the (index set)

$$I_{\Omega'} \doteq \{(\omega, A_\omega) \ : \ \omega \in \Omega' \ , \ A_\omega \in \mathcal{A}_\omega\} \ ,$$

that is, $I_{\Omega'}$ is the disjoint union $\sqcup_{\omega \in \Omega'} \mathcal{A}_\omega$ of the C^*-algebras $\mathcal{A}_\omega$, $\omega \in \Omega'$. Then, clearly, for any $\Omega' \subseteq \Omega$,

$$\left(\bigotimes\nolimits_{\omega \in \Omega} \mathcal{A}_\omega, a|_{I_{\Omega'}} \right)$$

is a C^*-representation of the family $\mathfrak{R}$ of polynomial relations of Definition 4.147. In particular, from the universal property of $\bigotimes_{\omega \in \Omega'} \mathcal{A}_\omega$, there is a unique morphism:

$$\bigotimes\nolimits_{\omega \in \Omega'} \mathcal{A}_\omega \to \bigotimes\nolimits_{\omega \in \Omega} \mathcal{A}_\omega$$

of C^*-representations, which is denoted here by $\mathfrak{i}_{\Omega\Omega'}$.

Notice that if, for all $\omega \in \Omega'$, $\mathcal{A}_\omega = \mathcal{L}(\mathbb{C}^{d(\omega)})$ for some $d(\omega) \in \mathbb{N}$, i.e., these C^*-algebras are matrix algebras, then $\mathfrak{i}_{\Omega\Omega'}$ is faithful, by Corollary 4.154. Thus, in this case $\bigotimes_{\omega \in \Omega'} \mathcal{A}_\omega$ is canonically seen as a C^*-subalgebra of $\bigotimes_{\omega \in \Omega} \mathcal{A}_\omega$.

From the existence of universal tensor products of unital C^*-algebras as universal C^*-algebras, we can easily make sense of tensor products of unital $*$-homomorphism of C^*-algebras.

Proposition 4.151 (Infinite Tensor Products of Unital $*$-Homomorphisms of C^*-Algebras) *Let Ω be a nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ and $\mathcal{A}'_\omega$ be two arbitrary unital C^*-algebras, and $\Theta_\omega : \mathcal{A}_\omega \to \mathcal{A}'_\omega$ a unital $*$-homomorphism. There exists a unique $*$-homomorphism*

$$\bigotimes\nolimits_{\omega \in \Omega} \Theta_\omega : \bigotimes\nolimits_{\omega \in \Omega} \mathcal{A}_\omega \to \bigotimes\nolimits_{\omega \in \Omega} \mathcal{A}'_\omega$$

such that, for all $\tilde{\omega} \in \Omega$ and $A_{\tilde{\omega}} \in \mathcal{A}_{\tilde{\omega}}$,

$$\left[\bigotimes\nolimits_{\omega \in \Omega} \Theta_\omega \right] (a(\tilde{\omega}, A_{\tilde{\omega}})) = a(\tilde{\omega}, \Theta_{\tilde{\omega}}(A_{\tilde{\omega}})) \ .$$

Proof Observe that the mapping $(\omega, A_\omega) \mapsto a(\omega, \Theta_\omega(A_\omega))$ from

$$I_\Omega \doteq \{(\omega, A_\omega) : \omega \in \Omega, A_\omega \in \mathcal{A}_\omega\}$$

to $\bigotimes_{\omega \in \Omega} \mathcal{A}'_\omega$ defines a C^*-representation of the tensor product relations for the family $\{\mathcal{A}_\omega\}_{\omega \in \Omega}$ of C^*-algebras. Thus, by the universality of $\bigotimes_{\omega \in \Omega} \mathcal{A}_\omega$, there is a unique $*$-homomorphism as stated in the proposition. See Definition 4.127 (iv–v).

□

In a similar way, we may introduce "permutation automorphisms" for tensor products of any fixed unital C^*-algebra.

Proposition 4.152 (Permutation Automorphisms) *Let Ω be a nonempty, possibly infinite set, and $\mathcal{A}$ some fixed unital C^*-algebra. For any bijection (permutation) $\pi : \Omega \to \Omega$, there is a unique $*$-automorphism τ_π of $\bigotimes_{\omega\in\Omega} \mathcal{A}$ such that, for all $\omega \in \Omega$ and $A \in \mathcal{A}$,*

$$\tau_\pi(a(\omega, A)) = a(\pi(\omega), A) . \tag{4.8}$$

For two bijections $\pi, \pi' : \Omega \to \Omega$, one additionally has the identity $\tau_\pi \circ \tau_{\pi'} = \tau_{\pi\circ\pi'}$.

Proof For any bijection $\pi : \Omega \to \Omega$, the existence of a unique $*$-homomorphism

$$\bigotimes_{\omega\in\Omega} \mathcal{A} \to \bigotimes_{\omega\in\Omega} \mathcal{A}$$

satisfying (4.8) is proven by means of an obvious adaptation of the proof of the last proposition, and we omit the details. Take now two bijections $\pi, \pi' : \Omega \to \Omega$. Then, for all $\omega \in \Omega$ and $A \in \mathcal{A}$,

$$\tau_\pi \circ \tau_{\pi'}(a(\omega, A)) = \tau_{\pi\circ\pi'}(a(\omega, A)) .$$

Hence, by Exercise 4.131 and Corollary 4.132, $\tau_\pi \circ \tau_{\pi'} = \tau_{\pi\circ\pi'}$. In particular, by choosing $\pi' = \pi^{-1}$ and then $\pi = \pi'^{-1}$, one concludes that τ_π is a $*$-isomorphism for any permutation π of Ω. □

Recall from Exercise 7.25 that, for any dimension $d \in \mathbb{N}$, the unital C^*-algebra $\mathcal{L}(\mathbb{C}^d)$ of $d\times d$ complex matrices[21] is simple. This property is preserved by universal tensor products.

Lemma 4.153 *Let Ω be any nonempty set and d any mapping $\Omega \to \mathbb{N}$. Then, the universal tensor product $\bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)})$ is a simple C^*-algebra. If the set Ω is finite, then $\bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)})$ is $*$-isomorphic to $\mathcal{L}(\mathbb{C}^{\prod_{\omega\in\Omega} d(\omega)})$.*

Proof Let $\mathcal{P}_f(\Omega)$ be the directed (with respect to the inclusion) set of all finite subsets of Ω. Observe that, for all $\Lambda = \{\omega_1, \ldots, \omega_{|\Lambda|}\} \in \mathcal{P}_f(\Omega)$, there is a unique unital $*$-homomorphism:

$$\Theta_\Lambda : \mathcal{L}(\mathbb{C}^{d(\omega_1)}) \otimes \cdots \otimes \mathcal{L}(\mathbb{C}^{d(\omega_{|\Lambda|})}) \to \bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)})$$

such that

$$\Theta_\Lambda(A_{\omega_1} \otimes \cdots \otimes A_{\omega_{|\Lambda|}}) = a(\omega_1, A_{\omega_1}) \cdots a(\omega_{|\Lambda|}, A_{\omega_{|\Lambda|}})$$

[21] The complex conjugation in this complex algebra is, by default, the usual operation of taking Hermitian conjugates of matrices.

for all $A_{\omega_k} \in \mathcal{L}(\mathbb{C}^{d(\omega_k)})$ and $k \in \{1, \ldots, |\Lambda|\}$. Recall from Proposition 7.62 that

$$\mathcal{L}(\mathbb{C}^{d(\omega_1)}) \otimes \cdots \otimes \mathcal{L}(\mathbb{C}^{d(\omega_{|\Lambda|})})$$

is a $*$-algebra. The uniqueness of Θ_Λ is clear. Its existence directly follows from the universal property of (algebraic) tensor products. See Proposition 7.10. Observe from Corollary 7.14 that $\mathcal{L}(\mathbb{C}^{d(\omega_1)})\otimes\cdots\otimes\mathcal{L}(\mathbb{C}^{d(\omega_{|\Lambda|})})$ and $\mathcal{L}(\mathbb{C}^{d(\omega_1)}\otimes\cdots\otimes\mathbb{C}^{d(\omega_{|\Lambda|})})$ are $*$-isomorphic $*$-algebras. Note also that $\mathcal{L}(\mathbb{C}^{d(\omega_1)+\cdots+d(\omega_{|\Lambda|})})$ and $\mathcal{L}(\mathbb{C}^{d(\omega_1)} \otimes \cdots \otimes \mathbb{C}^{d(\omega_{|\Lambda|})})$ are $*$-isomorphic, because the vector spaces $\mathbb{C}^{d(\omega_1)+\cdots+d(\omega_{|\Lambda|})}$ and $\mathbb{C}^{d(\omega_1)} \otimes \cdots \otimes \mathbb{C}^{d(\omega_{|\Lambda|})}$ have the same dimension, by virtue of Proposition 7.12. In particular, $\mathcal{L}(\mathbb{C}^{d(\omega_1)})\otimes\cdots\otimes\mathcal{L}(\mathbb{C}^{d(\omega_{|\Lambda|})})$ is a simple algebra, thanks to Exercise 7.25. Recall that the kernel of $*$-homomorphisms are always $*$-ideals (Exercise 7.70). Θ_Λ is not trivial, being a unital $*$-homomorphism, and, thus, it is necessarily faithful, i.e., injective. For every finite subset $\Lambda = \{\omega_1, \ldots, \omega_{|\Lambda|}\} \in \mathcal{P}_f(\Omega)\backslash\emptyset$, let

$$\mathcal{A}_\Lambda \doteq \Theta_\Lambda \left(\mathcal{L}(\mathbb{C}^{d(\omega_1)}) \otimes \cdots \otimes \mathcal{L}(\mathbb{C}^{d(\omega_{|\Lambda|})})\right) \subseteq \bigotimes\nolimits_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)}) \ .$$

By Theorem 4.87, $\{\mathcal{A}_\Lambda\}_{\Lambda\in\mathcal{P}_f(\Omega)}$ is an increasing net of simple unital C^*-subalgebras of the tensor product $\bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)})$. Note that we also use here the identification:

$$\mathcal{L}(\mathbb{C}^n) \equiv \mathcal{L}(\mathbb{C}^n) \otimes \mathfrak{1}_{\mathbb{C}^m} \subseteq \mathcal{L}(\mathbb{C}^n) \otimes \mathcal{L}(\mathbb{C}^m)$$

for any $n, m \in \mathbb{N}$, $\mathfrak{1}_{\mathbb{C}^m}$ being the unit of $\mathcal{L}(\mathbb{C}^m)$. Thus, by Proposition 4.57, one has to prove that the union

$$\mathcal{A}_\Omega \doteq \bigcup_{\Lambda\in\mathcal{P}_f(\Omega)} \mathcal{A}_\Lambda$$

is dense in the tensor product $\bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)})$. But, by construction, $\mathcal{A}_\Omega$ is a $*$-subalgebra for which

$$\{a(\omega, A_\omega) \ : \ A_\omega \in \mathcal{L}(\mathbb{C}^{d(\omega)}), \ \omega \in \Omega\} \subseteq \mathcal{A}_\Omega \ .$$

Thus, by Corollary 4.132, the closure of $\mathcal{A}_\Omega$ is the whole tensor product $\bigotimes_{\omega\in\Omega} \mathcal{L}(\mathbb{C}^{d(\omega)})$. The last part of the lemma is clear from the above arguments. □

From the above lemma, we obtain the following sufficient condition for the (canonical) $*$-homomorphisms of Definition 4.150 to be faithful.

Corollary 4.154 *Let Ω be a nonempty set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be an arbitrary unital C^*-algebra. For any $\Omega' \subseteq \Omega$ such that, for all $\omega \in \Omega'$, $\mathcal{A}_\omega = \mathcal{L}(\mathbb{C}^{d(\omega)})$ for some $d(\omega) \in \mathbb{N}$, the $*$-homomorphism*

$$\mathfrak{i}_{\Omega\Omega'} : \bigotimes\nolimits_{\omega\in\Omega'} \mathcal{A}_\omega \to \bigotimes\nolimits_{\omega\in\Omega} \mathcal{A}_\omega$$

of Definition 4.150 is faithful.

Proof Note that $\mathfrak{i}_{\Omega\Omega'}$ is nontrivial, because it is unital. Thus, it has to be faithful, $\bigotimes_{\omega\in\Omega'} \mathcal{A}_\omega$ being (by Lemma 4.153) a simple C^*-algebra. Recall that the kernel of any $*$-homomorphisms between C^*-algebras is a closed ideal. □

We now define an important class of states (Definition 4.61 (i)) on universal tensor products of unital C^*-algebras. It refers to the so-called product states.

Proposition 4.155 (Product States) *Let Ω be any nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be a unital C^*-algebra and $\rho_\omega \in E(\mathcal{A}_\omega)$ a state. There is a unique state on the tensor product, denoted by $\bigotimes_{\omega\in\Omega} \rho_\omega$, such that, for all $\tilde{\Lambda} \in \mathcal{P}_f(\Omega)\backslash\emptyset$ and $A_\omega \in \mathcal{A}_\omega$ with $\omega \in \tilde{\Lambda}$, one has*

$$\left[\bigotimes\nolimits_{\omega\in\Omega} \rho_\omega\right]\left(\prod_{\omega\in\tilde{\Lambda}} a(\omega, A_\omega)\right) = \prod_{\omega\in\tilde{\Lambda}} \rho_\omega(A_\omega) \ .$$

Proof From Corollary 4.132, the product property stated in the proposition uniquely defines the (product) state in a dense subspace of $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$, whence the product state has to be unique, if it exists. To show its existence, for every $\omega \in \Omega$, take any cyclic representation $(H_\omega, \pi_\omega, \Omega_\omega)$ of the C^*-algebra $\mathcal{A}_\omega$ associated with the state $\rho_\omega \in E(\mathcal{A}_\omega)$. Recall from Theorem 4.113 that such a representation always exists for each state on any C^*-algebra. See also Definition 4.88. Considering the representations (H_ω, π_ω) of the C^*-algebras $\mathcal{A}_\omega$, $\omega \in \Omega$, let the family $\{(\mathcal{A}_\Lambda, a_\Lambda)\}_{\Lambda\in\mathfrak{D}}$ of C^*-representations, as well as the family $\{i_{\Lambda\Lambda'}\}_{\Lambda,\Lambda'\in\mathfrak{D},\ \Lambda\supseteq\Lambda'}$ of $*$-homomorphisms, be defined exactly as in the proof of Proposition 4.149. Recall that they define a faithful directed system of C^*-representations. For all $\Lambda \in \mathfrak{D}$, define the state $\rho_\Lambda \in E(\mathcal{A}_\Lambda)$ by

$$\rho_\Lambda(A) \doteq \left\langle \bigotimes\nolimits_{\omega\in\tilde{\Lambda}} \Omega_\omega, A \bigotimes\nolimits_{\omega\in\tilde{\Lambda}} \Omega_\omega \right\rangle , \qquad A \in \mathcal{A}_\Lambda \ ,$$

i.e., ρ_Λ is the vector state on

$$\mathcal{A}_\Lambda \doteq \mathcal{B}\left(\bigotimes\nolimits_{\omega\in\tilde{\Lambda}} H_\omega\right)$$

associated with the unit vector

$$\bigotimes\nolimits_{\omega\in\tilde{\Lambda}} \Omega_\omega \in \bigotimes\nolimits_{\omega\in\tilde{\Lambda}} H_\omega \ . \tag{4.9}$$

Here, for all $\Lambda \in \mathfrak{D}$, $\tilde{\Lambda} \in \mathcal{P}_f(\Omega)$ and $\bigotimes_{\omega\in\tilde{\Lambda}} H_\omega$ are exactly defined as in the proof of Proposition 4.149. In particular, $\tilde{\Lambda} \in \mathcal{P}_f(\Omega)\backslash\{\emptyset\}$ stands for the projection of $\Lambda \in \mathfrak{D}$ over Ω. The vector (4.9) is then defined correspondingly. By construction, the family $\{\rho_\Lambda\}_{\Lambda\in\mathfrak{D}}$ is a compatible family of states for the directed system, in the sense of Definition 4.142. As the directed system of C^*-representations is faithful, it has an inductive limit that is a faithful cocone, thanks to Proposition 4.144. Let the C^*-representation $(\mathcal{A}, \tilde{a})$ denote the vertex of the inductive limit (as a cocone) and $\{\Theta_\Lambda\}_{\Lambda\in\mathfrak{D}}$ the corresponding family of morphisms of C^*-representations $\mathcal{A}_\Lambda \to \mathcal{A}$. Then, we invoke Proposition 4.143 to deduce the existence of a unique state $\tilde{\rho}$ on $\mathcal{A}$ satisfying $\tilde{\rho} \circ \Theta_\Lambda = \rho_\Lambda$ for all $\Lambda \in \mathfrak{D}$. Let

$$\Theta : \bigotimes_{\omega\in\Omega} \mathcal{A}_\omega \to \mathcal{A}$$

be the unique morphism of cocones. See Definition 4.137 (iii). Define the state $\rho \doteq \tilde{\rho} \circ \Theta$ of the tensor product. Note that ρ is a state because Θ is unital, the family of polynomial relations for the universal tensor product being unital. Then, by Definition 4.137 (ii), for all $\tilde{\Lambda} \in \mathcal{P}_f(\Omega)\backslash\emptyset$ and $A_\omega \in \mathcal{A}_\omega$ with $\omega \in \tilde{\Lambda}$,

$$\begin{aligned}
\rho\left(\prod_{\omega\in\tilde{\Lambda}} a(\omega, A_\omega)\right) &= \tilde{\rho} \circ \Theta\left(\prod_{\omega\in\tilde{\Lambda}} a(\omega, A_\omega)\right) \\
&= \rho_{\tilde{\Lambda}}\left(\prod_{\omega\in\tilde{\Lambda}} a_{\{\omega\}\times\mathcal{A}_\omega}(\omega, A_\omega)\right) \\
&= \prod_{\omega\in\tilde{\Lambda}} \rho_{\{\omega\}\times\mathcal{A}_\omega}\left(a_{\{\omega\}\times\mathcal{A}_\omega}(\omega, A_\omega)\right) \\
&= \prod_{\omega\in\tilde{\Lambda}} \rho_\omega(A_\omega)\ .
\end{aligned}$$

□

Corollary 4.156 *Let Ω be any nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be a unital C^*-algebra. For all $\tilde{\Lambda} \in \mathcal{P}_f(\Omega)\backslash\emptyset$ and $A_\omega \in \mathcal{A}_\omega$ with $\omega \in \tilde{\Lambda}$, one has the following identity of norms*

$$\left\|\prod_{\omega\in\Lambda} a(\omega, A_\omega)\right\| = \prod_{\omega\in\Lambda} \|a(\omega, A_\omega)\| = \prod_{\omega\in\Lambda} \|A_\omega\|\ .$$

In particular, for all $\omega \in \Omega$, $a(\omega, \cdot)$ is a faithful *$*$-homomorphism $\mathcal{A}_\omega \to \bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$.*

Proof From the identity $\|A^*A\| = \|A\|^2$ satisfied by the norm of any C^*-algebra, we may assume that, for any $\tilde{\Lambda} \in \mathcal{P}_f(\Omega)\backslash\emptyset$ and $\omega \in \tilde{\Lambda}$, the element $A_\omega \in \mathcal{A}_\omega$ and, consequently also $a(\omega, A_\omega)$ (Definition 4.147), is normal. Recalling that states are (by definition) norm-one linear functionals, the corollary directly follows from Propositions 4.78 (ii) and 4.155. □

To conclude this paragraph, we briefly discuss the relation of the above construction with the usual algebraic tensor product and general tensor products of C^*-algebras. For simplicity, we mainly consider twofold tensor products, but, of course, the discussion can be easily generalized for the tensor product of any finite family of C^*-algebras.

Lemma 4.157 *Let Ω be any nonempty, possibly infinite set. For every $\omega \in \Omega$, take some unital C^*-algebra $\mathcal{A}_\omega$. For any $\omega', \omega'' \in \Omega$, $\omega' \neq \omega''$, let $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ be the usual (algebraic) tensor product of $\mathcal{A}_{\omega'}$ and $\mathcal{A}_{\omega''}$ as complex vector spaces. Then, there is a unique linear transformation (inclusion mapping):*

$$\mathfrak{i}_{\Omega\{\omega',\omega''\}} \equiv \mathfrak{i}_{\{\omega',\omega''\}} : \mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''} \to \bigotimes\nolimits_{\omega\in\Omega} \mathcal{A}_\omega$$

such that, for all $A_{\omega'} \in \mathcal{A}_{\omega'}$ and $A_{\omega''} \in \mathcal{A}_{\omega''}$,

$$\mathfrak{i}_{\{\omega',\omega''\}}(A_{\omega'} \otimes A_{\omega''}) = a(\omega', A_{\omega'})a(\omega'', A_{\omega''}) \ . \tag{4.10}$$

This mapping is injective, and the algebraic tensor product $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ can thus be identified with the subspace

$$\mathfrak{i}_{\{\omega',\omega''\}}(\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}) = \mathrm{span}\{a(\omega', A_{\omega'})a(\omega'', A_{\omega''}) \ : A_{\omega'} \in \mathcal{A}_{\omega'} \ , \ A_{\omega''} \in \mathcal{A}_{\omega''}\}$$
$$\subseteq \bigotimes\nolimits_{\omega\in\Omega} \mathcal{A}_\omega \ .$$

Moreover, $\mathfrak{i}_{\{\omega',\omega''\}}$ is a (faithful) $*$-homomorphism. Recall that the (algebraic) tensor product of $*$-algebras is naturally a $*$-algebra, by Proposition 7.62.

Proof Clearly, by the polynomial relations defining the universal tensor product of C^*-algebras (Definition 4.147), the mapping

$$(A_{\omega'}, A_{\omega''}) \mapsto a(\omega', A_{\omega'})a(\omega'', A_{\omega''})$$

from $\mathcal{A}_{\omega'} \times \mathcal{A}_{\omega''}$ to $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ is bilinear. Thus, by the universal property of algebraic tensor products (Proposition 7.10), there is a unique linear mapping:

$$\mathfrak{i}_{\{\omega',\omega''\}} : \mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''} \to \bigotimes\nolimits_{\omega\in\Omega} \mathcal{A}_\omega$$

satisfying (4.10) for all $A_{\omega'} \in \mathcal{A}_{\omega'}$ and $A_{\omega''} \in \mathcal{A}_{\omega''}$. Let $\mathcal{B}' \subseteq \mathcal{A}_{\omega'}$ and $\mathcal{B}'' \subseteq \mathcal{A}_{\omega''}$ be (Hamel) bases for $\mathcal{A}_{\omega'}$ and $\mathcal{A}_{\omega''}$, respectively. By Proposition 7.12,

$$\{A' \otimes A'' \ : \ A' \in \mathcal{B}' \ , \ A'' \in \mathcal{B}''\} \subseteq \mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$$

is a (Hamel) basis for the algebraic tensor product $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$. In order to demonstrate the injectivity of $\mathfrak{i}_{\{\omega',\omega''\}}$, for any $n \in \mathbb{N}$ and arbitrary sequences $A'_1, \ldots, A'_n \in \mathcal{B}'$, $A''_1, \ldots, A''_n \in \mathcal{B}''$, $\alpha_1, \ldots, \alpha_n \in \mathbb{C}\backslash\{0\}$, one has to prove that

$$a(\omega', A'_1)a(\omega'', \alpha_1 A''_1) + \cdots + a(\omega', A'_n)a(\omega'', \alpha_n A''_n) \neq 0 \ .$$

Equivalently, one may prove that for $n \in \mathbb{N}$ arbitrary linearly independent elements $A'_1, \ldots, A'_n \in \mathcal{A}'_\omega$ and n arbitrary non-zero elements $A''_1, \ldots, A''_n \in \mathcal{A}_{\omega''}$,

$$a(\omega', A'_1)a(\omega'', A''_1) + \cdots + a(\omega', A'_n)a(\omega'', A''_n) \neq 0 \ .$$

Assume, by contradiction, that this element of $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ is zero. Let $\rho' \in E(\mathcal{A}_{\omega'})$ and $\rho'' \in E(\mathcal{A}_{\omega''})$ be arbitrary states of the C^*-algebras $\mathcal{A}_{\omega'}$ and $\mathcal{A}_{\omega''}$, respectively. By Proposition 4.155, there is a (product) state on $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ whose evaluation on the above element is

$$\rho'(A'_1)\rho''(A''_1) + \cdots + \rho'(A'_n)\rho''(A''_n) = \rho'(\rho''(A''_1)A'_1 + \cdots + \rho''(A''_n)A'_n) = 0 \ .$$

By Exercise 4.64 (i), one can choose $\rho'' \in E(\mathcal{A}_{\omega''})$ such that $\rho''(A''_k) \neq 0$ for some $k \in \{1, \ldots, n\}$. But then the above equality combined with Exercise 4.64 (i) would imply that $A'_1, \ldots, A'_n \in \mathcal{A}_{\omega'}$ are not linearly independent. The fact that $\mathfrak{i}_{\{\omega',\omega''\}}$ is a $*$-homomorphism is clear. □

From Corollary 4.156, by using the identification of the algebraic tensor product $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ with a subspace of the C^*-algebra $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$, $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ is endowed with a norm satisfying

$$\left\|A' \otimes A''\right\| = \left\|A'\right\| \left\|A''\right\| \ , \qquad A' \in \mathcal{A}_{\omega'} \ , \ A'' \in \mathcal{A}_{\omega''} \ .$$

Such norms for (algebraic) tensor products are called "crossnorms." In fact, by the last proposition, $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ is identified with a $*$-*subalgebra* of $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$. With this identification, $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ is a normed $*$-algebra whose norm satisfies

$$\left\|B^* B\right\| = \|B\|^2 \ , \qquad B \in \mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''} \ .$$

In particular, the completion of $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ with respect to this norm is a C^*-algebra. Such a norm for a tensor product of C^*-algebras is called a "C^*-tensor norm." In fact, any "C^*-tensor norm" is a special case of a crossnorm. See, for instance, [14, 11.3.10. Corollary].

Exercise 4.158 Let Ω be any nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be a unital C^*-algebra. For any $\omega', \omega'' \in \Omega$, $\omega' \neq \omega''$, show that the range of the unique morphism of C^*-representations $\bigotimes_{\omega\in\{\omega',\omega''\}} \mathcal{A}_\omega \to \bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ together with the inclusion mapping $\mathfrak{i}_{\{\omega',\omega''\}} : \mathcal{A}_\omega \otimes \mathcal{A}_{\omega'} \to \bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$ is a completion for $\mathcal{A}_\omega \otimes \mathcal{A}_{\omega'}$ as $*$-subalgebra of $\bigotimes_{\omega\in\Omega} \mathcal{A}_\omega$.

It turns out that the C^*-tensor norm induced in $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$ by the universal tensor product of these two (unital) C^*-algebras is the largest possible norm of this type for the $*$-algebra $\mathcal{A}_{\omega'} \otimes \mathcal{A}_{\omega''}$:

Corollary 4.159 *Let $\mathcal{A}_1, \mathcal{A}_2$ be two unital C^*-algebras and let $\|\cdot\|$ be any C^*-tensor norm for $\mathcal{A}_1 \otimes \mathcal{A}_2$. For all $B \in \mathcal{A}_1 \otimes \mathcal{A}_2$, one has $\|B\| \leq \|\mathfrak{i}_{\{1,2\}}(B)\|$, where*

$$\mathfrak{i}_{\{1,2\}} : \mathcal{A}_1 \otimes \mathcal{A}_2 \to \bigotimes_{\omega \in \{1,2\}} \mathcal{A}_\omega$$

is the unique linear transformation (inclusion mapping) of Lemma 4.157 for $\Omega = \{1, 2\}$.

Proof Let $(\mathcal{A}, \mathfrak{i})$ be a completion of the $*$-algebra $\mathcal{A}_1 \otimes \mathcal{A}_2$ with respect to the given C^*-tensor norm. Then, the mapping $a : (\{1\}\times \mathcal{A}_1) \cup (\{2\}\times \mathcal{A}_2) \to \mathcal{A}$ defined by

$$a(1, A_1) \doteq \mathfrak{i}(A_1 \otimes \mathfrak{1}) , \qquad A_1 \in \mathcal{A}_1 ,$$
$$a(2, A_2) \doteq \mathfrak{i}(\mathfrak{1} \otimes A_2) , \qquad A_2 \in \mathcal{A}_2 ,$$

is a C^*-representation of the polynomial relations defining the universal tensor product $\bigotimes_{\omega \in \{1,2\}} \mathcal{A}_\omega$. Thus, by Definition 4.127 (iv)–(v), there is a unique $*$-homomorphism

$$\Xi : \bigotimes_{\omega \in \{1,2\}} \mathcal{A}_\omega \to \mathcal{A}$$

satisfying

$$\Xi(a(1, A_1)a(2, A_2)) = \mathfrak{i}(A_2 \otimes A_2) .$$

Thus, $\Xi \circ \mathfrak{i}_{\{1,2\}} = \mathfrak{i}$. As $*$-homomorphisms of C^*-algebras are contractions (Proposition 4.97), it follows that, for all $B \in \mathcal{A}_1 \otimes \mathcal{A}_2$,

$$\|B\| = \|\mathfrak{i}(B)\| = \left\|\Xi \circ \mathfrak{i}_{\{1,2\}}(B)\right\| \leq \left\|\mathfrak{i}_{\{1,2\}}(B)\right\| .$$

□

In fact, in general, the C^*-tensor norms for tensor products of C^*-algebra are *not* unique. Similar to the "reasonable crossnorms" for the tensor products of normed spaces (see Definition 7.44 and the discussion on injective and projective norms following it), besides the largest C^*-tensor norm, referring to universal tensor products, as proven above, there is always a smallest C^*-tensor norm for any fixed finite family of C^*-algebras. See, for instance, [14, 11.3.9. Theorem]. They are called "spacial norms" for tensor products of C^*-algebras. In our setting, they refer to the C^*-representation (of the polynomial relations for universal tensor products) given in Proposition 4.149, when these are constructed for *faithful* representations of the corresponding C^*-algebras.

A C^*-algebra $\mathcal{A}$ is said to be "nuclear" if, for any other C^*-algebra $\mathcal{A}'$, the tensor product $\mathcal{A} \otimes \mathcal{A}'$ admits only one C^*-tensor norm (the one induced by the corresponding universal tensor products, in this case), i.e., the largest and the smallest C^*-tensor norms for $\mathcal{A} \otimes \mathcal{A}'$ are identical. Important examples of nuclear C^*-algebras are all the commutative ones [14, 11.3.13. Theorem] and all the finite-dimensional ones [14, 11.3.11. Lemma]. Further nuclear C^*-algebras are construct from these examples by using the following result.

Proposition 4.160 *Let $\mathcal{A}$ be any C^*-algebra together with a net $\{\mathcal{A}_i\}_{i\in I}$ of nuclear C^*-subalgebras. If the union $\cup_{i\in I}\mathcal{A}_i \subseteq \mathcal{A}$ is dense in $\mathcal{A}$, then $\mathcal{A}$ is also nuclear.*

Proof Exercise. *Hint*: Use the fact that any C^*-tensor norm is a crossnorm [14, 11.3.10. Corollary]. See also [14, 11.3.12. Proposition]. □

From the last proposition, it follows, in particular, that the universal tensor product of finite-dimensional unital C^*-subalgebras is always nuclear.

Corollary 4.161 *Let Ω be a nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be an arbitrary* finite-dimensional *unital C^*-algebra. The universal tensor product $\bigotimes_{\omega\in\Omega}\mathcal{A}_\omega$ corresponds to a nuclear unital C^*-algebra.*

Proof Exercise. □

By recalling the definition of algebraic tensor products as multilinear forms on spaces of linear forms (Definition 7.9) and generalizing Lemma 4.157 for n-fold, $n \in \mathbb{N}$, tensor products, we also obtain from the above results that product states separate points of universal tensor products.

Corollary 4.162 *Let Ω be a nonempty, possibly infinite set. For every $\omega \in \Omega$, let $\mathcal{A}_\omega$ be a unital C^*-algebra. For every $A \in \bigotimes_{\omega\in\Omega}\mathcal{A}_\omega$, one has that $\rho(A) = 0$ for any* product *state ρ only if $A = 0$. In other words, product states separate points of the C^*-algebra $\bigotimes_{\omega\in\Omega}\mathcal{A}_\omega$.*

Proof Exercise. □

4.8.2 The CAR C-Algebras*

We give in the present subsection a second important example of polynomial relations, the so-called canonical anticommutation relations (CAR), and define CAR C^*-algebras as being the universal C^*-algebras associated with them.

Definition 4.163 (CAR C^*-Algebras) Let H be any complex pre-Hilbert space and $G \subseteq H$ any nonempty subset. We take G as being an index set and define the following family of polynomial relations:

$$\begin{aligned}\mathfrak{R} \doteq \{&\bar{\alpha}_1 a(x_1) + \ldots + \bar{\alpha}_n a(x_n)\\ &= 0 : n \in \mathbb{N},\ x_1, \ldots, x_n \in G,\ \alpha_1, \ldots, \alpha_n \in \mathbb{C},\ \alpha_1 x_1\\ &\quad + \ldots + \alpha_n x_n = 0\}\end{aligned}$$

$$\cup\{a(x)a(x')^* + a(x')^*a(x) = \langle x, x'\rangle \mathbb{1} : x, x' \in G\}$$
$$\cup\{a(x)a(x') = -a(x')a(x) : x, x' \in G\} .$$

(a being again a mapping from G to some C^*-algebra). Then,

$$\mathrm{CAR}(G) \doteq (C^*(G, \mathfrak{R}), a) ,$$

where $(C^*(G, \mathfrak{R}), a)$ is any fixed universal C^*-representation of $\mathfrak{R}$. By an abuse of notation, sometimes $\mathrm{CAR}(G)$ also denotes only the universal C^*-algebra $C^*(G, \mathfrak{R})$. $\mathrm{CAR}(G)$ is called the "CAR C^*-algebra" associated with the subset $G \subseteq H$ of the complex pre-Hilbert space H.

Note that the first set of polynomial relations in the last definition is just saying that a behaves antilinearily if G contains linear combinations of other elements of G. In particular, if $G \subseteq H$ is a vector subspace of H, then this set of polynomial relations may be replaced with

$$\{a(\alpha x + x') = \bar{\alpha} a(x) + a(x) : \ x, x' \in G, \ \alpha \in \mathbb{C} \} .$$

In mathematical and theoretical physics, CAR C^*-algebras are those C^*-algebras related with fermionic particles. The corresponding Hilbert spaces, i.e., the completions of the linear spans of the families of vectors defining the CAR C^*-algebras, are called, in this context, "one-particle Hilbert spaces."

Exercise 4.164 Prove that the subset $G \subseteq H$ of a complex pre-Hilbert space H is separable (i.e., it includes a dense countable subset) only if the C^*-algebra $\mathrm{CAR}(G)$ is separable.

Like in the previous example of universal tensor products, Definition 4.163 assumes the existence of a universal C^*-representation of the above-defined family $\mathfrak{R}$ of polynomial relations. By Theorem 4.134, $\mathfrak{R}$ must be admissible, i.e., a C^*-representation $(\mathcal{A}, a)$ must exist and, for all $x \in G$,

$$\sup\{\|a(x)\|_{\mathcal{A}} \ : \ (\mathcal{A}, a) \text{ a } C^*\text{-representation of } \mathfrak{R}\} < \infty . \tag{4.11}$$

The last property is again easily deduced from the polynomial relations: Take any C^*-representation $(\mathcal{A}, a)$ of the family $\mathfrak{R}$ of relation of the last definition. Then, by the CAR, for all $x \in G$,

$$a(x)a(x)^* + a(x)^*a(x) = \|x\|^2 \mathbb{1} .$$

As $a(x)a(x)^*$ and $a(x)^*a(x)$ are positive elements (see Corollary 4.103), for all $x \in H$,

$$\|x\|^2 \mathbb{1} \geq a(x)a(x)^* \geq 0 \qquad \text{and} \qquad \|x\|^2 \mathbb{1} \geq a(x)^*a(x) \geq 0 .$$

This implies in turn that, for all $x \in G$,

$$\|a(x)\|^2 = \|a(x)^* a(x)\| \leq \|x\|^2 . \tag{4.12}$$

Hence, (4.11) holds true for all $x \in G$.

In order to show that the family $\mathfrak{R}$ of Definition 4.163 is admissible, one can use, as it is frequently done in the literature, the well-known "creation and annihilation operators" on the antisymmetric Fock space associated with the Hilbert space, which is the completion of $\operatorname{span}(G)$. See, e.g., [55, Section 5.2]. In fact, later on, we will construct the Fock spaces in an abstract way, by considering the cyclic representation associated with so-called Fock states of CAR C^*-algebras. See Definition 4.177. An alternative (and also well-known) argument to prove that the CAR relations $\mathfrak{R}$ are admissible is the construction of a C^*-representation of them in (universal) tensor products. This refers to the Jordan-Wigner transformation, and here we use this type of argument to show that the CAR are admissible polynomial relations. We start with the case of *finite orthonormal* families of vectors in complex pre-Hilbert spaces.

Lemma 4.165 (The CAR Are Admissible—Finite Orthonormal Case) *Let G be any (nonempty) finite orthonormal family of vectors of a complex pre-Hilbert space. The family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163 is admissible, simple, and unital.* $\mathrm{CAR}(G)$ *is $*$-isomorphic to $\mathcal{L}(\mathbb{C}^{2^{|G|}})$, the ($C^*$-)algebra of complex $2^{|G|} \times 2^{|G|}$ matrices.*

Proof

1. Assume that G has exactly $n \in \mathbb{N}$ elements, $e_1, \ldots, e_n \in G$. We will construct a C^*-representation for the CAR in the universal tensor product $\bigotimes_{\omega \in \Omega} \mathcal{L}(\mathbb{C}^2)$, where $\Omega \doteq \{1, \ldots, n\}$. Define the following 2×2 (Pauli) matrices, which are canonically identified with elements of the algebra $\mathcal{L}(\mathbb{C}^2)$:

$$\sigma^z \doteq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} , \qquad \sigma^+ \doteq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} .$$

For all $\omega \in \Omega$, we define the element $f_\omega^* \in \bigotimes_{\omega \in \Omega} \mathcal{L}(\mathbb{C}^2)$ by

$$f_\omega^* \doteq a(\omega, \sigma^+) \prod_{\omega' \in \Omega \,:\, \omega' < \omega} a(\omega', \sigma^z)$$

for $\omega \geq 2$ and $f_1^* \doteq a(1, \sigma^+)$. Note that

$$\sigma^z \sigma^+ = -\sigma^+ \sigma^z , \quad \sigma^z \sigma^z = \mathfrak{1} , \ \sigma^+ \sigma^+ = 0 , \ \text{and } \sigma^+ (\sigma^+)^* + (\sigma^+)^* \sigma^+ = \mathfrak{1} ,$$

which, combined with the relations defining the universal tensor product $\bigotimes_{\omega \in \Omega} \mathcal{L}(\mathbb{C}^2)$ (Definition 4.147), implies that, for all $\omega, \omega' \in \Omega$,

$$f_\omega^* f_{\omega'}^* = -f_{\omega'}^* f_\omega^* \qquad \text{and} \qquad (f_\omega^*)^* f_{\omega'}^* + f_{\omega'}^* (f_\omega^*)^* = \delta_{\omega,\omega'} \mathfrak{1} .$$

This proves that $\mathfrak{R}$ is admissible. In particular, it has a universal C^*-representation $\mathrm{CAR}(G)$, by Theorem 4.134 and Inequality (4.11).

2. To prove that $\mathfrak{R}$ is simple, we start by constructing a C^*-representation of the polynomial relations defining the universal tensor product $\bigotimes_{\omega\in G}\mathcal{L}(\mathbb{C}^2)$ on the C^*-algebra $\mathrm{CAR}(G)$. Define the mapping $a:\Omega\times\mathcal{L}(\mathbb{C}^2)\to\mathrm{CAR}(G)$ by

$$a\left(\omega,\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\right)\doteq a(e_\omega)^*a(e_\omega)A_{11}+a(e_\omega)a(e_\omega)^*A_{22}$$
$$+(A_{12}a(e_\omega)^*+A_{21}a(e_\omega))\prod_{\omega'\in\Omega\,:\,\omega'<\omega}$$
$$\times(\mathfrak{1}-2a(e_{\omega'})^*a(e_{\omega'}))$$

for all $\omega\in\Omega$ and $A\in\mathcal{L}(\mathbb{C}^2)$ canonically identified with the 2×2 matrix $(A_{i,j})_{i,j\in\{1,2\}}$ above, in the argument of a. Here, if $\omega=1$ then, by definition,

$$\prod_{\omega'\in\Omega\,:\,\omega'<\omega}(\mathfrak{1}-2a(e_{\omega'})^*a(e_{\omega'}))\doteq\mathfrak{1}\,.$$

By using the CAR, one can easily check that this mapping gives a C^*-representation for the relations defining the universal tensor product $\bigotimes_{\omega\in\Omega}\mathcal{L}(\mathbb{C}^2)$. In particular, by Definition 4.127 (iv–v), there is a unique $*$-homomorphism:

$$\Theta:\bigotimes_{\omega\in\Omega}\mathcal{L}(\mathbb{C}^2)\to\mathrm{CAR}(G)$$

satisfying

$$\Theta(a^{\otimes}(\omega,A))=a\left(\omega,\begin{pmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{pmatrix}\right)\,,\qquad \omega\in\Omega,\ A\in\mathcal{L}(\mathbb{C}^2)\,,$$

where $a^{\otimes}$ is the corresponding mapping of the C^*-representation $\bigotimes_{\omega\in\Omega}\mathcal{L}(\mathbb{C}^2)$. This $*$-homomorphism is unital, because

$$a\left(\omega,\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right)=\mathfrak{1}\,.$$

As the (universal) C^*-algebra $\bigotimes_{\omega\in\Omega}\mathcal{L}(\mathbb{C}^2)$ is simple (by Lemma 4.153), Θ has a trivial kernel and is thus faithful (i.e., injective). Simple calculations using the CAR show that

$$\{a(e_\omega)\ :\ \omega\in\Omega\}\subseteq\Theta\left(\bigotimes_{\omega\in\Omega}\mathcal{L}(\mathbb{C}^2)\right)\subseteq\mathrm{CAR}(G)\,.$$

As $\Theta\left(\bigotimes_{\omega\in\Omega}\mathcal{L}(\mathbb{C}^2)\right)$ is a C^*-subalgebra (by Theorem 4.87), it follows from Corollary 4.132 that Θ is surjective, and it is thus a $*$-isomorphism. Hence, $\mathrm{CAR}(G)$ is simple.

3. Recall from Lemma 4.153 that $\bigotimes_{\omega\in G}\mathcal{L}(\mathbb{C}^2)$ is $*$-isomorphic to $\mathcal{L}(\mathbb{C}^{2^{|G|}})$. The fact that $\mathfrak{R}$ is unital directly follows from

$$a(e_\omega)^* a(e_\omega) + a(e_\omega)a(e_\omega)^* = \mathbb{1}\ , \qquad \omega \in \Omega\ .$$

□

Now we use Proposition 4.146 to go from the case of finite families of orthonormal vectors to the general case of orthonormal families.

Proposition 4.166 *Let G be any (not necessarily finite) orthonormal family of vectors of a complex pre-Hilbert space. The family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163 is simple and unital. Moreover,* $\mathrm{CAR}(G)$ *is a nuclear C^*-algebra.*

Proof By combining the last lemma with Proposition 4.146, we conclude that $\mathfrak{R}$ is simple and unital, in the sense of Definition 4.127 (vi)–(vii). Thus, $\mathfrak{R}$ has a universal C^*-representation, and any such a representation corresponds to a C^*-algebra that is unital and simple, thanks to Proposition 4.128 (ii). By the proof of the last lemma, for any (nonempty) finite subset $G' \subseteq G$, the C^*-algebra $\mathrm{CAR}(G')$ is $*$-isomorphic to $\bigotimes_{\omega\in G'}\mathcal{L}(\mathbb{C}^2)$, which is finite dimensional. Then, by combining Proposition 4.160 with Corollary 4.132, one proves that $\mathrm{CAR}(G)$ is nuclear. □

The following corollary of the last proposition states that the CAR are admissible for any (not necessarily orthonormal) family of vectors of a complex pre-Hilbert space.

Corollary 4.167 *Let G be any (nonempty, not necessarily orthonormal) family of vectors of a complex pre-Hilbert space. The family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163 is simple and unital. Moreover,* $\mathrm{CAR}(G)$ *is a nuclear C^*-algebra.*

Proof In order to prove that $\mathfrak{R}$ is admissible in general, note that it suffices to show this property for a collection G of vectors in a Hilbert space H, because any complex pre-Hilbert space is unitarily equivalent to a subspace of a complete one. Assume without loss of generality that this Hilbert space H, together with some linear norm-preserving mapping ι, is the completion of $\mathrm{span}\,(G)$. Let $B \subseteq H$ be any Hilbert basis (i.e., a maximal orthonormal family) for this Hilbert space. Define the mapping $a^* : \mathrm{span}(B) \to \mathrm{CAR}(B)$ by

$$a^*(x) = \sum_{e\in B}\langle e, x\rangle\, a(e)^*\ , \qquad x \in \mathrm{span}(B)\ .$$

Observe that the above sum is finite for any $x \in \text{span}(B)$ and is thus well-defined. By simple computations using the CAR, one checks that this mapping is a linear contraction. Thus, it uniquely extends to a linear contraction $a^* : H \to \text{CAR}(B)$. By construction, $(\text{CAR}(B), a^*(\cdot)^*)$ is a C^*-representation of the family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163 with the index set H, which includes the dense subset span (G). Now, by restricting a^* to G, one verifies that $(\text{CAR}(B), a^*(\cdot)^*)$ leads to a C^*-representation $(\text{CAR}(B), a^* \circ \iota(\cdot)^*|_G)$ of the family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163 with the index set G. For more details, see for instance the proof of the next proposition. Therefore, this family $\mathfrak{R}$ has a universal C^*-representation $\text{CAR}(G)$, by Theorem 4.134 and Inequality (4.11). Assume now without loss of generality that $G = H$ is a Hilbert space. For any finite subset $G' \subseteq G$, let B' be any (finite) orthonormal family of vectors in the complex pre-Hilbert space so that $\text{span}(B') = \text{span}(G')$. By using Exercise 4.131 and Corollary 4.132, we see that the unique morphism of C^*-representations $\text{CAR}(B') \to \text{CAR}(G')$ is surjective. As $\text{CAR}(B')$ is simple (by the last proposition), it is also one-to-one. Thus, $\text{CAR}(G')$ is also a simple C^*-algebra. With this remark, the fact that $\mathfrak{R}$ is simple and unital is proven exactly in the same way as in the orthonormal case, via Proposition 4.146. The fact that $\text{CAR}(G)$ is a nuclear C^*-algebra is again proven by combining Proposition 4.160 with Corollary 4.132, with an obvious adaptation of the argument used in the orthonormal case. □

Similar to the case of universal tensor products of unital C^*-algebras (see Definition 4.150), from the universal property of CAR C^*-algebras, we define natural $*$-homomorphisms $\text{CAR}(G') \to \text{CAR}(G)$ for every nonempty subset $G' \subseteq G$:

Definition 4.168 (Canonical Inclusions of CAR C^*-Algebras) Let G be any nonempty family of vectors of a complex pre-Hilbert space. For any nonempty subset $G' \subseteq G$, clearly, for any $\Omega' \subseteq \Omega$, $(\text{CAR}(G), a|_{G'})$ is a C^*-representation of the CAR for the (index) set G'. In particular, from the universal property of $\text{CAR}(G')$, there is a unique morphism $\text{CAR}(G') \to \text{CAR}(G)$ of C^*-representations, that is denoted here by $\mathfrak{i}_{GG'}$.

As a direct consequence of Corollary 4.167, it turns out that the canonical inclusions of CAR C^*-algebras are always faithful.

Corollary 4.169 *Let G be any nonempty family of vectors of a complex pre-Hilbert space. For any nonempty subset $G' \subseteq G$, the natural $*$-homomorphism*

$$\mathfrak{i}_{GG'} : \text{CAR}(G') \to \text{CAR}(G)$$

is faithful. In particular, $\text{CAR}(G')$ *is canonically identified with a C^*-subalgebra of* $\text{CAR}(G)$.

Proof To show this assertion, it suffices to straightforwardly adapt the arguments used to prove Corollary 4.154. □

In the next proposition, we show that the equivalence class of the C^*-algebra $\mathrm{CAR}(G)$ is invariant with respect to the operations of taking the closure and the linear span of G.

Proposition 4.170 *Let G be any (nonempty) family of vectors of a complex pre-Hilbert space, whose closure is denoted by $\overline{G}$. Take any universal C^*-representation $(\mathrm{CAR}(G), a)$ of the CAR, i.e., the family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163, for the index set G:*

(i) *The universal C^*-algebras $\mathrm{CAR}(G)$, $\mathrm{CAR}(\mathrm{span}(G))$, and $\mathrm{CAR}(\overline{G})$ are $*$-isomorphic to each other.*
(ii) *Let a_{span} be the unique antilinear extension of a from G to $\mathrm{span}(G)$. Then, $(\mathrm{CAR}(G), a_{\mathrm{span}})$ is a universal C^*-representation of the CAR for the index set $\mathrm{span}(G)$.*
(iii) *Let a_{cl} be the unique continuous extension of a from G to $\overline{G}$. Then, $(\mathrm{CAR}(G), a_{\mathrm{cl}})$ is a universal C^*-representation of the CAR for the index set $\overline{G}$.*
(iv) *Conversely, for any universal C^*-representations $(\mathrm{CAR}(\mathrm{span}(G)), a)$ and $(\mathrm{CAR}(\overline{G}), a)$, one has that $(\mathrm{CAR}(\mathrm{span}(G)), a|_G)$ and $(\mathrm{CAR}(\overline{G}), a|_G)$ are universal C^*-representations of the CAR for the index set G.*

Proof

1. Take any universal C^*-representation $(\mathrm{CAR}(G), a)$. Observe from the CAR that, for any finite sequences $x_1, \ldots, x_n \in G$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$, one has that $\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$ only if $\bar{\alpha}_1 a(x_1) + \cdots + \bar{\alpha}_n a(x_n) = 0$. Thus, there is a unique antilinear mapping:

$$a_{\mathrm{span}} : \mathrm{span}(G) \to \mathrm{CAR}(G)$$

extending a from G to $\mathrm{span}(G)$. From the CAR, $(\mathrm{CAR}(G), a_{\mathrm{span}})$ is a C^*-representation of the family $\mathfrak{R}$ of polynomial relations defined in Definition 4.163 for the index set $\mathrm{span}(G)$. For any universal C^*-representation $(\mathrm{CAR}(\mathrm{span}(G)), a)$ (which exists, thanks to Corollary 4.167), let

$$\mathfrak{i}_{G\mathrm{span}(G)} : \mathrm{CAR}(\mathrm{span}(G)) \to \mathrm{CAR}(G)$$

be the unique morphism of C^*-representations such that

$$a_{\mathrm{span}} = \mathfrak{i}_{G\mathrm{span}(G)} \circ a \ .$$

See Definition 4.127 (iv)–(v). By Exercise 4.131 and Corollary 4.132, it follows that

$$\mathfrak{i}_{G\mathrm{span}(G)}(\mathrm{CAR}(\mathrm{span}(G)))$$

is dense in $\mathrm{CAR}(G)$. Recall from Theorem 4.87 that this subspace of $\mathrm{CAR}(G)$ is also a unital C^*-subalgebra. Hence, $\mathfrak{i}_{G\mathrm{span}(G)}$ is surjective. Additionally, $\mathfrak{i}_{G\mathrm{span}(G)}$ is also injective, $\mathrm{CAR}(\mathrm{span}(G))$ being simple (Corollary 4.167). Thus, $\mathfrak{i}_{G\mathrm{span}(G)}$ is a $*$-isomorphism.

2. Let $(\mathcal{A}, a)$ be any C^*-representation of the CAR for the index set $\mathrm{span}(G)$ and

$$\Theta : \mathrm{CAR}(\mathrm{span}(G)) \to \mathcal{A}$$

the unique morphism of C^*-representations. See again Definition 4.127 (iv)–(v). Then, clearly,

$$\Theta \circ \mathfrak{i}_{G\mathrm{span}(G)}^{-1} : \mathrm{CAR}(G) \to \mathcal{A}$$

is a morphism of C^*-representations from $(\mathrm{CAR}(G), a_{\mathrm{span}})$ to $(\mathcal{A}, a)$. This morphism has to be unique because, otherwise, Θ would not be unique, $\mathfrak{i}_{G\mathrm{span}(G)}$ being a $*$-isomorphism. Thus, $(\mathrm{CAR}(G), a_{\mathrm{span}})$ is a universal C^*-representation of the CAR for the index set $\mathrm{span}(G)$.

3. From the CAR, the mapping $a_{\mathrm{span}} : \mathrm{span}(G) \to \mathrm{CAR}(G)$ is continuous and antilinear. Similar to the linear case, as $\mathrm{CAR}(G)$ is a complete normed space, it has a unique continuous antilinear extension to the closure of $\mathrm{span}(G)$. Let a_{cl} denote the restriction of this new mapping to the closure of G. This is the unique continuous extension of $a : G \to \mathrm{CAR}(G)$ to the closure of G. Exactly in the same way as done above for $(\mathrm{CAR}(G), a_{\mathrm{span}})$, we show that $(\mathrm{CAR}(G), a_{\mathrm{cl}})$ is a universal C^*-representation for the index sets $\overline{G}$. Mutatis mutandis for $(\mathrm{CAR}(\mathrm{span}(G)), a|_G)$ and $(\mathrm{CAR}(\overline{G}), a|_G)$.

□

In the following we show that universal tensor products of CAR C^*-algebras always lead to simple C^*-algebras. This fact will be used later on to construct so-called product states of even CAR C^*-algebras.

Proposition 4.171 *Let G be any (nonempty) orthonormal family of vectors of a complex pre-Hilbert space. Let $\mathfrak{P}$ be any collection of nonempty disjoint subsets of G. The universal tensor product $\bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}(\Lambda)$ is a simple nuclear C^*-algebra.*

Proof For any fixed finite subset $\Lambda \in \mathcal{P}_f(G)\backslash\{\emptyset\}$ of G, let

$$\mathfrak{P}(\Lambda) \doteq \left\{\Lambda \cap \Lambda' \ : \ \Lambda' \in \mathfrak{P},\ \Lambda \cap \Lambda' \neq \emptyset\right\} \subseteq \mathcal{P}_f(G) \ .$$

By the universal property of $\bigotimes_{\Lambda'\in\mathfrak{P}(\Lambda)} \mathrm{CAR}(\Lambda')$ (Definition 4.127 (iv)–(v)), there is a unique $*$-homomorphism:

$$\Theta_{\mathfrak{P}(\Lambda)} : \bigotimes_{\Lambda'\in\mathfrak{P}(\Lambda)} \mathrm{CAR}(\Lambda') \to \bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}(\Lambda)$$

mapping

$$a(\Lambda', A_{\Lambda'}) \in \bigotimes_{\Lambda' \in \mathfrak{P}(\Lambda)} \mathrm{CAR}(\Lambda') ,$$

for $\Lambda' \in \mathfrak{P}(\Lambda)$ and $A_{\Lambda'} \in \mathrm{CAR}(\Lambda')$, to

$$a(\Lambda(\Lambda'), i_{\Lambda(\Lambda')\Lambda'}(A)) \in \bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda) ,$$

where $\Lambda(\Lambda') \in \mathfrak{P}$ is the unique element of $\mathfrak{P}$ such that $\Lambda' \subseteq \Lambda(\Lambda')$ while $i_{\Lambda(\Lambda')\Lambda'}$ is the canonical inclusion $\mathrm{CAR}(\Lambda') \to \mathrm{CAR}(\Lambda(\Lambda'))$. By a slight abuse of notation, note that we also use the same letter a for the mapping of each C^*-representation. As $a(\Lambda(\Lambda'), \cdot)$ is injective (see Corollary 4.156), $\Theta_{\mathfrak{P}(\Lambda)}$ is not trivial. Recalling from Lemma 4.165 that, for all $\Lambda' \in \mathfrak{P}(\Lambda)$, $\mathrm{CAR}(\Lambda')$ is $*$-isomorphic to $\mathcal{L}(\mathbb{C}^{2^{|\Lambda'|}})$ and we can thus infer from Lemma 4.153 that $\bigotimes_{\Lambda' \in \mathfrak{P}(\Lambda)} \mathrm{CAR}(\Lambda')$ is simple. $\Theta_{\mathfrak{P}(\Lambda)}$ is therefore faithful. For all $\Lambda \in \mathcal{P}_f(G)\backslash\{\emptyset\}$, define the C^*-subalgebra

$$\mathcal{A}_\Lambda \doteq \Theta_{\mathfrak{P}(\Lambda)} \left(\bigotimes_{\Lambda' \in \mathfrak{P}(\Lambda)} \mathrm{CAR}(\Lambda') \right) \subseteq \bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda)$$

By construction and finite dimensionality, $\{\mathcal{A}_\Lambda\}_{\Lambda \in \mathcal{P}_f(G)\backslash\{\emptyset\}}$ is an increasing net of simple C^*-subalgebras whose union is dense in $\bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda)$. Thus, by Proposition 4.57, $\bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda)$ is simple. Note $\mathcal{A}_\Lambda$ is also nuclear for any $\Lambda \in \mathcal{P}_f(G)\backslash\{\emptyset\}$, by finite dimensionality [14, 11.3.11. Lemma]. As a consequence, applying Proposition 4.160, we also conclude that $\bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda)$ is nuclear. □

The next proposition states that any universal tensor product of the form $\bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda)$ is $*$-isomorphic to a C^*-subalgebra of some CAR C^*-algebra, in a natural way.

Proposition 4.172 *Let G be any (nonempty) orthonormal family of vectors of a complex pre-Hilbert space. Let $\mathfrak{P}$ be any collection of nonempty disjoint subsets of G. Assume that there is a one-to-one mapping $p : \mathfrak{P} \to (\cup\mathfrak{P})^c$, that is, p injectively maps the elements of $\mathfrak{P}$ to points in the complement of the union of all subsets of G that are in $\mathfrak{P}$. Then, there is a unique $*$-homomorphism:*

$$\Theta_{\mathfrak{P}} : \bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda) \to \mathrm{CAR}((\cup\mathfrak{P}) \cup p(\mathfrak{P}))$$

such that, for all $\Lambda \in \mathfrak{P}$ and $e \in \Lambda$,

$$\Theta_{\mathfrak{P}}(a^{\otimes}(\Lambda, a(e))) = a(e)(a(p(\Lambda))^* - a(p(\Lambda))) ,$$

$a^{\otimes}$ being the mapping of the C^*-representation $\bigotimes_{\Lambda \in \mathfrak{P}} \mathrm{CAR}(\Lambda)$, while the one of $\mathrm{CAR}(\Lambda)$ is always denoted by a for any subset $\Lambda \subseteq G$. Additionally, this $*$-homomorphism is faithful.

Proof The uniqueness of $\Theta_{\mathfrak{P}}$ follows from Corollary 4.132 and Exercise 4.131. Observing, for any fixed $\Lambda \in \mathfrak{P}$, that the mapping

$$e \mapsto a(e)(a(p(\Lambda))^* - a(p(\Lambda)))$$

from Λ to $\mathrm{CAR}((\cup\mathfrak{P}) \cup p(\mathfrak{P}))$ gives a C^*-representation of the CAR for the index set Λ , by Definition 4.127 (iv)–(v), there is a unique $*$-homomorphism

$$\Theta_\Lambda : \mathrm{CAR}(\Lambda) \to \mathrm{CAR}((\cup\mathfrak{P}) \cup p(\mathfrak{P}))$$

satisfying

$$\Theta_\Lambda(a(e)) = a(e)(a(p(\Lambda))^* - a(p(\Lambda))) , \qquad e \in \Lambda .$$

By simple computations using the CAR, one checks that the mapping

$$(\Lambda, A_\Lambda) \mapsto \ \Theta_\Lambda(A_\Lambda) \in \mathrm{CAR}((\cup\mathfrak{P}) \cup p(\mathfrak{P})) , \qquad \Lambda \in \mathfrak{P},\ A_\Lambda \in \mathrm{CAR}(\Lambda) ,$$

gives a C^*-representation of the polynomial relations defining the universal tensor product $\bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}(\Lambda)$. Thus, $\Theta_{\mathfrak{P}}$ is the unique morphism:

$$\bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}(\Lambda) \to \mathrm{CAR}((\cup\mathfrak{P}) \cup p(\mathfrak{P}))$$

of C^*-representations (see again Definition 4.127 (iv)–(v)). $\Theta_{\mathfrak{P}}$ is nontrivial, because every Θ_Λ, $\Lambda \in \mathfrak{P}$, is nontrivial. As $\bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}(\Lambda)$ is simple (Proposition 4.171), the $*$-homomorphism $\Theta_{\mathfrak{P}}$ is faithful. □

The last proposition allows us to define states of CAR C^*-algebra via product states in universal tensor products of smaller CAR C^*-algebras. This construction will be discussed in detail below, in the context of even CAR C^*-algebras.

Any CAR C^*-algebra is naturally endowed with a state called the "Fock state" associated with this C^*-algebra. We first define these states in CAR C^*-algebras corresponding to finite orthonormal families of vectors in a complex pre-Hilbert space and then extend the definition for the general case by means of Proposition 4.143. Let $G = \{e_1, \ldots, e_n\}$, $n \in \mathbb{N}$, denote any such a family. We start with a few simple technical remarks. For any subset $G' \subseteq G$, define $a(G') \in \mathrm{CAR}(G)$ by

$$a(G') \doteq \prod_{k\in\{1,\ldots,n\}\ ,\ e_k\in G'} a(e_k) ,$$

if G' is nonempty, and $a(\emptyset) \doteq \mathfrak{1}$. We say that $A \in \mathrm{CAR}(G)$ is a "normally ordered monomial" if

$$A = a(G')^* a(G'')$$

for subsets $G', G'' \subseteq G$ such that $G' \cup G'' \neq \emptyset$. Let $\mathrm{NO}(G) \subseteq \mathrm{CAR}(G)$ denote the subspace of all linear combinations of normally ordered monomials.

Lemma 4.173 *Let $G = \{e_1, \ldots, e_n\}$, $n \in \mathbb{N}$, be a finite orthonormal family of vectors in a complex pre-Hilbert space. Then,*

$$\mathrm{CAR}(G) = \mathbb{C}\mathbb{1} \oplus \mathrm{NO}(G) ,$$

i.e., any $A \in \mathrm{CAR}(G)$ can be uniquely decomposed as $A = c\mathbb{1} + B$ with $c \in \mathbb{C}$ and $B \in \mathrm{NO}(G)$.

Proof Recall that $\mathrm{CAR}(G)$ has finite dimension (for G is finite), thanks to Lemma 4.165. From the CAR,

$$\mathrm{CAR}(G) = \mathrm{span}(\{\mathbb{1}\} \cup \mathrm{NO}(G)) .$$

Thus, for any $A \in \mathrm{CAR}(G)$, there are $c \in \mathbb{C}$ and $B \in \mathrm{NO}(G)$ such that $A = c\mathbb{1} + B$. To prove uniqueness of this decomposition, we observe the following: From the proof of Lemma 4.165 (Point 2), there is a $*$-isomorphism:

$$\Theta : \bigotimes_{\omega \in \{1,\ldots,n\}} \mathcal{L}(\mathbb{C}^2) \to \mathrm{CAR}(G)$$

satisfying

$$\Theta\left(a^{\otimes}\left(1, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right) \cdots a^{\otimes}\left(n, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right)\right) = \prod_{k \in \{1,\ldots,n\}} a(e_k)^* a(e_k) = a(G)^* a(G) ,$$

where $a^{\otimes}$ is the corresponding mapping of the C^*-representation $\bigotimes_{\omega \in \{1,\ldots,n\}} \mathcal{L}(\mathbb{C}^2)$. As $*$-isomorphisms of C^*-algebras are norm-preserving, from Corollary 4.156 we conclude that $\|a(G)^* a(G)\| = 1$. Assume now that $A = c'\mathbb{1} + B' = c\mathbb{1} + B$ for some $c, c' \in \mathbb{C}$ and $B, B' \in \mathrm{NO}(G)$. Then, from the CAR,

$$a(G)^* A a(G) = c a(G)^* a(G) = c' a(G)^* a(G) .$$

Thus, $c' = c$ and, consequently, $B = B'$. □

By the last lemma, there is a unique linear form $\rho_{\mathrm{Fock}} : \mathrm{CAR}(G) \to \mathbb{C}$ satisfying

$$\rho^G_{\mathrm{Fock}}(c\mathbb{1} + B) = c , \qquad c \in \mathbb{C}, \ B \in \mathrm{NO}(G) .$$

This functional turns out to be a state.

Lemma 4.174 (Fock State—Finite Case) *Let* $G = \{e_1, \ldots, e_n\}$*,* $n \in \mathbb{N}$*, be a finite orthonormal family of vectors in a complex pre-Hilbert space. Then,* ρ_{Fock}^{G} *is a state on* $\text{CAR}(G)$.

Proof Observe from the last lemma that, for all $A \in \text{CAR}(G)$ and all (possibly empty) subsets $G', G'' \subseteq G$, there are constants $c(G', G'') \in \mathbb{C}$ such that

$$A = \sum_{G', G'' \subseteq G} c(G', G'') a(G')^* a(G'') \ .$$

From simple computations using the CAR, it follows that

$$A^* A = \sum_{G' \subseteq G} |c(G', \emptyset)|^2 \mathbf{1} + B \ ,$$

where $B \in \text{NO}(G)$. Recall that any positive element of a C^*-algebra is of the form $A^* A$ (Corollary 4.103). As a consequence, ρ_{Fock}^{G} is a positive functional. Clearly, $\rho_{\text{Fock}}^{G}(\mathbf{1}) = 1$ and, therefore, ρ_{Fock}^{G} is a state. See Definition 4.62 (or Definition 4.61 together with Proposition 4.44). □

Now, by means of Proposition 4.143, we extend the above definition of Fock states to the CAR C^*-algebra of any (not necessarily finite) orthonormal family of vectors in a complex pre-Hilbert space.

Corollary 4.175 (Fock States—General Orthogonal Case) *Let* G *be any (nonempty) orthonormal family of vectors in a complex pre-Hilbert space. There is a unique state* $\rho_{\text{Fock}}^{G} \in E(\text{CAR}(G))$ *such that, for every* $e_1, \ldots, e_m \in G$ *and* $e'_1, \ldots, e'_n \in G$*,* $m, n \in \mathbb{N}_0$*,* $m + n \geq 1$*, one has that*

$$\rho_{\text{Fock}}^{G}(a(e_1)^* \cdots a(e_m)^* a(e_1) \cdots a(e_n)) = 0 \ .$$

Proof Observe that the above conditions uniquely define the (Fock) state (provided it exists). This follows from Corollary 4.132 combined with the CAR. To prove its existence, let $\mathfrak{D} \doteq \mathcal{P}_f(G) \backslash \emptyset$ be all nonempty finite subsets of the orthonormal set G of vectors. Then, the family $\{(\text{CAR}(\Lambda), a)\}_{\Lambda \in \mathfrak{D}}$ is a directed system of universal C^*-representations in the sense of Definition 4.136. Note that this system is faithful, because the CAR are polynomial relations that are simple and unital (Corollary 4.167). See, for instance, the proof of Proposition 4.146 (Point 2). In particular, the directed system has a faithful inductive limit, by Proposition 4.144. As the directed system consists of universal C^*-representations, the corresponding inductive limit can be identified with a universal C^*-representation $(\mathcal{A}, a)$ of the CAR for the index set $\cup \mathfrak{D} = G$. See Definitions 4.137 (iii) and 4.139 as well as Exercise 4.140. The family $\{\rho_{\text{Fock}}^{\Lambda}\}_{\Lambda \in \mathfrak{D}}$ of (Fock) states $\rho_{\text{Fock}}^{\Lambda} \in E(\text{CAR}(\Lambda))$ (Lemma 4.174) is clearly compatible (see Definition 4.142) with the directed system, again by Corollary 4.132. Thus, by Proposition 4.143, there is a unique state $\rho \in E(\mathcal{A})$ satisfying

$$\rho \circ \Theta_\Lambda = \rho_{\mathrm{Fock}}^{\Lambda} , \qquad \Lambda \in \mathfrak{D} .$$

Finally, let Θ be the unique morphism $\mathrm{CAR}(G) \to \mathcal{A}$ of C^*-representations. See Definitions 4.127 (iv)–(v) and 4.163. Then, the state $\rho_{\mathrm{Fock}}^{G} \doteq \rho \circ \Theta$ has the required property. □

Notice from the Cauchy-Schwarz inequality for states[22] that, for all $e \in G$ and $A \in \mathrm{CAR}(G)$,

$$\rho_{\mathrm{Fock}}^{G}(Aa(e)) = \rho_{\mathrm{Fock}}^{G}(a(e)^* A) = 0 .$$

In fact, it is easy to see that this property uniquely defines ρ_{Fock}^{G}, and we use it to define Fock states for the CAR associated with a general index set.

Proposition 4.176 (Fock States—General Case) *Let G be any (nonempty, not necessarily orthogonal) family of vectors in a complex pre-Hilbert space. There is a unique state $\rho_{\mathrm{Fock}}^{G} \in E(\mathrm{CAR}(G))$ such that, for all $x \in G$ and $A \in \mathrm{CAR}(G)$,*

$$\rho_{\mathrm{Fock}}^{G}(Aa(x)) = \rho_{\mathrm{Fock}}^{G}(a(x)^* A) = 0 .$$

Proof Observe the above condition uniquely defines the (Fock) state (provided it exists). This follows again from Corollary 4.132 combined with the CAR. To prove its existence, let G' be any Hilbert basis of a completion $(H, \mathfrak{i})$ of $\mathrm{span}(G)$ and let

$$\Theta : \mathrm{CAR}(G) \to \mathrm{CAR}(G')$$

be the unique morphism of C^*-representations between $(\mathrm{CAR}(G), a)$ and $(\mathrm{CAR}(G'), a_{\overline{\mathrm{span}(G')}} \circ \mathfrak{i})$. See Proposition 4.170 and its proof for the definition of the mapping:

$$\overline{\mathrm{span}(G')} = H \to \mathrm{CAR}(G') .$$

Define the state

$$\rho_{\mathrm{Fock}}^{G} \doteq \rho_{\mathrm{Fock}}^{G'} \circ \Theta \in E(\mathrm{CAR}(G)) .$$

By construction,

$$\rho_{\mathrm{Fock}}^{G'}(Aa_{\overline{\mathrm{span}(G')}}(x)) = \rho_{\mathrm{Fock}}^{G'}(a_{\overline{\mathrm{span}(G')}}(x)^* A) = 0 , \qquad x \in G', \ A \in \mathrm{CAR}(G') .$$

Recall that states and $*$-homomorphisms of C^*-algebras are always continuous. Thus, by continuity and antilinearity of $a_{\overline{\mathrm{span}(G')}}$ (see, e.g., Definition 4.163 and

[22] Recall that $|\rho(A^*B)|^2 \le \rho(A^*A)\rho(B^*B)$ for $A, B \in \mathcal{A}$ and any state ρ on a C^*-algebra $\mathcal{A}$.

Inequality (4.12)),

$$\rho_{\mathrm{Fock}}^{G'}(Aa_{\overline{\mathrm{span}(G')}}(x)) = \rho_{\mathrm{Fock}}^{G'}(a_{\overline{\mathrm{span}(G')}}(x)^* A) = 0 \ , \qquad x \in H, \ A \in \mathrm{CAR}(G') \ .$$

Hence,

$$\rho_{\mathrm{Fock}}^{G}(Aa(x)) = \rho_{\mathrm{Fock}}^{G}(a(x)^* A) = 0 \ , \qquad x \in G, \ A \in \mathrm{CAR}(G) \ .$$

□

In the following we introduce antisymmetric Fock spaces and the corresponding annihilation and creation operators via the cyclic representations associated with Fock states (see Definition 4.88 and Theorem 4.113).

Definition 4.177 (Fermionic Fock Spaces) Let G be any (nonempty) family of vectors of a complex pre-Hilbert space. Let $(H_{\mathrm{Fock}}^G, \pi_{\mathrm{Fock}}^G, \Omega_{\mathrm{Fock}}^G)$ be any cyclic representation of $\mathrm{CAR}(G)$ associated with the Fock state $\rho_{\mathrm{Fock}}^G \in E(\mathrm{CAR}(G))$, i.e.,

$$\rho_{\mathrm{Fock}}^G(\cdot) = \left\langle \Omega_{\mathrm{Fock}}^G, \pi_{\mathrm{Fock}}^G(\cdot)\Omega_{\mathrm{Fock}}^G \right\rangle \ .$$

(i) The Hilbert space H_{Fock}^G is named "(antisymmetric or fermionic) Fock space associated with G," while $\Omega_{\mathrm{Fock}}^G \in H_{\mathrm{Fock}}^G$ is the corresponding "Fock vacuum vector."

(ii) For all $x \in G$, the operator

$$A(x) \doteq \pi_{\mathrm{Fock}}^G(a(x)) \in \mathcal{B}(H_{\mathrm{Fock}}^G)$$

is called the "annihilation operator" associated with x. Its adjoint $A^*(x) \doteq A(x)^*$ is named the "creation operator" associated with x.

(iii) For any finite sequence $x_1, \ldots, x_n \in G$, $n \in \mathbb{N}$, we define the element

$$x_1 \wedge \cdots \wedge x_n \doteq A(x_1)^* \cdots A(x_n)^* \Omega_{\mathrm{Fock}}^G \in H_{\mathrm{Fock}}^G \ .$$

Further, we use the following notation for the linear span of such vectors:

$$\mathfrak{G}_G \doteq \mathrm{span}\{\Omega_{\mathrm{Fock}}^G\} \cup \{x_1 \wedge \cdots \wedge x_n \ : \ x_1, \ldots, x_n \in G, \ n \in \mathbb{N}\} \subseteq H_{\mathrm{Fock}}^G \ .$$

Recall from Lemma 4.112 and Theorem 4.113 that any state on a C^*-algebra $\mathcal{A}$ is associated with a cyclic representation of $\mathcal{A}$ that is unique up to a unitary equivalence. In particular, the Fock space H_{Fock}^G is uniquely defined up to unitary equivalence.

Proposition 4.178 *Let G be any (nonempty) family of vectors of a complex pre-Hilbert space and H_{Fock}^G the associated Fock space. Then, $\mathfrak{G}_G \subseteq H_{\mathrm{Fock}}^G$ is a* *dense subspace of* H^G_{Fock} *and, for any two finite sequences* $x_1, \ldots, x_m \in G$ *and* $x'_1, \ldots, x'_n \in G$*,* $m, n \in \mathbb{N}$*, one has that*

$$\left\langle x_1 \wedge \cdots \wedge x_m, \Omega^G_{\text{Fock}} \right\rangle^G_{\text{Fock}} = 0 \qquad \textit{and}$$

$$\left\langle x_1 \wedge \cdots \wedge x_m, x'_1 \wedge \cdots \wedge x'_n \right\rangle^G_{\text{Fock}} = \det \left[\left\langle x_i, x'_j \right\rangle \right]^n_{i,j=1}$$

if $m = n$ *and* $\langle x_1 \wedge \cdots \wedge x_m, x'_1, \ldots, x'_n \rangle^G_{\text{Fock}} = 0$*, else. Here,* $\langle \cdot, \cdot \rangle^G_{\text{Fock}}$ *denotes the scalar product of the Hilbert space* H^G_{Fock}*, while* $\langle \cdot, \cdot \rangle$ *is the scalar product of the complex pre-Hilbert space containing* G*.*

Proof Exercise. □

The following fact directly follows from the last proposition.

Exercise 4.179 Let G be any (nonempty) family of vectors of a complex pre-Hilbert space and define

$$\widetilde{\mathfrak{G}}_G \doteq \text{span} \left(\{\mathfrak{1}\} \cup \{a(x_1)^* \cdots a(x_n)^* \ : \ x_1, \ldots, x_n \in G, \ n \in \mathbb{N}\} \right) \subseteq \text{CAR}(G) \ .$$

Show that the linear mapping $\widetilde{\mathfrak{G}}_G \to \mathfrak{G}_G$, $A \mapsto \pi^G_{\text{Fock}}(A)\Omega^G_{\text{Fock}}$, is a one-to-one correspondence.

Hint: Observe that any $A \in \widetilde{\mathfrak{G}}_G$ can be written as a linear combination of monomials of the form $a_{\text{span}(G)}(x_1)^* \cdots a_{\text{span}(G)}(x_n)^*$, $n \in \mathbb{N}$, where $x_1, \ldots, x_n \in \text{span}(G)$ is a orthonormal sequence, and use Proposition 4.178 together with basic properties of determinants.

It turns out that $\mathfrak{G}_G$ has naturally the structure of a Grassmann algebra, i.e., it is an algebra generated by a family of anticommuting elements. This follows from the last exercise.

Corollary 4.180 *Let* G *be any (nonempty) family of vectors of a complex pre-Hilbert space. Then, there is a unique product (i.e., a bilinear mapping)* $\wedge$: $\mathfrak{G}_G \times \mathfrak{G}_G \to \mathfrak{G}_G$ *such that, for any two finite sequences* $x_1, \ldots, x_m \in G$ *and* $x'_1, \ldots, x'_n \in G$*,* $m, n \in \mathbb{N}$*, one has that,*

$$(x_1 \wedge \cdots \wedge x_m) \wedge \Omega^G_{\text{Fock}} = \Omega^G_{\text{Fock}} \wedge (x_1 \wedge \cdots \wedge x_m) = x_1 \wedge \cdots \wedge x_m \ ,$$

$$(x_1 \wedge \cdots \wedge x_m) \wedge (x'_1 \wedge \cdots \wedge x'_n) = x_1 \wedge \cdots \wedge x_m \wedge x'_1 \wedge \cdots \wedge x'_n \ .$$

In particular, $\Omega^G_{\text{Fock}} \in \mathfrak{G}_G$ *is the unit of the algebra* $\mathfrak{G}_G$*.*

Proof The uniqueness of the product is clear. In order to prove its existence, consider $\widetilde{\mathfrak{G}}_G \subseteq \text{CAR}(G)$ as defined in the last exercise. Notice that $\widetilde{\mathfrak{G}}_G$ is the smallest subalgebra of $\text{CAR}(G)$ containing the family $\{\mathfrak{1}\} \cup \{a(x)^* : x \in G\} \subseteq \text{CAR}(G)$, i.e., $\widetilde{\mathfrak{G}}_G$ is the subalgebra of $\text{CAR}(G)$ generated by this family. As the elements of the family anticommute (by the CAR), i.e.,

$$a(x)^* a(x')^* = -a(x')^* a(x)^* , \qquad x, x' \in G ,$$

the unital algebra $\widetilde{\mathfrak{G}}_G$ is, by definition, a "Grassmann algebra." By Exercise 4.179, the linear mapping $A \mapsto \pi^G_{\mathrm{Fock}}(A)\Omega^G_{\mathrm{Fock}}$ identifies the vector spaces $\mathfrak{G}_G$ and $\widetilde{\mathfrak{G}}_G$. The algebra product of $\mathfrak{G}_G$ inherited from $\widetilde{\mathfrak{G}}_G$ via this identification has the required properties. □

We introduce in the following the notion of "Bogoliubov $*$-homomorphisms" between CAR C^*-algebras. To this end, we exploit the universal property of these algebras. See Definitions 4.127 and 4.163.

Definition 4.181 (Bogoliubov $*$-Homomorphisms) Let $G_1 \subseteq \tilde{H}_1$ and $G_2 \subseteq \tilde{H}_2$ be two nonempty families of vectors in complex pre-Hilbert spaces $\tilde{H}_1$ and $\tilde{H}_2$. For any function $u : G_1 \to G_2$ preserving the scalar product, i.e.,

$$\langle u(x), u(x') \rangle_{\tilde{H}_2} = \langle x, x' \rangle_{\tilde{H}_1} , \qquad x, x' \in G_1 ,$$

the $*$-homomorphism $\mathrm{Bog}(u) : \mathrm{CAR}(G_1) \to \mathrm{CAR}(G_2)$ is defined to be the unique morphism of C^*-representations, from $(\mathrm{CAR}(G_1), a)$ to $(\mathrm{CAR}(G_2), a \circ u)$. $\mathrm{Bog}(u)$ is called the "Bogoliubov $*$-homomorphisms" associated with the transformation u.

Observe that transformations $u : G_1 \to G_2$ preserving scalar products uniquely extend to isometries $\overline{\mathrm{span}(G_1)} \to \overline{\mathrm{span}(G_2)}$. Conversely, the restriction of such an isometry to G_1 preserves scalar products, by the polarization identity (Theorem 7.204). Thus, alternatively, we can equivalently define Bogoliubov $*$-homomorphisms via isometries (i.e., unitary transformations) between (pre)Hilbert spaces.

It turns out that Bogoliubov $*$-homomorphisms behave functorially, i.e., they commute with compositions.

Lemma 4.182 *Let G_1, G_2, G_3 be three nonempty families of vectors in complex pre-Hilbert spaces. For any functions $u : G_1 \to G_2$ and $u : G_2 \to G_3$ preserving the scalar product, one has the identity*

$$\mathrm{Bog}(u' \circ u) = \mathrm{Bog}(u') \circ \mathrm{Bog}(u) .$$

Proof The identity directly follows from the universality of $(\mathrm{CAR}(G_1), a)$. See Definitions 4.127 and 4.163. □

The lemma has the following important, albeit simple, corollary.

Corollary 4.183 *Let G be any nonempty family of vectors in a complex pre-Hilbert space. Let $\mathfrak{U}$ be any group of invertible transformations $G \to G$ preserving the scalar product. Then, the mapping*

$$\mathrm{Bog}(\cdot) : \mathfrak{U} \to \mathrm{Aut}(\mathrm{CAR}(G))$$

is a group homomorphism, i.e., $\mathrm{Bog}(u \circ u') = \mathrm{Bog}(u) \circ \mathrm{Bog}(u')$ *and* $\mathrm{Bog}(u^{-1}) = \mathrm{Bog}(u)^{-1}$ *for all* $u, u' \in \mathfrak{U}$*, where* $\mathrm{Aut}(\mathrm{CAR}(G))$ *denotes the group of* $*$*-automorphisms of* $\mathrm{CAR}(G)$ *(i.e., invertible* $*$*-isomorphisms* $\mathrm{CAR}(G) \to \mathrm{CAR}(G)$*).*

Note that, for any set $\mathfrak{A}$ of $*$-automorphisms of a C^*-algebra $\mathcal{A}$, the subset

$$\mathcal{A}_{\mathfrak{A}} \doteq \{A \in \mathcal{A} \ : \ \Theta(A) = A \text{ for all } \Theta \in \mathfrak{A}\} \subseteq \mathcal{A}$$

is a C^*-subalgebra $\mathcal{A}$. This follows from the continuity of $*$-homomorphisms of C^*-algebras. See Theorem 4.87. This motivates the following definition.

Definition 4.184 (Invariant Subalgebras of CAR C^*-Algebras) Let G be any nonempty family of vectors in a complex pre-Hilbert space. Let $\mathfrak{U}$ be any group of invertible transformations $G \to G$ preserving the scalar product. We define the unital C^*-subalgebra

$$\begin{aligned}\mathrm{CAR}_{\mathfrak{U}}(G) &\doteq \mathrm{CAR}(G)_{\mathrm{Bog}(\mathfrak{U})} \\ &= \{A \in \mathrm{CAR}(G) \ : \ \mathrm{Bog}(u)(A) = A \text{ for all } u \in \mathfrak{U}\} \subseteq \mathrm{CAR}(G) \ .\end{aligned}$$

The following example of groups of invertible transformations preserving the scalar products, along with the corresponding groups of Bogoliubov automorphisms, is very important for the physics of fermions.

Definition 4.185 (Gauge and Parity Automorphisms of CAR C^*-Algebras) Let G be any nonempty family of vectors in a complex pre-Hilbert space:

(i) Let the group $\mathfrak{U}_\circ$ of invertible transformations $\mathrm{span}(G) \to \mathrm{span}(G)$ that preserve the scalar product be defined by

$$\mathfrak{U}_\circ \doteq \{\mathrm{e}^{i\phi}\mathrm{id}_{\mathrm{span}(G)} \ : \ \phi \in [0, 2\pi)\} \ .$$

The Bogoliubov automorphisms $\alpha_\phi \doteq \mathrm{Bog}(\mathrm{e}^{i\phi}\mathrm{id}_{\mathrm{span}(G)})$, $\phi \in [0, 2\pi)$, of $\mathrm{CAR}(\mathrm{span}(G))$ are canonically seen as $*$-automorphisms of $\mathrm{CAR}(G)$ (by Proposition 4.170) and are named "(global) gauge automorphisms." The unital C^*-subalgebra

$$\mathrm{CAR}_\circ(G) \doteq \mathrm{CAR}_{\mathfrak{U}_\circ}(\mathrm{span}(G)) \subseteq \mathrm{CAR}(\mathrm{span}(G)) \equiv \mathrm{CAR}(G)$$

is the so-called gauge-invariant C^*-subalgebra of $\mathrm{CAR}(G)$, or the "gauge-invariant CAR C^*-algebra associated with G."

(ii) In the special case $\phi = \pi$, α_π is called the "parity automorphism" of $\mathrm{CAR}(G)$, and we define the subgroup:

$$\mathfrak{U}_\mathrm{e} \doteq \{\alpha_0, \alpha_\pi\} \subseteq \mathfrak{U}_\circ \ .$$

The "even subalgebra" of CAR(G) or "even CAR C^*-algebra associated with G" is the unital C^*-subalgebra:

$$\mathrm{CAR}_{\mathrm{e}}(G) \doteq \mathrm{CAR}_{\mathfrak{U}_{\mathrm{e}}}(\mathrm{span}(G)) \subseteq \mathrm{CAR}(\mathrm{span}(G)) \equiv \mathrm{CAR}(G) \ .$$

In particular, $\mathrm{CAR}_{\mathrm{e}}(G) \supseteq \mathrm{CAR}_{\circ}(G)$.

Note that $\mathrm{e}^{i\phi}G$, $\phi \in [0, 2\pi)$, may not be equal to G, but, of course, $\mathrm{span}(G) = \mathrm{e}^{i\phi}\mathrm{span}(G)$. This is why CAR(span($G$)) is used in the above definition. By the canonical identification of $(\mathrm{CAR}(\mathrm{span}(G)), a|_G)$ with $(\mathrm{CAR}(G), a)$ used above (see Proposition 4.170), the C^*-subalgebras $\mathrm{CAR}_{\circ}(G)$ and $\mathrm{CAR}_{\mathrm{e}}(G)$ of CAR(span(G)) are then viewed as C^*-subalgebras of CAR(G).

In quantum physics, the even CAR C^*-subalgebra is seen as more fundamental than the full CAR C^*-algebra, because of the local causality of quantum field theory. In particular, physically relevant Hamiltonians of fermionic systems are even, i.e., they are built in some even CAR C^*-algebra. In quantum statistical mechanics, these Hamiltonians are frequently (globally) gauge-invariant, and not only even, i.e., they refer to some gauge-invariant CAR C^*-algebra. This case corresponds to models conserving the particle number. Note, however, that non-gauge-invariant (approximating) models do appear, for instance, in the BCS theory of conventional superconductivity. See Sect. 6.6.2.

The next exercise gives a more concrete representation of even CAR C^*-algebras.

Exercise 4.186 Let G be any nonempty family of vectors in a complex pre-Hilbert space. Show that, for all $x \in G$, $\alpha_\pi(a(x)) = -a(x)$, where α_π is the parity automorphism of CAR(G) defined in Definition 4.185 (ii) and that $\mathrm{CAR}_{\mathrm{e}}(G)$ is the smallest C^*-subalgebra of CAR(G) such that

$$\{a(x)a(x') \ : \ x, x' \in G\} \cup \{a(x)^* a(x') \ : \ x, x' \in G\} \subseteq \mathrm{CAR}_{\mathrm{e}}(G) \ .$$

Prove additionally that, for any nonempty subset $G' \subseteq G$, one has

$$\mathfrak{i}_{GG'}(\mathrm{CAR}_{\mathrm{e}}(G')) = \mathrm{CAR}_{\mathrm{e}}(G) \cap i_{GG'}(\mathrm{CAR}(G')) \subseteq \mathrm{CAR}_{\mathrm{e}}(G) \subseteq \mathrm{CAR}(G) \ ,$$

where $\mathfrak{i}_{GG'}$ is the canonical inclusion $\mathrm{CAR}(G') \to \mathrm{CAR}(G)$.

Because of the property of even CAR C^*-algebras stated in the last exercise, for any nonempty subset $G' \subseteq G$, we canonically see $\mathrm{CAR}_{\mathrm{e}}(G')$ as a C^*subalgebra of $\mathrm{CAR}_{\mathrm{e}}(G)$ (and CAR(G)).

If the linear span of G is infinite dimensional, it turns out that $\mathrm{CAR}_{\mathrm{e}}(G)$ is a simple nuclear C^*-algebra, exactly as the full CAR C^*-algebra CAR(G) (Corollary 4.167). In order to prove this fact, we use the following property of even CAR C^*-algebras.

Lemma 4.187 *Let G be any nonempty orthonormal family of vectors in a complex pre-Hilbert space. For any nonempty proper*[23] *subset $G' \subsetneq G$ and any $e \in G \backslash G'$, there is a C^*-subalgebra of $\mathrm{CAR}_{\mathrm{e}}(G' \cup \{e\})$ containing the even C^*-algebra $\mathrm{CAR}_{\mathrm{e}}(G')$, which is $*$-isomorphic to the full CAR C^*-algebra $\mathrm{CAR}(G')$ associated with the index set G'.*

Proof Fix $G' \subsetneq G$ and $e \in G \backslash G'$. Define the mapping $b : G' \to \mathrm{CAR}_{\mathrm{e}}(G' \cup \{e\})$ by

$$b(e') \doteq a(e')(a(e)^* - a(e)) , \qquad e' \in G' .$$

(Compare this definition with the properties of the unique $*$-homomorphism stated in Proposition 4.172.) Observe that $(\mathrm{CAR}_{\mathrm{e}}(G' \cup \{e\}), b)$ is a C^*-representation of the CAR for the index set G'. Thus, by the universal property of the C^*-representation $\mathrm{CAR}(G')$ (Definitions 4.127 (iv)–(v) and 4.163), there is a unique morphism $\Theta : \mathrm{CAR}(G') \to \mathrm{CAR}_{\mathrm{e}}(G' \cup \{e\})$ of C^*-representations satisfying

$$\Theta(a(e')) = a(e')(a(e)^* - a(e)) .$$

It is nontrivial as the CAR are unital polynomial relations. See Definition 4.163. As $\mathrm{CAR}(G')$ is simple (Corollary 4.167), Θ is faithful and, thus,

$$\Theta(\mathrm{CAR}(G')) \subseteq \mathrm{CAR}_{\mathrm{e}}(G' \cup \{e\})$$

is a C^*-subalgebra that is $*$-isomorphic to $\mathrm{CAR}(G')$. For any $e', e'' \in G'$, observe from the CAR that

$$\Theta(a(e')a(e'')) = a(e')a(e'') , \quad \Theta(a(e')^* a(e'')) = a(e')^* a(e'') ,$$
$$\Theta(a(e')a(e'')^*) = a(e')a(e'')^* .$$

Thus, by Exercises 4.131 and 4.186, the restriction of Θ to $\mathrm{CAR}_{\mathrm{e}}(G')$ is nothing else than the canonical inclusion $*$-isomorphism $\mathrm{i}_{GG'} : \mathrm{CAR}(G') \to \mathrm{CAR}(G' \cup \{e\})$ (restricted to $\mathrm{CAR}_{\mathrm{e}}(G')$). In particular, by Exercise 4.186, $\Theta(\mathrm{CAR}_{\mathrm{e}}(G')) = \mathrm{CAR}_{\mathrm{e}}(G')$. □

Proposition 4.188 (The Even CAR C^*-Algebra Is Simple and Nuclear) *Let G be any nonempty family of vectors in a complex pre-Hilbert space. Then, $\mathrm{CAR}_{\mathrm{e}}(G)$ is a nuclear C^*-algebra. If $\mathrm{span}(G)$ has infinite dimension, then $\mathrm{CAR}_{\mathrm{e}}(G)$ is simple. If G is separable, then $\mathrm{CAR}_{\mathrm{e}}(G)$ is separable.*

Proof If $\mathrm{span}(G)$ has finite dimension, then, by Proposition 4.170, $\mathrm{CAR}(G)$ is $*$-isomorphic to the CAR C^*-algebra associated with a finite index set, by taking any basis of the vector space $\mathrm{span}(G)$. By Lemma 4.165, $\mathrm{CAR}(G)$ and $\mathrm{CAR}_{\mathrm{e}}(G)$ have

[23] That is, a subset which is not equal to the whole set G.

finite dimension, in this case, and are thus nuclear, thanks to [14, 11.3.11. Lemma]. Assume now that span(G) has infinite dimension. By Proposition 4.170, we may assume that G is an orthonormal family. Keeping in mind that any CAR C^*-algebra is simple (Corollary 4.167), we combine the last lemma with Lemma 4.165 to deduce, for every finite subset $\Lambda \in \mathcal{P}_f(G)\backslash\{\emptyset\}$, the existence of a finite-dimensional C^*-subalgebra $\mathcal{A}_\Lambda \subseteq \mathrm{CAR}_{\mathrm{e}}(G)$ such that $\mathrm{CAR}_{\mathrm{e}}(\Lambda) \subseteq \mathcal{A}_\Lambda$. From Exercise 4.186, the union of this family is dense in $\mathrm{CAR}_{\mathrm{e}}(G)$. It also follows from the last lemma that the family $\{\mathcal{A}_\Lambda\}_{\Lambda\in\mathcal{P}_f(G)\backslash\{\emptyset\}}$ forms a lattice[24] of C^*-subalgebras of $\mathrm{CAR}_{\mathrm{e}}(G)$. Thus, from Proposition 4.57, $\mathrm{CAR}_{\mathrm{e}}(G)$ is simple. From Proposition 4.160, this C^*-algebra is also nuclear. Recall from Exercise 4.164 that $\mathrm{CAR}(G)$ is separable, whenever G is separable. Observing that the linear mapping

$$\alpha_0 + \alpha_\pi : \mathrm{CAR}(G) \to \mathrm{CAR}_{\mathrm{e}}(G)$$

(see Definition 4.185 for the definition of the $*$-automrphisms α_0 and α_π) is continuous and surjective, we conclude that $\mathrm{CAR}_{\mathrm{e}}(G)$ is also separable whenever G is separable. □

To conclude this paragraph, we discuss some important general properties of states of even CAR C^*-algebras. In particular, we show that there is a natural notion of product states, similar to the case of universal tensor products of C^*-algebras. In fact, we prove a version of Proposition 4.155 for even CAR C^*-algebras. To this end, the following corollary of Proposition 4.172 is pivotal.

Corollary 4.189 *Let G be any (nonempty) orthonormal family of vectors of a complex pre-Hilbert space. Let $\mathfrak{P}$ be any collection of nonempty disjoint subsets of G whose union is G. That is, $\mathfrak{P}$ is a partition of G. Then, there is a unique $*$-homomorphism:*

$$\Theta_{\mathfrak{P}} : \bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}_{\mathrm{e}}(\Lambda) \to \mathrm{CAR}_{\mathrm{e}}(G)$$

such that, for all $\Lambda \in \mathfrak{P}$,

$$\Theta_{\mathfrak{P}}(a(\Lambda, A_\Lambda)) = \mathfrak{i}_{G\Lambda}(A_\Lambda)\ , \qquad A_\Lambda \in \mathrm{CAR}_{\mathrm{e}}(\Lambda)\ ,$$

a being the mapping of the C^*-representation $\bigotimes_{\Lambda\in\mathfrak{P}} \mathrm{CAR}(\Lambda)$ and $\mathfrak{i}_{G\Lambda}$ being the canonical inclusion $\mathrm{CAR}(\Lambda) \to \mathrm{CAR}(G)$. This $*$-homomorphism is faithful.

Proof Exercise. *Hint*: Use Proposition 4.172 to show that $\Theta_{\mathfrak{P}}$ is faithful. □

[24] Recall that a partially ordered set $(P, \succeq)$ is called a "lattice" if any two elements $p, p' \in P$ have an infimum and a supremum. In this case, take $P = \{\mathcal{A}_\Lambda\}_{\Lambda\in\mathcal{P}_f(G)\backslash\{\emptyset\}}$ endowed with the inclusion relation.

We now define the particular states of CAR C^*-algebras that are naturally identified with those of the corresponding even or gauge-invariant CAR C^*-subalgebras.

Definition 4.190 (Even and Gauge-Invariant States of CAR C^*-Algebras) Let G be any nonempty family of vectors in a complex pre-Hilbert space:

(i) A state $\rho \in E(\mathrm{CAR}(G))$ is "(globally) gauge-invariant" if $\rho \circ \alpha_\phi = \rho$ for all $\phi \in [0, 2\pi)$, where α_ϕ, $\phi \in [0, 2\pi)$, are the parity automorphisms of $\mathrm{CAR}(G)$ defined in Definition 4.185 (i). The set of all even states on $\mathrm{CAR}(G)$ is denoted by $E_\circ(\mathrm{CAR}(G))$.

(ii) A state $\rho \in E(\mathrm{CAR}(G))$ is said to be an "even state" on $\mathrm{CAR}(G)$, whenever $\rho \circ \alpha_\pi = \rho$, where α_π is the parity automorphism of $\mathrm{CAR}(G)$ defined in Definition 4.185 (ii). The set of all even states on $\mathrm{CAR}(G)$ is denoted by $E_\mathrm{e}(\mathrm{CAR}(G))$. Clearly, $E_\circ(\mathrm{CAR}(G)) \subseteq E_\mathrm{e}(\mathrm{CAR}(G))$.

The following example of a (globally) gauge-invariant state is very important.

Exercise 4.191 Show that the Fock state on any CAR C^*-algebra is (globally) gauge-invariant. In particular, it is a even state.

Even states of any CAR C^*-algebra are in one-to-one correspondence with the states of the corresponding even CAR C^*-algebra.

Lemma 4.192 *Let G be any nonempty family of vectors in a complex pre-Hilbert space. The restriction of states of* $\mathrm{CAR}(G)$ *to* $\mathrm{CAR}_\mathrm{e}(G)$ *(*$\mathrm{CAR}_\circ(G)$*) is a one-to-one correspondence between* $E_\mathrm{e}(\mathrm{CAR}(G))$ *and* $E(\mathrm{CAR}_\mathrm{e}(G))$ *(*$E_\circ(\mathrm{CAR}(G))$ *and* $E(\mathrm{CAR}_\circ(G))$*). If G is separable, this correspondence is an affine homeomorphism with respect to the weak* topology for states (see Sect. 4.5.1).*

Proof The restriction of states is clearly an affine mapping $E(\mathrm{CAR}(G)) \rightarrow E(\mathrm{CAR}_\mathrm{e}(G))$ that is continuous in the separable case. Recall that $\mathrm{CAR}(G)$ is separable when G is separable (Exercise 4.164). In particular, in this case, $E(\mathrm{CAR}(G))$ is a compact metric space (Proposition 4.84). By Exercise 7.173, the continuous mapping $E(\mathrm{CAR}(G)) \rightarrow E(\mathrm{CAR}_\mathrm{e}(G))$ is a homeomorphism whenever it is bijective. For any state $\rho \in E(\mathrm{CAR}_\mathrm{e}(G))$, the mapping $\mathrm{CAR}(G) \rightarrow \mathbb{C}$ defined by

$$A \mapsto \frac{1}{2}\rho\left(A + \alpha_\pi(A)\right) , \qquad A \in \mathrm{CAR}(G) ,$$

where α_π is the parity automorphism of $\mathrm{CAR}(G)$ defined in Definition 4.185 (ii), is an even state on $\mathrm{CAR}(G)$. To show this, note first that $A + \alpha_\pi(A) \in \mathrm{CAR}_\mathrm{e}(G)$ and the mapping is thus well-defined. In fact, this mapping is clearly a positive linear functional, and its operator norm is at most one, for $*$-homomorphisms of C^*-algebras are contractions (Proposition 4.97). Additionally,

$$\frac{1}{2}\rho\left(\mathfrak{1} + \alpha_\pi(\mathfrak{1})\right) = \mathfrak{1}$$

and thus, by Proposition 4.47 (iv) and Definition 4.190, the positive linear functional defined above is an even state on $\mathrm{CAR}(G)$. The restriction of this functional to $\mathrm{CAR}_{\mathrm{e}}(G)$ is the original state $\rho \in E(\mathrm{CAR}_{\mathrm{e}}(G))$, and, thus, any state on $\mathrm{CAR}_{\mathrm{e}}(G)$ is the restriction of some state on $\mathrm{CAR}(G)$. Conversely, if ρ is a even state on $\mathrm{CAR}(G)$, then

$$\frac{1}{2}\rho|_{\mathrm{CAR}_{\mathrm{e}}(G)}\left(A+\alpha_{\pi}(A)\right)=\frac{1}{2}\rho\left(A+\alpha_{\pi}(A)\right)=\rho(A)\ , \qquad A\in\mathrm{CAR}(G)\ .$$

Hence, the restriction operation is a one-to-one correspondence $E(\mathrm{CAR}(G)) \to E(\mathrm{CAR}_{\mathrm{e}}(G))$.

The proof of the lemma for the gauge-invariant case is done by means of a simple adaptation of the above arguments, by observing that, for all $A \in \mathrm{CAR}(G)$,

$$\frac{1}{2\pi}\int_0^{2\pi}\alpha_{\phi}(A)\mathrm{d}\phi\in\mathrm{CAR}_{\circ}(G)\ .$$

In fact, note that, for all $A \in \mathrm{CAR}(G)$, the mapping $\phi \mapsto \alpha_{\phi}(A)$, from $\mathbb{R}$ to $\mathrm{CAR}(G)$, is continuous and the above integral is thus well-defined as a Riemann integral taking values in a Banach space. □

The following result is a version of Proposition 4.155 for even states of CAR C^*-algebras.

Proposition 4.193 (Products of Even States) *Let G be any (nonempty) orthonormal family of vectors of a complex pre-Hilbert space. Let $\mathfrak{P}$ be any collection of nonempty disjoint subsets of G whose union is G. That is, $\mathfrak{P}$ is a partition of G. For every $\Lambda \in \mathfrak{P}$, let $\rho_\Lambda \in E_{\mathrm{e}}(\mathrm{CAR}(\Lambda))$ be an arbitrary* even *state on* $\mathrm{CAR}(\Lambda)$. *There is a unique even state $\otimes_{\Lambda\in\mathfrak{P}}\rho_\Lambda \in E_{\mathrm{e}}(\mathrm{CAR}(G))$ such that, for any finite sequences $\Lambda_1, \ldots, \Lambda_n \in \mathfrak{P}$ and*

$$A_1 \in \mathrm{CAR}(\Lambda_1), \ldots, A_n \in \mathrm{CAR}(\Lambda_n)\ ,$$

one has

$$(\otimes_{\Lambda\in\mathfrak{P}}\rho_\Lambda)(\mathrm{i}_{G\Lambda_1}(A_1)\cdots\mathrm{i}_{G\Lambda_n}(A_n))=\rho_{\Lambda_1}(A_1)\cdots\rho_{\Lambda_n}(A_n)\ ,$$

$\mathrm{i}_{G\Lambda}$ *being the canonical inclusion* $\mathrm{CAR}(\Lambda) \to \mathrm{CAR}(G)$. *This even state is called the "product state" associated with the family $\{\rho_\Lambda\}_{\Lambda\in\mathfrak{P}}$ of even states.*

Proof Observe that the uniqueness of the product state follows from the fact that the linear span of elements of the form

$$\mathrm{i}_{G\Lambda_1}(A_1)\cdots\mathrm{i}_{G\Lambda_n}(A_n)\ , \qquad A_1\in\mathrm{CAR}(\Lambda_1),\ldots,A_n\in\mathrm{CAR}(\Lambda_n),$$
$$\Lambda_1,\ldots,\Lambda_n\in\mathfrak{P},$$

is dense in $\mathrm{CAR}(G)$. This can be deduced from Exercise 4.131 and Corollary 4.132, as already done many times before. Let $\otimes_{\Lambda\in\mathfrak{P}}(\rho_\Lambda|_{\mathrm{CAR}_e(\Lambda)})$ be the product state on the universal tensor product $\bigotimes_{\Lambda\in\mathfrak{P}}\mathrm{CAR}_e(\Lambda)$ whose existence is stated in Proposition 4.155. By Corollary 4.189, $\bigotimes_{\Lambda\in\mathfrak{P}}\mathrm{CAR}_e(\Lambda)$ can be identified with a C^*-subalgebra of $\mathrm{CAR}_e(G)$ and, thus, $\otimes_{\Lambda\in\mathfrak{P}}(\rho_\Lambda|_{\mathrm{CAR}_e(\Lambda)})$ can now be seen as a state on this C^*-subalgebra. Let $\otimes_{\Lambda\in\mathfrak{P}}\rho_\Lambda$ be any extension of the latter to the whole $\mathrm{CAR}_e(G)$. Observe that such an extension exists, by Lemma 4.75 (i). Recall that we naturally identify $\otimes_{\Lambda\in\mathfrak{P}}\rho_\Lambda$ with an even state on $\mathrm{CAR}(G)$. Again by Corollary 4.189, this state has the required properties. □

4.8.3 Self-Dual CAR C^*-Algebras and Fermionic Quasi-Free States

We introduce in the present subsection another very useful viewpoint on CAR algebras. This approach originates from Araki's works [72] aiming at a mathematical setting for fermion systems, which is independent of the (global) gauge invariance of models. In other words, it allows the study of gauge and non-gauge-invariant fermion systems on the same basis. It refers to so-called *self-dual* CAR algebras. Similar to usual CAR algebras constructed from a complex (pre-)Hilbert space, self-dual CAR algebras are constructed from "self-dual Hilbert spaces," which are defined as follows.

Definition 4.194 (Self-Dual Hilbert Spaces)

(i) A complex Hilbert space H is a "self-dual Hilbert space" if it is a $*$-normed space in the sense of Definition 7.75, i.e., it is endowed with a complex conjugation that preserves the norm. In particular (from the polarization identity), one has the identity

$$\langle x, x'\rangle = \langle x'^*, x^*\rangle \ , \qquad x, x' \in H \ .$$

(ii) Let H_1 and H_2 be two self-dual Hilbert spaces. The unitary transformation $U : H_1 \to H_2$ is an "equivalence of self-dual Hilbert space" or a "Bogoliubov transformation" between these spaces, whenever U is self-conjugate in the sense of Definition 7.55, or a $*$-morphism (see Definition 7.54), i.e.,

$$U(x^*) = U(x)^* \ , \qquad x \in H \ .$$

In this case, H_1 and H_2 are said to be equivalent self-dual Hilbert spaces.

In a self-dual Hilbert space H, note that one can naturally decompose any vector $x \in H$ in its real and imaginary part, as is usual in any $*$-vector space:

$$\mathrm{Re}\{x\} \doteq \frac{1}{2}(x + x^*) \qquad \text{and} \qquad \mathrm{Im}\{x\} \doteq \frac{1}{2i}(x - x^*) .$$

The following example of a self-dual Hilbert space is probably the most important one.

Definition 4.195 (Self-Dual Hilbert Space Associated with a Complex Hilbert Space) Let H be any complex Hilbert space. Define a new (complex) Hilbert space by

$$H_{\mathrm{sd}} \doteq H \oplus_2 H^{\mathrm{td}} ,$$

where H^{td} denotes the topological dual space of H, i.e., the space of continuous linear forms on H, while the notation $H \oplus_2 H^{\mathrm{td}}$ stands for the vector space $H \times H^{\mathrm{td}}$ endowed with the norm

$$\|(x, \varphi)\| \doteq \sqrt{\|x\|_H^2 + \|x\|_{\mathrm{op}}^2} , \qquad x \in H , \ \varphi \in H^{\mathrm{td}} .$$

We define the complex conjugation of $H \oplus_2 H^{\mathrm{td}}$ as being the unique one satisfying

$$(x, \langle x', \cdot\rangle)^* = (x', \langle x, \cdot\rangle), \qquad x, x' \in H .$$

It is easy to check that the Hilbert space $H \oplus_2 H^{\mathrm{td}}$ is a (Hilbert) direct sum as defined in Definition 7.211. Observe meanwhile that the equality for the complex conjugation appearing in the definition uniquely defines a complex conjugation of $H \oplus_2 H^{\mathrm{td}}$, by the Riesz-Fréchet theorem (Theorem 7.214). As we will see below, up to a quite general technical condition, any self-dual Hilbert space is equivalent to some self-dual Hilbert space of the above form.

For any Hilbert space H, as already discussed, the space of bounded operators $\mathcal{B}(H)$ is a $*$-normed space (it is even a C^*-algebra, as we already know) with respect to the operation of taking adjoints of such operators. Additionally, if H is a self-dual Hilbert space, then $\mathcal{L}(H)$ has a natural complex conjugation, H being a $*$-vector space. See Definition 7.55. As H is a $*$-normed space, $\mathcal{B}(H) \subseteq \mathcal{L}(H)$ is a self-conjugate subspace with respect to this natural complex conjugation of $\mathcal{L}(H)$. See Exercise 7.56 (v). Thus, beyond the operation of taking adjoints, $\mathcal{B}(H)$ has a second complex conjugation, the one stemming from the $*$-normed space structure of H, which is here denoted by $(\cdot)^\star$ in order to avoid confusion with the usual adjoints of bounded operators, denoted as before by $(\cdot)^*$. Bounded operators on H, on which both complex conjugations act identically, up to a minus sign, are important in the theory of self-dual CAR C^*-algebras.

Definition 4.196 (Self-Dual Bounded Operators) Let H be any self-dual Hilbert space. A bounded operator $A \in \mathcal{B}(H)$ is a "self-dual (bounded) operator" if the equality $A^\star = -A^*$ is satisfied.

Similarly, orthogonal projectors in a self-dual Hilbert space, on which $(\cdot)^\star$ coincides with the operation of taking the orthogonal complement of ranges, are pivotal for the theory of self-dual CAR C^*-algebras.

Definition 4.197 (Basis Projections) Let H be any self-dual Hilbert space. An orthogonal projector $P \in \mathcal{B}(H)$ is called a "basis projection" if $P^\star \in \mathcal{B}(H)$ is the orthogonal projector whose range $\mathrm{ran}(P^\star)$ is the orthogonal complement of the range $\mathrm{ran}(P)$ of P, i.e., $P^\star = \mathrm{id}_H - P$.

Note that $\mathrm{id}_H^\star = \mathrm{id}_H$ (see Definition 7.55) and thus, $\mathrm{id}_H - P$ is a basis projection iff P is a basis projection.

Lemma 4.198 *Let H be a self-dual Hilbert space and $P \in \mathcal{B}(H)$ any basis projection. The mapping $x \mapsto \langle x^*, \cdot \rangle$ is a unitary transformation from $\mathrm{ran}(P)^\perp$ to $\mathrm{ran}(P)^{\mathrm{td}}$. There is a unique unitary transformation $U_P : H \to \mathrm{ran}(P)_{\mathrm{sd}}$ (see Definition 4.195) satisfying*

$$U_P(x + x') = x + \left\langle x'^*, \cdot \right\rangle , \qquad x \in \mathrm{ran}(P) \ , \ x' \in \mathrm{ran}(P)^\perp .$$

This unitary transformation is a Bogoliubov transformation.

Proof Exercise. □

From the last lemma, it follows, in particular, that a self-dual Hilbert space on which there is some basis projection is equivalent to a self-dual Hilbert space of the form given in Definition 4.195. Note additionally from the lemma that a finite-dimensional self-dual Hilbert space having a basis projection must be of *even* dimension.

Another consequence of the above lemma is that any two basis projections on the same self-dual Hilbert space are intertwined by some (unique) Bogoliubov transformation and are thus equivalent to each other in a quite strong sense.

Corollary 4.199 *Let H be a self-dual Hilbert space. For any two basis projections $P, P' \in \mathcal{B}(H)$, there is a unique Bogoliubov transformation $U : H \to H$ satisfying $P'U = UP$.*

Proof The uniqueness of U is clear. To prove its existence, note from the last lemma that the self-dual Hilbert spaces H, $\mathrm{ran}(P)_{\mathrm{sd}}$ and $\mathrm{ran}(P')_{\mathrm{sd}}$ are equivalent. It follows in particular from this fact that $\mathrm{ran}(P)$ and $\mathrm{ran}(P')$ are unitarily equivalent. To prove this property, one may use Hilbert bases. We omit the details. Let

$$V : H \to \mathrm{ran}(P)_{\mathrm{sd}} \qquad \text{and} \qquad V' : H \to \mathrm{ran}(P')_{\mathrm{sd}}$$

be Bogoliubov transformations having the property stated in Lemma 4.198, respectively for P and P'. Let $u : \mathrm{ran}(P) \to \mathrm{ran}(P')$ be any unitary transformation. Define further the unitary mapping:

$$\tilde{U} : \mathrm{ran}(P)_{\mathrm{sd}} \to \mathrm{ran}(P')_{\mathrm{sd}}$$

by

$$\tilde{U}(x, \langle x', \cdot\rangle) = (u(x), \langle u(x'), \cdot\rangle)\ , \qquad x, x' \in \mathrm{ran}(P)\ .$$

Observing that $\tilde{U}$ is a Bogoliubov transformation, the unitary operator $U \doteq (V')^*\tilde{U}V$, which is again a Bogoliubov transformation, has the required properties. □

We now prove that if the dimension of a self-dual Hilbert space is infinite or even, then there are basis projections for this space. This is a consequence of the following lemma.

Lemma 4.200 *Let H be a self-dual Hilbert space. There is a Hilbert basis B of H whose elements are all real, i.e., every $e \in B$ satisfies $e^* = e$.*

Proof For any non-zero vector $x \in H$ either $\|\mathrm{Re}\{x\}\| > 0$ or $\|\mathrm{Im}\{x\}\| > 0$. Thus, H contains at least one non-zero real vector. In particular the collection of all orthonormal families of *real* vectors of H is nonempty. By a simple application of Zorn's lemma, there is a maximal family in this collection, that is, some orthonormal family $B \subseteq H$ of real vectors that is not contained in some strictly bigger family of this type. By Lemma 7.218, one has to prove that the closure of the linear span of B is the whole space H. Thus, assume, by contradiction, that the orthogonal complement $B^\perp$ of B is not empty. Observe that $B^\perp$ is the same as the orthogonal complement of the closure of the linear span of this set, by Lemma 7.206 (iii). Take any non-zero $x \in B^\perp$ and note that $x^* \in B^\perp$, because the elements of B are all real. Thus, $\mathrm{Re}\{x\}, \mathrm{Im}\{x\} \in B^\perp$. Recalling that either $\|\mathrm{Re}\{x\}\| > 0$ or $\|\mathrm{Im}\{x\}\| > 0$, we conclude that $B^\perp$ contains a non-zero real vector, which would imply that B is not maximal in the collection of all orthonormal families of *real* vectors of H. □

Corollary 4.201 (Existence of Basis Projectors) *Let H be a self-dual Hilbert space. If the dimension of H is infinite or even, then H admits a basis projection.*

Proof Take any Hilbert basis B of H whose elements are real. Such a basis exists, by the last lemma. Notice that if B has a even or infinite number of elements, then there is a subset $C \subseteq B$ and a bijection $\xi : C \to C^c \doteq B\backslash C$. The existence of the mapping ξ is obvious if B has a finite even number of elements. If B is infinite, the existence of ξ follows from the fact that, in this case, B has the same cardinality as the disjoint union of two copies of it. For all $e \in C$, define the vector

$$f_e \doteq e + i\xi(e)\ .$$

Let P be the orthogonal projector of H whose range is the closure of the linear span of $\{f_e : e \in C\}$. Observing that this family of vectors is orthogonal, one has

$$P(e^*)^* = P(e)^* = \frac{1}{2}(e + i\xi(e))^* = e - P(e)$$

for every $e \in C$. Similarly, for every $e \in C$,

$$P(\xi(e)^*)^* = P(\xi(e))^* = \left(-\frac{i}{2}(e + i\xi(e))\right)^* = \frac{1}{2}(ie + \xi(e)) = \xi(e) - P(\xi(e)) \; .$$

In other words, for all $e \in B$, $P(e^*)^* = e - P(e)$. Hence, $P^\star = \mathrm{id}_H - P$ and P is thus a basis projection; see Definition 4.197. □

In the following we introduce the notion of a "basis symbol," which generalizes the one of a basis projection. We also discuss some of their important, albeit simple, properties. These objects play a crucial role in the theory of states of self-dual CAR C^*-algebras and, consequently, in statistical mechanics of fermions.

Definition 4.202 (Basis Symbols) Let H be any self-dual Hilbert space. A positive operator $S \in \mathcal{B}(H)$ is a "basis symbol" if its spectrum lies in the interval [0, 1], i.e., $\sigma(S) \subseteq [0, 1]$ and $S^\star = \mathrm{id}_H - S$. In particular, any basis projection is a basis symbol.

Recall that $\mathrm{id}_H^\star = \mathrm{id}_H$ and thus, $\mathrm{id}_H - S$ is a basis symbol iff S is a basis symbol. In the following lemma, we show that any basis symbol can be seen as a "diagonal block" of some basis projection for some enlarged self-dual Hilbert space.

Lemma 4.203 *Let H be any self-dual Hilbert space and define $\tilde{H} \doteq H \times H$, which is a self-dual Hilbert space with the norm*

$$\left\|(x, x')\right\| \doteq \sqrt{\|x\|^2 + \|x'\|^2}, \qquad x, x' \in H \; ,$$

and complex conjugation

$$(x, x')^* \doteq (x^*, -x'^*) \; , \qquad x, x' \in H \; .$$

Take a basis symbol $S \in \mathcal{B}(H)$ and define the operator $P_S \in \mathcal{B}(\tilde{H})$ by

$$P_S(x, x') \doteq \left(S(x) + S^{\frac{1}{2}}(\mathrm{id}_H - S)^{\frac{1}{2}}(x'), S^{\frac{1}{2}}(\mathrm{id}_H - S)^{\frac{1}{2}}(x) + x' - S(x')\right) \; , \quad x, x' \in H \; .$$

Then, P_S is a basis projection on $\tilde{H}$ for which

$$\left\langle (x, 0), P_S(x', 0)\right\rangle = \left\langle x, S(x')\right\rangle \; , \qquad x, x' \in H \; .$$

Proof The assertion follows from direct computations, using among other things the equality

$$(S^{\frac{1}{2}}(\mathrm{id}_H - S)^{\frac{1}{2}})^\star = (\mathrm{id}_H - S)^{\frac{1}{2}} S^{\frac{1}{2}} \; . \tag{4.13}$$

This last statement is deduced as follows: Since $S \in \mathcal{B}(H)$ is a basis symbol, $\sigma((\mathrm{id}_H - S)S) \subseteq [0,1]$ and $((\mathrm{id}_H - S)S)^* = (\mathrm{id}_H - S)S$. In other words, by Definition 4.36, $(\mathrm{id}_H - S)S \geq 0$ is a positive operator. As a consequence, there is a unique $A_S \in \mathcal{B}(H)^+$ such that $(\mathrm{id}_H - S)S = A_S^2$. See, e.g., Proposition 4.102. Note from the properties of the basis symbol S and the complex conjugation in H that

$$(A_S^\star)^2 = (A_S^2)^\star = ((\mathrm{id}_H - S)S)^\star = (\mathrm{id}_H - S)^\star S^\star = (\mathrm{id}_H - S)S \ ,$$

while $A_S \geq 0$ yields $A_S^\star \geq 0$. By uniqueness of the square root of a positive operator [51, Theorem 2.2.10], we arrive at the equality $A_S^\star = A_S$, which can be rewritten as (4.13). □

After this brief review on important basic properties of self-dual Hilbert spaces, we are in a position to define and study the so-called self-dual CAR polynomial relations.

Definition 4.204 (CAR C^*-Algebra Associated with a Self-Dual Hilbert Space) Let H be any self-dual Hilbert space. We take H as being an index set and define the following family of relations:

$$\begin{aligned}\mathfrak{R} \doteq \{a(\alpha x + x') = \bar{\alpha} a(x) + a(x') : \ x, x' \in H, \ \alpha \in \mathbb{C}\ \} \cup \{a(x^*) = a(x)^*\} \\ \cup \{a(x')a(x) + a(x)a(x') = \langle x', x^* \rangle \mathfrak{1} : x, x' \in H\} \ .\end{aligned}$$

(a being again a mapping from H to some C^*-algebra).

(i) The "self-dual CAR C^*-algebra" associated with the self-dual Hilbert space H is

$$\mathrm{sCAR}(H) \doteq (C^*(H, \mathfrak{R}), a) \ ,$$

where $(C^*(H, \mathfrak{R}), a)$ is any fixed universal C^*-representation of $\mathfrak{R}$. By a slight abuse of notation, sometimes $\mathrm{sCAR}(H)$ also denotes only the universal C^*-algebra $C^*(H, \mathfrak{R})$.

(ii) For any closed self-conjugate subspace $G \subseteq H$, $\mathfrak{i}_{HG}$ denotes the natural inclusion

$$\mathrm{sCAR}(G) \to \mathrm{sCAR}(H),$$

i.e., the unique morphism of C^*-representations from $(\mathrm{sCAR}(G), a)$ to $(\mathrm{sCAR}(H), a|_G)$.

The first set of relations in the last definition is just saying that a is an antilinear mapping, exactly as in the (usual) CAR case. The second set, along with the first one, says that $a(\cdot)^*$ is a $*$-morphism of $*$-vector states.

In mathematical and theoretical physics, self-dual CAR C^*-algebras, similar to the usual CAR C^*-algebras, are associated with fermionic particles. They are particularly useful in dealing with states of these particles that break the (global) gauge invariance, as in the BCS theory of superconductivity. In fact, as C^*-algebras the self-dual CAR C^*-algebras are equivalent to (usual) CAR C^*-algebras, but the presentation of these C^*-algebras as universal C^*-algebras of polynomial self-dual CAR is particularly useful in many important situations.

Exercise 4.205 Prove that the self-dual Hilbert space H is separable only if the C^*-algebra $\mathrm{sCAR}(H)$ is separable.

Like in the previous examples of universal C^*-algebras (universal tensor products and CAR C^*-algebras), Definition 4.204 assumes the existence of a universal C^*-representation of the above-defined family $\mathfrak{R}$ of polynomial relations. By Theorem 4.134, $\mathfrak{R}$ must be admissible, i.e., a C^*-representation $(\mathcal{A}, a)$ must exist and, for all $x \in H$,

$$\sup\{\|a(x)\|_{\mathcal{A}} \ : \ (\mathcal{A}, a) \text{ a } C^*\text{-representation of } \mathfrak{R}\} < \infty \ . \tag{4.14}$$

The last property is again easily deduced from the polynomial relations: Take any C^*-representation $(\mathcal{A}, a)$ of the family $\mathfrak{R}$ of relation of the last definition. Then, for all $x \in H$,

$$a(x)a(x)^* + a(x)^*a(x) = a(x)a(x^*) + a(x^*)a(x) = \|x\|^2 \mathfrak{1} \ .$$

Exactly as in the case of the (usual) CAR, this implies that, for all $x \in H$,

$$\|a(x)\|^2 = \left\|a(x)^*a(x)\right\| \leq \|x\|^2 \ .$$

Hence, (4.14) holds true for all $x \in H$.

If the self-dual Hilbert space admits a basis projection, then one can easily construct a C^*-representation of the self-dual CAR, showing that the family $\mathfrak{R}$ of Definition 4.204 is admissible. Then, from this special case, we deduce that $\mathfrak{R}$ is admissible in any case.

Lemma 4.206 (The Self-Dual CAR Are Admissible) *Let H be any self-dual Hilbert space. The family $\mathfrak{R}$ of polynomial relations defined in Definition 4.204 is admissible and unital.*

Proof $\mathfrak{R}$ is clearly unital. Assume first that H admits a basis projection P. Define, in this case, a mapping $\tilde{a} : H \to \mathrm{CAR}(\mathrm{ran}(P))$ by

$$\tilde{a}(x) \doteq a(P(x)) + a((x - P(x))^*)^* \ , \qquad x \in H \ .$$

Observe that $(x - P(x))^* \in \mathrm{ran}(P)$ for every $x \in H$, because P is a basis projection (Definition 4.197). Then, by the CAR (Definition 4.163) and the fact that P is a basis projection, we obtain that

$$\begin{aligned}\tilde{a}(x)\tilde{a}(x') + \tilde{a}(x')\tilde{a}(x) &= a(P(x))a((x' - P(x'))^*)^* + a((x' - P(x'))^*)^* a(P(x)) \\ &\quad + a((x - P(x))^*)^* a(P(x')) + a(P(x'))a((x - P(x))^*)^* \\ &= \left\langle P(x), (x' - P(x'))^* \right\rangle \mathbb{1} + \left\langle P(x'), (x - P(x))^* \right\rangle \mathbb{1} \\ &= \left\langle P(x), P(x'^*) \right\rangle \mathbb{1} + \left\langle P(x'), P(x^*) \right\rangle \mathbb{1} \\ &= \left\langle P(x'^*)^*, P(x)^* \right\rangle \mathbb{1} + \left\langle P(x'), P(x^*) \right\rangle \mathbb{1} \\ &= \left\langle x' - P(x'), x^* - P(x^*) \right\rangle \mathbb{1} \\ &\quad + \left\langle P(x'), P(x^*) \right\rangle \mathbb{1} = \left\langle x', x^* \right\rangle \mathbb{1}\end{aligned}$$

Thus, $(\mathrm{CAR}(\mathrm{ran}(P)), \tilde{a})$ is a C^*-representation of the self-dual CAR. By Corollary 4.201, H admits a basis projection if its dimension is infinite or (finite) even. Thus, $\mathfrak{R}$ is admissible in this two cases. If H has (finite) odd dimension, then we can canonically construct a new self-dual Hilbert space by considering the direct sum of two copies of H and extending the complex conjugation in an obvious way. This leads to a self-dual Hilbert space whose dimension is even and contains H as a self-conjugate subspace. See the definition of the self-dual Hilbert space $\tilde{H}$ in Lemma 4.203. Thus, we get in this manner a C^*-representation of the self-dual CAR for H, by restriction of the one for the doubled space. Hence, $\mathfrak{R}$ is admissible in any case. □

Corollary 4.207 (From CAR to Self-Dual CAR) *For any complex Hilbert space H', there is a unique $*$-isomorphism $\Theta : \mathrm{CAR}(H') \to \mathrm{sCAR}(H'_{\mathrm{sd}})$ satisfying $\Theta(a(x) + a(x')^*) = a(x + \langle x', \cdot \rangle)$ for all $x, x' \in H$.*

Proof Let H' be any complex Hilbert space. Then, by the self-dual CAR, $(\mathrm{sCAR}(H'_{\mathrm{sd}}), a|_{H'})$ is a C^*-representation of the (usual) CAR for the index set H'. Let $\Theta : \mathrm{CAR}(H') \to \mathrm{sCAR}(H'_{\mathrm{sd}})$ be the unique $*$-homomorphism satisfying $\Theta(a(x)) = a(x)$ for all $x \in H'$. See Definition 4.127 (iv)–(v). Then

$$\Theta(a(x)^*) = \Theta(a(x))^* = a(x)^* = a(\langle x, \cdot \rangle)$$

for all $x \in H' \subseteq H'_{\mathrm{sd}}$. See Definition 4.195. Thus,

$$\Theta(a(x) + a(x')^*) = a(x + \left\langle x', \cdot \right\rangle) , \qquad x, x' \in H' .$$

Observe additionally that Θ is faithful, because $\mathfrak{R}$ is unital and $\mathrm{CAR}(H')$ is simple (Corollary 4.167). Meanwhile, Θ is also surjective, because its image is a C^*-subalgebra of $\mathrm{sCAR}(H'_{\mathrm{sd}})$ containing $\{a(x) : x \in H'_{\mathrm{sd}}\}$, which, by Corollary 4.132, is a family generating this (universal) C^*-algebra. □

Corollary 4.208 (The Self-Dual CAR Are Simple and Nuclear) *Let H be any self-dual Hilbert space for which a basis projection $P \in \mathcal{B}(H)$ exists. Then, $\mathfrak{R}$ is simple and nuclear. Addionally,* $\mathrm{sCAR}(H)$ *is $*$-isomorphic to* $\mathrm{CAR}(\mathrm{ran}(P))$.

Proof To prove that $\mathfrak{R}$ is simple and nuclear, it suffices to combine Lemma 4.206 with Corollary 4.167. Finally, by Lemma 4.198, if H admits a basis projection P, then it is equivalent to $\mathrm{ran}(P)_{\mathrm{sd}}$ as a self-dual Hilbert space. Using Corollary 4.207, one easily verifies that $\mathrm{sCAR}(H)$ is $*$-isomorphic to $\mathrm{CAR}(\mathrm{ran}(P))$. □

We call the $*$-isomorphism Θ of Corollary 4.207 the "natural $*$-isomorphism $\Theta : \mathrm{CAR}(H') \to \mathrm{sCAR}(H'_{\mathrm{sd}})$." Similar to the case of (usual) CAR C^*-algebras, we may define (Bogoliubov) $*$-homomorphisms of self-dual CAR C^*-algebras via isometries of the underlying Hilbert spaces.

Definition 4.209 (Self-Dual Bogoliubov $*$-Homomorphisms) Let H_1 and H_2 be any two self-dual Hilbert spaces. The self-dual Bogoliubov $*$-homomorphism associated with a $*$-morphism $U : H_1 \to H_2$ preserving the scalar product[25] is the $*$-homomorphism

$$\mathrm{Bog}(U) : \mathrm{sCAR}(H_1) \to \mathrm{sCAR}(H_2)$$

defined to be the unique morphism between the C^*-representations $(\mathrm{sCAR}(H_1), a)$ and $(\mathrm{sCAR}(H_2), a \circ u)$, both being C^*-representations of the self-dual CAR for the index set H_1. Note that $\mathrm{Bog}(U)$ is a $*$-isomorphism, whenever U is a Bogoliubov transformation, i.e., U is unitary.

Exactly as in the (usual) CAR case, self-dual Bogoliubov $*$-homomorphisms behave functorially, i.e., they commute with compositions. See Lemma 4.182. As in the previous case (Corollary 4.183), one has the following important, albeit simple, corollary.

Corollary 4.210 *Let H be any self-dual Hilbert space and $\mathfrak{U} \subseteq \mathcal{B}(H)$ any group of Bogoliubov transformations (see Definition 4.194). Then,*

$$\mathrm{Bog}(\cdot) : \mathfrak{U} \to \mathrm{Aut}(\mathrm{sCAR}(H))$$

is a group homomorphism, where $\mathrm{Aut}(\mathrm{CAR}(H))$ *denotes the group of automorphism of* $\mathrm{CAR}(H)$ *(i.e., invertible $*$-isomorphisms* $\mathrm{CAR}(H) \to \mathrm{CAR}(H)$*).*

Similar to Definition 4.184, we use Bogoliubov transformations and self-dual Bogoliubov $*$-homomorphisms to define important subalgebras of self-dual CAR C^*-algebras.

Definition 4.211 (Invariant Subalgebras of Self-Dual CAR C^*-Algebras) Let H be any self-dual Hilbert space and $\mathfrak{U} \subseteq \mathcal{B}(H)$ any group of Bogoliubov transformations. We define the unital C^*-subalgebra

$$\mathrm{sCAR}_{\mathfrak{U}}(H) \doteq \{A \in \mathrm{sCAR}(H) \ : \ \mathrm{Bog}(u)(A) = A \text{ for all } u \in \mathfrak{U}\} \subseteq \mathrm{sCAR}(H)\ .$$

[25] U is linear and satisfies $U(x^*) = U(x)^*$ and $\langle U(x), U(x')\rangle_{H_1} = \langle x, x'\rangle_{H_1}$ for all $x, x' \in H_1$.

If $\mathfrak{U} = \{\mathrm{id}_H, -\mathrm{id}_H\}$, then $\mathrm{CAR}_{\mathfrak{U}}(\mathrm{H})$ is denoted by $\mathrm{CAR}_{\mathrm{e}}(H)$. This unital C^*-subalgebra is called here the "even subalgebra" of $\mathrm{sCAR}(G)$ or "even self-dual CAR C^*-algebra associated with H."

For unitary transformations of the form $\mathrm{e}^{i\phi}\mathrm{id}_H$, $\phi \in [0, 2\pi)$, on an arbitrary self-dual Hilbert space H, notice that only the two choices $\phi \in \{0, \pi\}$ give a Bogoliubov transformation, because

$$(\mathrm{e}^{i\phi}\mathrm{id}_H)^\star = \mathrm{e}^{-i\phi}\mathrm{id}_H \neq \mathrm{e}^{i\phi}\mathrm{id}_H \ , \qquad \phi \in (0, \pi) \cup (\pi, 2\pi) \ .$$

Thus, the gauge automorphisms α_ϕ, $\phi \in [0, 2\pi)$, of Definition 4.185 (i) for the CAR case have no obvious analogue in the sCAR case. In fact, as explained above, the sCAR formalism was proposed, among other things, in order to describe both gauge-invariant and non-gauge-invariant fermion systems on the same basis.

The next exercise gives a more concrete representation of even self-dual CAR C^*-algebras.

Exercise 4.212 Let H be any self-dual Hilbert space. Show the following assertions:

(i) For all $x \in H$, $\mathrm{Bog}(\pm\mathrm{id}_H)(a(x)) = \pm a(x)$.
(ii) $\mathrm{sCAR}_{\mathrm{e}}(H)$ is the smallest C^*-subalgebra of $\mathrm{sCAR}(H)$ containing the set

$$\{a(x)a(x') \ : \ x, x' \in H\} \subseteq \mathrm{sCAR}_{\mathrm{e}}(G) \ .$$

(iii) For any closed self-conjugate subspace $G \subseteq H$, one has

$$\mathrm{i}_{HG}(\mathrm{sCAR}_{\mathrm{e}}(G)) = \mathrm{sCAR}_{\mathrm{e}}(H) \cap i_{HG}(\mathrm{sCAR}(G)) \subseteq \mathrm{sCAR}_{\mathrm{e}}(H) \subseteq \mathrm{sCAR}(H) \ ,$$

where i_{HG} is the natural inclusion $\mathrm{sCAR}(G) \to \mathrm{sCAR}(H)$.
(iv) For any complex Hilbert space H', the natural $*$-isomorphism $\mathrm{CAR}(H') \to \mathrm{sCAR}(H'_{\mathrm{sd}})$ (see Corollary 4.207) defines, by restriction, a $*$-isomorphism $\mathrm{CAR}_{\mathrm{e}}(H') \to \mathrm{sCAR}_{\mathrm{e}}(H'_{\mathrm{sd}})$.

Because of Exercise 4.212 (iii), for any closed self-conjugate subspace $G \subseteq H$, we canonically see $\mathrm{sCAR}_{\mathrm{e}}(G)$ as a C^*subalgebra of $\mathrm{sCAR}_{\mathrm{e}}(H)$ (and $\mathrm{sCAR}(H)$). One can further verify from Exercise 4.212 (ii) and (iv) and Proposition 4.188 combined with Corollaries 4.201 and 4.208 that, if H is infinite dimensional, then $\mathrm{sCAR}_{\mathrm{e}}(H)$ is a simple nuclear C^*-algebra, exactly as the full self-dual CAR C^*-algebra $\mathrm{sCAR}(H)$ (see again Corollaries 4.201 and 4.208).

In the following we discuss important properties of the states of even self-dual CAR C^*-algebras. In particular, we will introduce so-called quasi-free states, which are seen in theoretical physics as the states corresponding to free (or noninteracting) fermion gases at equilibrium. Also in the theory of interacting fermions at equilibrium, this type of state plays an important role, as it provides a very efficient approximation procedure, known as the "Hartree-Fock" theory. This method will be presented and discussed in some detail in Sects. 5.7 and 6.10.

We start by defining even states on self-dual CAR C^*-algebras, similar to what is done for states on CAR C^*-algebras (Definition 4.190).

Definition 4.213 (Even States of Self-Dual CAR C^*-Algebras) Let H be any self-dual Hilbert space. An arbitrary state $\rho \in E(\mathrm{sCAR}(H))$ is said to be an "even state" of $\mathrm{sCAR}(H)$ if $\rho \circ \mathrm{Bog}(-\mathrm{id}_H) = \rho$. The set of all even states of $\mathrm{sCAR}(G)$ is denoted by $E_\mathrm{e}(\mathrm{sCAR}(H))$.

Exactly as in the case of (usual) CAR C^*-algebra, the even states of any self-dual CAR C^*-algebra are in one-to-one correspondence with the states of the corresponding even self-dual CAR C^*-algebra, via the restriction of states of the full self-adjoint CAR C^*-algebras to the corresponding even subalgebras. In other words, a version of Lemma 4.192 holds true for self-dual CAR C^*-algebras. With this identification of states, by using the natural $*$-isomorphism $\mathrm{CAR}(H) \to \mathrm{sCAR}(H_\mathrm{sd})$ (see Corollary 4.207) for any complex Hilbert space H, we infer from Proposition 4.193 the following corollary.

Corollary 4.214 (Products of Even States) *Let H be any self-dual Hilbert space. Let $\{G_i\}_{i\in I}$ be any family of closed self-conjugate subspaces of H that are mutually orthogonal and span the whole space H, that is, the linear span of the union $\cup\{G_i : i \in I\} \subseteq H$ is dense in H. For any family $\rho_i \in E_\mathrm{e}(\mathrm{sCAR}(G_i))$, $i \in I$, of even states, there is a unique even state $\otimes_{i\in I}\rho_i \in E_\mathrm{e}(\mathrm{sCAR}(H))$ of $\mathrm{sCAR}(H)$ such that, for any finite sequences $i_1, \ldots, i_n \in I$ and*

$$A_1 \in \mathrm{sCAR}(G_{i_1}), \ldots, A_n \in \mathrm{CAR}(G_{i_n}) ,$$

one has

$$(\otimes_{i\in I}\rho_i)(\mathfrak{i}_{HG_{i_1}}(A_1)\cdots \mathfrak{i}_{HG_{i_n}}(A_n)) = \rho_{i_1}(A_1)\cdots \rho_{i_n}(A_n) .$$

This even state is called the "product state" associated with the family $\{\rho_i\}_{i\in I}$ of even states.

The following corollary of Proposition 4.176 gives another important example of even states of self-dual CAR C^*-algebras.

Corollary 4.215 (Fock States of Self-Dual CAR C^*-Algebras) *Let H be any self-dual Hilbert space and P any basis projection on this space. Then, there is a unique state $\rho^P_{\mathrm{Fock}} \in E(\mathrm{sCAR}(H))$ such that, for all $x \in \mathrm{ran}(P) \subseteq H$ and $A \in \mathrm{sCAR}(H)$, one has $\rho^P_{\mathrm{Fock}}(Aa(x)) = 0$. This state is even and satisfies $\rho^P_{\mathrm{Fock}}(a(x)A) = 0$ for all $x \in \mathrm{ran}(P)^\perp \subseteq H$ and $A \in \mathrm{sCAR}(H)$.*

Proof To prove the corollary, besides Proposition 4.176, use the natural identification of the C^*-algebras $\mathrm{sCAR}(\mathrm{ran}(P)_\mathrm{sd})$ and $\mathrm{CAR}(\mathrm{ran}(P))$, as well as the identification of the C^*-algebras $\mathrm{sCAR}(\mathrm{ran}(P)_\mathrm{sd})$ and $\mathrm{sCAR}(H)$ via the self-dual Bogoliubov $*$-isomorphism $\mathrm{Bog}(U_P)$, where U_P is the Bogoliubov transformation of Lemma 4.198. We omit the details. □

We call the state $\rho_{\mathrm{Fock}}^{P} \in E(\mathrm{sCAR}(H))$ of the above corollary the "self-dual Fock state" associated with the basis projection P.

To conclude the present section, we introduce and discuss important basic properties of so-called quasi-free states of even usual and self-dual CAR C^*-algebras. We start by defining basis symbols associated with states of these algebras.

Exercise 4.216 Let H be any self-dual Hilbert space. Let $\rho \in E(\mathrm{sCAR}(H))$ be any state on $\mathrm{sCAR}(H)$. Show that there is a unique basis symbol (see Definition 4.202) $S_\rho \in \mathcal{B}(H)$ such that the following equality is satisfied:

$$\rho(a(x)a(x')) = \left\langle x, S_\rho(x'^*)\right\rangle , \qquad x, x' \in H .$$

Hint: Use the Riesz representation theorem for sesquilinear forms (Corollary 7.216). We call $S_\rho \in \mathcal{B}(H)$ the "basis symbol associated with the state $\rho \in E(\mathrm{sCAR}(H))$."

Using the definition

$$\mathbb{O}_{i,j}(A_1, A_2) \doteq \begin{cases} A_1A_2 & \text{for } i < j, \\ -A_2A_1 & \text{for } i > j, \\ 0 & \text{for } i = j, \end{cases}$$

for elements $A_1, A_2 \in \mathcal{U}$ and natural numbers $i, j \in \mathbb{N}$, we define quasi-free states from Pfaffians[26]

$$\mathrm{Pf}\left[M_{i,j}\right]_{i,j=1}^{2n} \doteq \frac{1}{2^n n!} \sum_{\pi \in \Pi_{2n}} (-1)^\pi \, M_{\pi(1)\pi(2)} \cdots M_{\pi(2n-1)\pi(2n)}$$

of skew-symmetric $2n \times 2n$ matrices as follows.

Definition 4.217 (Quasi-Free States of Self-Dual CAR C^*-Algebras) Let H be any self-dual Hilbert space. We say that the even state $\rho \in E_{\mathrm{e}}(\mathrm{sCAR}(H))$ is a "quasi-free state" if, for every finite sequence of vectors $x_1, \ldots, x_{2n} \in H$, $n \in \mathbb{N}$, one has that

$$\rho(a(x_1)\cdots a(x_{2n})) = \mathrm{Pf}[\rho(\mathbb{O}_{i,j}(a(x_i), a(x_j)))]_{i,j=1}^{2n} .$$

It turns out that self-dual Fock states are quasi-free states whose basis symbols are the corresponding basis projections.

Exercise 4.218 Let H be any self-dual Hilbert space and P any basis projection on this space. Show that the self-dual Fock state $\rho_{\mathrm{Fock}}^{P} \in E(\mathrm{sCAR}(H))$ is the unique quasi-free state on $\mathrm{sCAR}(H)$ whose basis symbol is P.

[26] Here, Π_{2n} denotes the set of all permutations of $\{1, \ldots, 2n\}$ and $(-1)^\pi$ the sign of the permutation $\pi \in \Pi_{2n}$.

By combining the last exercise with Lemma 4.203, we obtain that quasi-free states of any self-dual CAR C^*-algebra $\mathrm{sCAR}(H)$ are in one-to-one correspondence to the basis symbols on H:

Proposition 4.219 (Quasi-Free States) *Let H be any self-dual Hilbert space and S any basis symbol on this space. There is a unique quasi-free state $\rho^S \in E(\mathrm{sCAR}(H))$ of $\mathrm{sCAR}(H)$ whose basis symbol is S.*

Proof The proposition directly follows from the last exercise, along with Lemma 4.203, by identifying $\mathrm{sCAR}(H)$ with the C^*-subalgebra of $\mathrm{sCAR}(H \oplus_2 H)$ associated with the closed self-conjugate subspace corresponding to the first component of the direct sum $H \oplus_2 H$. □

For any family G of vectors in a complex pre-Hilbert space, let $(H, \mathfrak{i})$ be any fixed completion of $\mathrm{span}(G)$. Recall that $\mathrm{CAR}(G)$ is naturally identified with $\mathrm{CAR}(H)$ via the unique morphism of C^*-representation from $(\mathrm{CAR}(G), a)$ to $(\mathrm{CAR}(H), a \circ \mathfrak{i}|_G)$. In turn, $\mathrm{CAR}(H)$ is naturally identified with $\mathrm{sCAR}(H_{\mathrm{sd}})$. Thus, one may use these identifications of C^*-algebras to assign a basis symbol S_ρ on H_{sd} to any state $\rho \in E(\mathrm{CAR}(G))$ of $\mathrm{CAR}(G)$, as well as a (quasi-free) state ρ^S of $\mathrm{CAR}(G)$ to any basis symbol $S \in \mathcal{B}(H_{\mathrm{sd}})$. We call these states the "quasi-free states" of $\mathrm{CAR}(G)$.

Basis symbols $S \in \mathcal{B}(H_{\mathrm{sd}})$ that satisfy

$$\left\langle x, S(x'^*) \right\rangle = 0 , \qquad x, x' \in H \subseteq H_{\mathrm{sd}},$$

can be uniquely associated with a positive operator on $H \subseteq H_{\mathrm{sd}}$ whose spectrum lies in the interval $[0, 1]$:

Lemma 4.220 *Let H be any complex Hilbert space. Let $S \in \mathcal{B}(H_{\mathrm{sd}})$ be any basis symbol satisfying $\left\langle x, S(x'^*) \right\rangle = 0$ for all $x, x' \in H \subseteq H_{\mathrm{sd}}$. Then S is completely determined by $PSP \in \mathcal{B}(H)$, where P is the basis projection whose range is $H \subseteq H_{\mathrm{sd}}$. The operator PSP is positive and its spectrum lies in the interval $[0, 1]$. Conversely, if $s \in \mathcal{B}(H)$ is such an operator, then there is a unique basis symbol $S_s \in \mathcal{B}(H_{\mathrm{sd}})$ satisfying $PS_sP = s$ and $\left\langle x, S_s(x'^*) \right\rangle = 0$ for all $x, x' \in H \subseteq H_{\mathrm{sd}}$.*

Proof Exercise. □

If the quasi-free state $\rho \in E(\mathrm{CAR}(G))$ has an associated basis symbol as in the lemma, we say that it is a "simple" quasi-free state of $\mathrm{CAR}(G)$. We say that ρ is an "extended" quasi-free state, otherwise. Thus, the lemma is saying that simple quasi-free states of $\mathrm{CAR}(G)$ are in one-to-one correspondence to the positive operators on a completion of $\mathrm{span}(G)$, whose spectrum lies in the interval $[0, 1]$. We call here such operators "simple symbols" on G. The following proposition gives a more concrete characterization of simple quasi-free states.

Proposition 4.221 (Simple Quasi-Free States of CAR C^*-Algebras) *Let G be any (nonempty) family of vectors in a complex pre-Hilbert space and let $(H, \mathfrak{i})$ be any completion of $\mathrm{span}(G)$:*

(i) *A quasi-free state* $\rho \in E(\mathrm{CAR}(G))$ *is simple iff, for all* $x, x' \in G$, $\rho(a(x)a(x')) = 0$.

(ii) *For any simple symbol* $s \in \mathcal{B}(H)$ *on* G*, there is a unique simple quasi-free state* $\rho^s \in E(\mathrm{CAR}(G))$ *satisfying, for all* $x_1, \ldots, x_n, x'_1, \ldots, x'_n \in G$, $n \in \mathbb{N}$,

$$\rho(a(x_1) \cdots a(x_n)a(x'_n)^* \cdots a(x'_1)^*) = \det\left[\left\langle \mathrm{i}(x_i), s \circ \mathrm{i}(x'_j)\right\rangle_H\right]_{i,j=1}^n .$$

This state is (globally) gauge-invariant (see Definition 4.190). In particular, for any $m, n \in \mathbb{N}$ *with* $m \neq n$ *and* $x_1, \ldots, x_m, x'_1, \ldots, x'_n \in G$,

$$\rho(a(x_1) \cdots a(x_m)a(x'_n)^* \cdots a(x'_1)^*) = 0 .$$

Proof The assertions follow from Proposition 4.219 and direct computations. We omit the details. □

In the literature, quasi-free states are frequently defined to be gauge-invariant. Here, this case corresponds to the states that we call "simple quasi-free states," as shown in the last proposition. By contrast, [69, Definition 3.1, Condition (3.1)] only imposes the quasi-free states to be even, which is a strictly weaker property than being gauge-invariant. This more general setting refers here to the states of CAR C^*-algebras, which we call "extended quasi-free states." Quasi-free states of both kinds are very important in quantum statistical mechanics, because, among other things, the equilibrium states of free fermions turn out to be quasi-free states.

Chapter 5
Thermodynamic Equilibrium in Infinite Volume

5.1 Algebraic Framework

We begin by introducing simplified notations: For a fixed $d \in \mathbb{N}$ (space dimension), we define the cubic lattice $\Gamma \doteq \mathbb{Z}^d$ and use the symbol

$$\mathcal{P}_f \equiv \mathcal{P}_f(\Gamma) \doteq \{\Lambda \subsetneq \Gamma : |\Lambda| < \infty\} \subsetneq 2^\Gamma$$

for the set of all finite subsets of Γ. Here, $|\Lambda| \in \mathbb{N}_0$ is the "volume" of the set $\Lambda \subsetneq \Gamma$, that is, the number of points contained in Λ. For all $\Lambda \in \mathcal{P}_f$, we define the "diameter of Λ" by

$$d(\Lambda) \doteq \max\{|x - y| : x, y \in \Lambda\} < \infty . \tag{5.1}$$

To define the thermodynamic limit of quantum lattice systems, we use the sequence of cubic boxes defined, for each natural number $\ell \in \mathbb{N}$, by

$$\Lambda_\ell \doteq \{(x_1, \ldots, x_d) \in \Gamma : |x_i| \leq \ell\} \in \mathcal{P}_f . \tag{5.2}$$

In the sequel, we consider quantum spins and fermions on cubic lattices, whose observable algebras correspond to specific universal C^*-algebras (see Sect. 4.8) that we describe in the next two paragraphs.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_5

5.1.1 Quantum Spin Systems

Fix once and for all $N \in \mathbb{N}$. For any subset $\Lambda \subseteq \Gamma$, let

$$I_\Lambda \doteq \Lambda \times \mathcal{L}(\mathbb{C}^N) .$$

Recall that we canonically identify $\mathcal{L}(\mathbb{C}^N)$ with the (C^*-)algebra of $N \times N$ complex matrices. The spin C^*-algebra of the cubic lattice Γ for a fixed (once and for all) $N \in \mathbb{N}$ is the universal C^*-algebra of Definition 4.147 for the index set I_Γ, that is, the universal tensor product

$$\bigotimes_{x\in\Gamma} \mathcal{L}(\mathbb{C}^N) .$$

We use here the notation $\mathrm{Spin}(N, \Gamma)$ for this universal C^*-algebra. We similarly define the C^*-algebra $\mathrm{Spin}(N, \Lambda)$ for any nonempty subset $\Lambda \subseteq \Gamma$. Observe from Lemma 4.153 and Exercise 4.148 that all the spin algebras $\mathrm{Spin}(N, \Lambda)$, $\Lambda \subseteq \Gamma$, are simple and separable C^*-algebras. For every $\Lambda \subseteq \Gamma$, let

$$\mathrm{i}_{\Gamma\Lambda} : \mathrm{Spin}(N, \Lambda) \to \mathrm{Spin}(N, \Gamma)$$

denote the corresponding natural $*$-homomorphism (Definition 4.150). These $*$-homomorphisms are faithful, thanks to Corollary 4.154. Thus, we canonically identify $\mathrm{Spin}(N, \Lambda)$, $\Lambda \subseteq \Gamma$, $\Lambda \neq \emptyset$, with the C^*-subalgebra $\mathrm{i}_{\Gamma\Lambda}(\mathrm{Spin}(N, \Lambda)) \subseteq \mathrm{Spin}(N, \Gamma)$ of the (full) spin C^*-algebra of the cubic lattice Γ. In fact, the family $\{\mathrm{Spin}(N, \Lambda)\}_{\Lambda\subseteq\Gamma,\Lambda\neq\emptyset}$ defines a directed system of universal C^*-representations. See Definition 4.136 and remarks thereafter. With these identifications, observe that

$$\bigcup_{\Lambda\in\mathcal{P}_f} \mathrm{Spin}(N, \Lambda) \tag{5.3}$$

is a dense $*$-subalgebra, thanks to Exercise 4.131 and Corollary 4.132.

As we have seen in Sect. 4.8, the universal property of C^*-representations provides a very elegant way to define important families of $*$-automorphisms of universal C^*-algebras. In fact, Proposition 4.152 directly yields the existence of so-called translation automorphisms for the (full) spin algebra $\mathrm{Spin}(N, \Gamma)$: Note that, for all $x \in \mathbb{Z}^d$, the mapping $\pi_x : y \mapsto y+x$ is a bijection $\Gamma \to \Gamma$ (i.e., a permutation of Γ). In other words, by Proposition 4.152, for any fixed $x \in \mathbb{Z}^d$, there is a unique $*$-automorphism

$$\tau_x : \mathrm{Spin}(N, \Gamma) \to \mathrm{Spin}(N, \Gamma)$$

such that

$$\tau_x(a(A, y)) = a(y + x, A) , \qquad A \in \mathcal{L}(\mathbb{C}^N),\ y \in \Gamma . \tag{5.4}$$

Observe from Exercise 4.131 and Corollary 4.132 that, for all $\Lambda \subseteq \Gamma$ and all $x \in \mathbb{Z}^d$,

$$\tau_x(\mathrm{Spin}(N, \Lambda)) = \mathrm{Spin}(N, \Lambda + x) \ ,$$

where

$$\Lambda + x \doteq \{y + x \ : \ y \in \Lambda\} \subseteq \Gamma \ .$$

The above-defined translation automorphism $\{\tau_x\}_{x\in\mathcal{Z}^d}$ allows one to define translation invariant states of the spin algebra:

Definition 5.1 (Invariant States of Quantum Spins) A state $\rho \in E(\mathrm{Spin}(N, \Gamma))$ is "(translation) invariant," whenever $\rho = \rho \circ \tau_x$ for all $x \in \mathbb{Z}^d$. The convex set of all invariant states on $\mathrm{Spin}(N, \Gamma)$ is denoted by

$$E_1(\mathrm{Spin}(N, \Gamma)) \subseteq E(\mathrm{Spin}(N, \Gamma)) \ .$$

Extreme points of this convex set are called "ergodic states."[1] The set of all ergodic states on $\mathrm{Spin}(N, \Gamma)$ is denoted by $\mathcal{E}_1(\mathrm{Spin}(N, \Gamma))$.

Note that the set of invariant states is nonempty and weak*-compact. For instance, using the product states of Proposition 4.155 for copies of the same state of $\mathcal{L}(\mathbb{C}^N)$ for all $x \in \mathbb{Z}^d$, one constructs an invariant state: Explicitly, for any $\rho \in E_1(\mathrm{Spin}(N, \{0\})) \equiv E_1(\mathcal{L}(\mathbb{C}^N))$, one has that

$$\bigotimes_{x\in\mathbb{Z}^d} \rho \circ \tau_{-x}|_{\mathrm{Spin}(N,\{x\})} \in E_1(\mathrm{Spin}(N, \Gamma)) \ .$$

The set of ergodic states

$$\mathcal{E}_1(\mathrm{Spin}(N, \Gamma)) \subseteq E_1(\mathrm{Spin}(N, \Gamma))$$

is also never empty. This follows from Proposition 7.334.

Similar to the translation automorphism defined above, we can easily define so-called gauge automorphisms of the spin algebra, thanks to the universal property of this algebra, this time via Proposition 4.151: Let ϕ be any function $\Gamma \to \mathrm{SU}(N)$, where, as is usual, $\mathrm{SU}(N)$[2] denotes the group of *special* unitary $N \times N$ (complex) matrices, that is, unitary matrices with determinant equal to one. Then,

[1] We adopt this terminology, because of the formal analogy to the classical case: Recall that the invariant probability measures of a given classical dynamical systems are ergodic, in the usual sense, iff they are extreme in the convex set of all invariant measures of the system.

[2] In the present construction, one may even take the group $\mathrm{U}(N)$ of all unitary $N \times N$ matrices. However, we restrict ourselves to the so-called special unitary matrices, because this case is the most relevant one in physics.

by Proposition 4.151, there is a unique $*$-automorphism α_ϕ of the spin algebra $\mathrm{Spin}(N, \Gamma)$ satisfying

$$\alpha_\phi(a(x, A)) = a(x, \phi(x)A\phi^*(x)) \ , \qquad x \in \Gamma \ .$$

Observe that Proposition 4.151 only says that α_ϕ is a $*$-homomorphism. The fact that it is a $*$-automorphism, i.e., it is bijective, directly follows, once again, from Exercise 4.131 and Corollary 4.132. Because of the formal similarity with the so-called local $\mathrm{SU}(N)$ gauge transformations of gauge fields in quantum field theory, such $*$-automorphisms of $\mathrm{Spin}(N, \Gamma)$ are named here "general, or local, gauge automorphisms of $\mathrm{Spin}(N, \Gamma)$." Gauge automorphisms of $\mathrm{Spin}(N, \Gamma)$ associated with *constant* functions $\phi : \Gamma \to \mathrm{SU}(N)$, which, in turn, are canonically identified with the elements of $\mathrm{SU}(N)$, are called "global gauge automorphisms" of $\mathrm{Spin}(N, \Gamma)$. Clearly, the elements of the C^*-algebra $\mathrm{Spin}(N, \Gamma)$ that are invariant under global gauge automorphisms form a unital C^*-subalgebra of $\mathrm{Spin}(N, \Gamma)$, which is denoted by

$$\mathrm{Spin}_\circ(N, \Gamma) \doteq \{A \in \mathrm{Spin}(N, \Gamma) \ : \ \alpha_\phi(A) = A \text{ for all } \phi \in \mathrm{SU}(N)\} \ .$$

We call this C^*-algebra the "gauge-invariant spin algebra" of the cubic lattice Γ.

Notice that the *global* gauge automorphism of $\mathrm{Spin}(N, \Gamma)$ commute with the translation ones, that is, $\alpha_\phi \circ \tau_x = \tau_x \circ \alpha_\phi$ for all $x \in \mathbb{Z}^d$ and all $\phi \in \mathrm{SU}(N)$. In particular,

$$\tau_x(\mathrm{Spin}_\circ(N, \Gamma)) = \mathrm{Spin}_\circ(N, \Gamma) \ .$$

Thus, $\tau_x|_{\mathrm{Spin}_\circ(N,\Gamma)}$ are $*$-automorphisms of the gauge-invariant spin algebra $\mathrm{Spin}_\circ(N, \Gamma)$. Moreover, the *general* gauge automorphism preserves the local spin algebras, i.e., for any function $\phi : \Gamma \to \mathrm{SU}(N)$ and all $\Lambda \in \mathcal{P}_f$,

$$\alpha_\phi(\mathrm{Spin}(N, \Lambda)) = \mathrm{Spin}(N, \Lambda) \ .$$

In particular, $\alpha_\phi|_{\mathrm{Spin}(N,\Lambda)}$ is a $*$-automorphism of the $*$-subalgebra $\mathrm{Spin}(N, \Lambda) \subseteq \mathrm{Spin}(N, \Gamma)$ for every $\Lambda \in \mathcal{P}_f$. Again, both properties directly follow from Exercise 4.131 combined with Corollary 4.132.

5.1.2 Lattice Fermion Systems

Fix once and for all a finite subset Ω. For any nonempty subset $\Lambda \subseteq \Gamma$, define the Hilbert space

$$\ell^2(\Omega \times \Lambda) \doteq \left\{ f \in \mathcal{F}(\Omega \times \Lambda; \mathbb{C}) : \|f\|_2^2 \doteq \sum_{x \in \Omega \times \Lambda} |f(x)|^2 < \infty \right\}$$

of square summable functions from $\Omega \times \Gamma$ to $\mathbb{C}$. See Definition 7.228 and Exercise 7.229. We canonically identify the vector space $\ell^2(\Omega \times \Lambda)$ with a vector subspace of $\ell^2(\Omega \times \Gamma)$, by setting the functions of $\ell^2(\Omega \times \Lambda)$ to zero outside the subset $\Omega \times \Lambda \subseteq \Omega \times \Gamma$.

The CAR C^*-algebra of the cubic lattice Γ for a fixed (once and for all) finite subset Ω is the universal C^*-algebra of Definition 4.163 for the index set $G = \ell^2(\Omega \times \Gamma)$, that is, the CAR C^*-algebra $\mathrm{CAR}(\ell^2(\Omega \times \Gamma))$. We use here the notation $\mathrm{CAR}(\Omega, \Gamma)$ for this universal C^*-algebra. We similarly define the C^*-algebra $\mathrm{CAR}(\Omega, \Lambda)$ for any nonempty subset $\Lambda \subseteq \Gamma$. Observe from Corollary 4.167 and Exercise 4.164 that all the CAR C^*-algebras $\mathrm{CAR}(\Omega, \Lambda)$, $\Lambda \subseteq \Gamma$, are simple and separable C^*-algebras. Recall that the Hilbert spaces $\ell^2(\Omega \times \Lambda)$, $\Lambda \subseteq \Gamma$, are separable, the domains $\Omega \times \Lambda$ being countable sets. Since for every nonempty $\Lambda \subseteq \Gamma$, $\ell^2(\Omega \times \Lambda)$ is canonically identified with a vector subspace of $\ell^2(\Omega \times \Gamma)$, by Definition 4.150 and Corollary 4.169, we canonically identify $\mathrm{CAR}(\Omega, \Lambda)$ with a C^*-subalgebra of $\mathrm{CAR}(\Omega, \Gamma)$ for every nonempty subsets $\Lambda \subseteq \Gamma$. In fact, as in the quantum spin case, the family

$$\{\mathrm{CAR}(\Omega, \Lambda)\}_{\Lambda \subseteq \Gamma, \Lambda \neq \emptyset}$$

defines a directed system of universal C^*-representations. See Definition 4.136 and remarks thereafter. With these identifications, again similar to the quantum spin case,

$$\bigcup\nolimits_{\Lambda \in \mathcal{P}_f} \mathrm{CAR}(\Omega, \Lambda) \tag{5.5}$$

is a dense $*$-subalgebra, thanks to Exercise 4.131 and Corollary 4.132.

In contrast to the quantum spin case, for fermions, it is important to consider the corresponding even CAR C^*-algebras, which are denoted here by

$$\mathrm{CAR}_{\mathrm{e}}(\Omega, \Lambda) \doteq \mathrm{CAR}_{\mathrm{e}}(\ell^2(\Omega \times \Lambda)) \subseteq \mathrm{CAR}(\Omega, \Lambda) , \qquad \Lambda \subseteq \Gamma.$$

See Definition 4.185 (ii) for $G = \ell^2(\Omega \times \Gamma)$. Observe from Exercise 4.186 that, for all nonempty $\Lambda, \Lambda' \subseteq \Gamma$ with $\Lambda' \subseteq \Lambda$, one has

$$\mathrm{CAR}_{\mathrm{e}}(\Omega, \Lambda') = \mathrm{CAR}(\Omega, \Lambda') \cap \mathrm{CAR}_{\mathrm{e}}(\Omega, \Lambda) .$$

Like for spin algebras, we define general (or local) gauge automorphisms. Here, they refer to special cases of Bogoliubov $*$-homomorphisms (see Definition 4.181): For every function $\phi : \Gamma \to \mathbb{R}$, we define the unitary transformation $U_\phi \in \mathcal{B}(\ell^2(\Omega \times \Gamma))$ by

$$U_\phi(f)(s, x) \doteq \mathrm{e}^{i\phi(x)} f(s, x) , \qquad s \in \Omega, \ x \in \Gamma .$$

Clearly, for any pair of functions $\phi, \phi' : \Gamma \to \mathbb{R}$,

$$U_\phi \circ U_{\phi'} = U_{\phi+\phi'} \ . \tag{5.6}$$

In other words,

$$\mathfrak{U}_\circ^{\mathrm{loc}} \doteq \{U_\phi \ : \ \phi \text{ is a function } \Gamma \to \mathbb{R}\} \subseteq \mathcal{B}(\ell^2(\Omega \times \Gamma))$$

is a group of invertible transformations preserving scalar products. Then, the general gauge automorphisms of $\mathrm{CAR}(\Omega, \Gamma)$ are the Bogoliubov $*$-automorphisms associated with this group. Observe that these automorphisms refer to the so-called local $U(1)$-gauge transformations of quantum field theory. To fix the notation, for all functions $\phi : \Gamma \to \mathbb{R}$,

$$\alpha_\phi \doteq \mathrm{Bog}(U_\phi) \in \mathrm{Aut}(\mathrm{CAR}(\Omega, \Gamma)) \ ,$$

where $\mathrm{Aut}(\mathrm{CAR}(\Omega, \Gamma))$ denotes the group of automorphism of $\mathrm{CAR}(\Omega, \Gamma)$. By Corollary 4.183 and Eq. (5.6),

$$\alpha_\phi \circ \alpha_{\phi'} = \alpha_{\phi+\phi'} \ , \qquad \phi, \phi' : \Gamma \to \mathbb{R} \ .$$

The gauge automorphisms corresponding to *constant* functions $\phi : \Gamma \to \mathbb{R}$ are called "global gauge automorphisms" of $\mathrm{CAR}(\Omega, \Gamma)$. We again canonically identify such functions with real numbers. In other words, the global gauge automorphisms of $\mathrm{CAR}(\Omega, \Gamma)$ are those Bogoliubov $*$-automorphisms related with the subgroup

$$\mathfrak{U}_\circ \doteq \{U_\phi \ : \ \phi \in \mathbb{R}\} \subseteq \mathfrak{U}_\circ^{\mathrm{loc}} \ .$$

In fact, for all $\phi \in [0, 2\pi)$, α_ϕ is nothing else than the gauge automorphism of Definition 4.185 (i) for $G = \ell^2(\Omega \times \Gamma)$.

As in the case of spin algebras, by Exercise 4.131 and Corollary 4.132, (general) gauge automorphisms of $\mathrm{CAR}(\Omega, \Gamma)$ preserve the local CAR subalgebras, that is,

$$\alpha_\phi(\mathrm{CAR}(\Omega, \Lambda)) = \mathrm{CAR}(\Omega, \Lambda) \ , \qquad \Lambda \in \mathcal{P}_f \ .$$

Thus, for all $\Lambda \in \mathcal{P}_f$, $\alpha_\phi|_{\mathrm{CAR}(\Omega,\Lambda)}$ is a -automorphism of $\mathrm{CAR}(\Omega, \Lambda)$. As in the case of spin algebras, the elements of the C^*-algebra $\mathrm{CAR}(\Omega, \Gamma)$ that are invariant under the global gauge automorphisms form a unital C^*-subalgebra of $\mathrm{CAR}(\Omega, \Gamma)$, which is denoted by

$$\mathrm{CAR}_\circ(\Omega, \Gamma) \doteq \{A \in \mathrm{CAR}(\Omega, \Gamma) \ : \ \alpha_\phi(A) = A \text{ for all } \phi \in \mathbb{R}\} \ .$$

Clearly, $\mathrm{CAR}_\circ(\Omega, \Gamma) \subseteq \mathrm{CAR}_\mathrm{e}(\Omega, \Gamma)$.

Analogous to the translation automorphisms of quantum spin systems (5.4), we can define translation automorphisms also for the CAR algebra of the cubic lattice. Like the above-defined gauge automorphisms, in the fermion case, such $*$-automorphisms of $\mathrm{CAR}(\Omega, \Gamma)$ are Bogoliubov $*$-homomorphisms (see Definition 4.181): For all $x \in \mathbb{Z}^d$, define the unitary operator $U_x \in \mathcal{B}(\ell^2(\Omega \times \Gamma))$ by

$$U_x(f)(s, y) \doteq f(s, y - x) , \qquad s \in \Omega, \ y \in \Gamma . \tag{5.7}$$

It is easy to check that

$$U_x \circ U_{x'} = U_{x+x'} , \qquad x, x' \in \mathbb{Z}^d . \tag{5.8}$$

In other words,

$$\mathfrak{U}_t \doteq \{U_x \ : \ x \in \mathbb{Z}^d\} \subseteq \mathcal{B}(\ell^2(\Omega \times \Gamma))$$

is a group of invertible transformations preserving scalar products. Then, the translation automorphisms of $\mathrm{CAR}(\Omega, \Gamma)$ are the Bogoliubov $*$-automorphisms associated with this group. To fix notation, for all $x \in \mathbb{Z}^d$,

$$\tau_x \doteq \mathrm{Bog}(U_x) \in \mathrm{Aut}(\mathrm{CAR}(\Omega, \Gamma)) . \tag{5.9}$$

Again by Corollary 4.183 and Eq. (5.8),

$$\tau_x \circ \tau_{x'} = \tau_{x+x'} , \qquad x, x' \in \mathbb{Z}^d . \tag{5.10}$$

Additionally, like for spin algebras, the global gauge automorphisms of $\mathrm{CAR}(\Omega, \Gamma)$ commute with translation automorphisms, that is, for all $x \in \mathbb{Z}^d$ and $\phi \in \mathbb{R}$, one has $\tau_x \circ \alpha_\phi = \alpha_\phi \circ \tau_x$. This is again a consequence of Exercise 4.131 combined with Corollary 4.132. In particular, like in the quantum spin case, for all $x \in \mathbb{Z}^d$,

$$\tau_x(\mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma)) = \mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma) \quad \text{and} \quad \tau_x(\mathrm{CAR}_{\circ}(\Omega, \Gamma)) = \mathrm{CAR}_{\circ}(\Omega, \Gamma) .$$

In other words, for all $x \in \mathbb{Z}^d$, the restrictions $\tau_x|_{\mathrm{CAR}_{\mathrm{e}}(\Omega,\Gamma)}$ and $\tau_x|_{\mathrm{CAR}_{\circ}(\Omega,\Gamma)}$ are, respectively, $*$-automorphisms of the even and gauge-invariant CAR algebras, $\mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma)$ and $\mathrm{CAR}_{\circ}(\Omega, \Gamma)$.

Similar to (translation) invariant states of Definition 5.1 for quantum spin systems, we introduce such a notion for lattice fermions:

Definition 5.2 (Invariant States of Lattice Fermions) A state $\rho \in E(\mathrm{CAR}(\Omega, \Gamma))$ is "(translation) invariant," whenever $\rho = \rho \circ \tau_x$ for every $x \in \mathbb{Z}^d$. The (nonempty) convex set of all invariant states on $\mathrm{CAR}(\Omega, \Gamma)$ is denoted by

$$E_1(\mathrm{CAR}(\Omega, \Gamma)) \subseteq E(\mathrm{CAR}(\Omega, \Gamma)) .$$

Extreme points in this convex set are called "ergodic states." The set of all ergodic states on $\mathrm{CAR}(\Omega, \Gamma)$ is denoted by $\mathcal{E}_1(\mathrm{CAR}(\Omega, \Gamma))$.

Like in the quantum spin case, the sets

$$E_1(\mathrm{CAR}(\Omega, \Gamma)) \subseteq E(\mathrm{CAR}(\Omega, \Gamma)) \quad \text{and} \quad \mathcal{E}_1(\mathrm{CAR}(\Omega, \Gamma)) \subseteq E_1(\mathrm{CAR}(\Omega, \Gamma))$$

are nonempty and weak*-compact. To get an invariant state, similar to the quantum spin case, use, for instance, the product $\bigotimes_{x\in\Gamma} \rho_x$ of even states given by Proposition 4.193, where $\rho_x \in E_{\mathrm{e}}(\mathrm{CAR}(\Omega, \{x\}))$ is defined by $\rho_x \doteq \rho \circ \tau_{-x}$ for all $x \in \Gamma$ and some fixed even state $\rho \in E_{\mathrm{e}}(\mathrm{CAR}(\Omega, \{0\}))$. In particular, ergodic states exist, thanks to Proposition 7.334.

Importantly, observe that, besides the fact that translation automorphisms preserve the even CAR C^*-algebra $\mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma) \subseteq \mathrm{CAR}(\Omega, \Gamma)$, as stated above, invariant states of $\mathrm{CAR}(\Omega, \Gamma)$ turn out to be necessarily even, in the sense of Definition 4.190:

Theorem 5.3 *Let Ω be any finite set and fix $d \in \mathbb{N}$. Then* $E_1(\mathrm{CAR}(\Omega, \Gamma)) \subseteq E_{\mathrm{e}}(\mathrm{CAR}(\Omega, \Gamma))$.

Proof See [1, Lemma 1.8] for more details. □

By combining Theorem 5.3 with Lemma 4.192, the set $E_1(\mathrm{CAR}(\Omega, \Gamma))$ of all invariant states can be identified with a subset of $E(\mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma))$.

Remark 5.4 Note that we have taken the CAR formulation of fermion systems which allows us to define global gauge automorphisms as Bogoliubov automorphisms. Alternatively, one can also use the self-dual approach described in Sect. 4.8.3.

5.1.3 General Notation Encoding Both Fermion and Quantum Spin Systems

We want to consider, as far as possible, both the spin and fermion cases at the same time. We therefore use a unified notation for these two cases:

- The spin and fermion C^*-algebras $\mathrm{Spin}(N, \Gamma)$ and $\mathrm{CAR}(\Omega, \Gamma)$ are denoted by $\mathcal{U}$. We explicitly use the previous notation, $\mathrm{Spin}(N, \Gamma)$ and $\mathrm{CAR}(\Omega, \Gamma)$, only when distinguishing between both cases is mandatory.
- In the fermion case, the even (CAR) C^*-subalgebra $\mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma)$ of $\mathrm{CAR}(\Omega, \Gamma)$ is fundamental. It is thus denoted by $\mathcal{U}^{\mathrm{e}}$. If one considers the quantum spin case, $\mathcal{U}^{\mathrm{e}}$ is just the original algebra, i.e., $\mathcal{U}^{\mathrm{e}} = \mathrm{Spin}(N, \Gamma)$.
- $\mathcal{U}^{\circ}$ stands for the (globally) gauge-invariant subalgebra in both fermion and quantum spin cases, that is, $\mathcal{U}^{\circ} \doteq \mathrm{CAR}_{\circ}(\Omega, \Gamma)$ in the fermion case, while $\mathcal{U}^{\circ} \doteq \mathrm{Spin}_{\circ}(N, \Gamma)$ for quantum spins. Clearly, $\mathcal{U}^{\circ} \subseteq \mathcal{U}^{e}$ in both cases.

- The dense sets (5.3) and (5.5) of local elements are denoted in both cases by

$$\mathcal{U}_{\mathrm{loc}} \doteq \bigcup\nolimits_{\Lambda \in \mathcal{P}_f} \mathcal{U}_\Lambda \ ,$$

  where $\mathcal{U}_\Lambda \doteq \mathrm{Spin}(N, \Lambda)$ (quantum spin case) or $\mathcal{U}_\Lambda \doteq \mathrm{CAR}(\Omega, \Lambda)$ (fermion case) for any finite subset $\Lambda \in \mathcal{P}_f$. By definition, $\mathcal{U}_\emptyset \doteq \mathbb{C}\mathfrak{1}$. In addition, $\mathcal{U}^{\mathrm{e}}_{\mathrm{loc}} \doteq \mathcal{U}_{\mathrm{loc}} \cap \mathcal{U}^{\mathrm{e}}$ and $\mathcal{U}^{\mathrm{e}}_\Lambda \doteq \mathcal{U}_\Lambda \cap \mathcal{U}^{\mathrm{e}}$ for $\Lambda \in \mathcal{P}_f$.
- The set of all invariant states is denoted by $E_1 \equiv E_1(\mathcal{U})$, with $\mathcal{E}_1 \equiv \mathcal{E}_1(\mathcal{U})$ being the corresponding set of ergodic states. In other words, in the quantum spin case, $E_1 = E_1(\mathrm{Spin}(N, \Gamma))$ and $\mathcal{E}_1 = \mathcal{E}_1(\mathrm{Spin}(N, \Gamma))$, while for fermions, $E_1 = E_1(\mathrm{CAR}(\Omega, \Gamma))$ and $\mathcal{E}_1 = \mathcal{E}_1(\mathrm{CAR}(\Omega, \Gamma))$.
- The notation α_ϕ, with $\phi \in \mathbb{R}$ in the fermion case and $\phi \in \mathrm{SU}(N)$ in the quantum spin case, denotes an arbitrary global gauge automorphism of $\mathcal{U}$.

In the sequel, we consider only *translation invariant* quantum spin or fermion systems. Recall that translations in $\mathcal{U}$ refer to (unique) $*$-automorphisms $\tau_x : \mathcal{U} \to \mathcal{U}$, $x \in \mathbb{Z}^d$, defined by (5.4), in the quantum spin case, and by (5.9), in the fermion case. The "translation automorphisms" represent in the given observable algebra $\mathcal{U}$ the group of (physical) translations in the lattice Γ, as one sees from the construction of these $*$-automorphisms explained above.

5.2 Interactions of Infinite Spin and Fermion Systems on the Lattice

Every physical system of particles belongs to some finite region $\Lambda \subseteq \Gamma$ of space, here represented by the lattice Γ, and they become macroscopic in the large volume limit $|\Lambda| \to \infty$. Thus, infinitely extended systems refer to a family of local Hamiltonians (associated with any finite region Λ of space) that are related to a global *interaction* defined as follows:

Definition 5.5 (Spin and Fermion Interactions on the Lattice Γ)

(i) We say that the mapping $\Phi : \mathcal{P}_f \to \mathcal{U}$ is a "(quantum spin or fermion) interaction" on the lattice Γ if, for all $\Lambda \in \mathcal{P}_f$,

$$\Phi(\Lambda) = \Phi(\Lambda)^* \in \mathcal{U}^e_\Lambda \ ,$$

where we recall that $\mathcal{U}^{\mathrm{e}}_\Lambda \doteq \mathcal{U}_\Lambda \cap \mathcal{U}^{\mathrm{e}}$, where $\mathcal{U}^{\mathrm{e}} \doteq \mathcal{U} \doteq \mathrm{Spin}(N, \Gamma)$ and $\mathcal{U}_\Lambda \doteq \mathrm{Spin}(N, \Lambda)$, in the quantum spin case, while $\mathcal{U}^{\mathrm{e}} = \mathrm{CAR}_{\mathrm{e}}(\Omega, \Gamma)$ and $\mathcal{U}_\Lambda \doteq \mathrm{CAR}(\Omega, \Lambda)$ in the fermion case.

(ii) We say that the (spin or fermion) interaction Φ has "finite range" if there is $R < \infty$ such that, for all $\Lambda \in \mathcal{P}_f$, $\Phi(\Lambda) = 0$ whenever $d(\Lambda) > R$ (see (5.1)). In this case, the smallest $R \geq 0$ with this property is called the "range" of Φ.

(iii) The (spin or fermion) interaction Φ is said to be "invariant" if, for all $\Lambda \in \mathcal{P}_f$ and $x \in \mathbb{Z}^d$,

$$\tau_x(\Phi(\Lambda)) = \Phi(\Lambda + x) ,$$

where $\tau_x : \mathcal{U} \to \mathcal{U}$, $x \in \Gamma$, are the translation automorphisms on $\mathcal{U}$ defined by (5.4), in the quantum spin case, and (5.9), in the fermion case, while

$$\Lambda + x \doteq \{y + x \ : \ y \in \Lambda\} .$$

(iv) The (spin or fermion) interaction Φ is "(globally) gauge-invariant" if, for all $\Lambda \in \mathcal{P}_f$, $\Phi(\Lambda) \in \mathcal{U}^{\circ}$.

Note that the set of all (spin or fermion) interactions naturally has the structure of a real vector space $\mathcal{V}$: For all interactions $\Phi, \Phi' \in \mathcal{V}$ and constant $\alpha \in \mathbb{R}$, the real-valued functions $\Phi + \Phi'$ and $\alpha\Phi$, defined on $\mathcal{P}_f$ by

$$(\Phi + \Phi')(\Lambda) \doteq \Phi(\Lambda) + \Phi'(\Lambda) \qquad \text{and} \qquad (\alpha\Phi)(\Lambda) \doteq \alpha(\Phi(\Lambda)) , \tag{5.11}$$

are clearly interactions. The invariant interactions form a subspace of this vector space $\mathcal{V}$. In the following, we introduce a norm for invariant interactions:

Definition 5.6 (A Banach Space of Invariant Interactions) For any invariant (spin or fermion) interaction Φ, define the following quantity:

$$\|\Phi\| \doteq \sum_{\Lambda \in \mathcal{P}_f,\, 0 \in \Lambda} \frac{1}{|\Lambda|} \|\Phi(\Lambda)\| \in \mathbb{R}_0^+ \cup \{\infty\} .$$

It is a norm in the vector space

$$\mathcal{W}_1 \doteq \{\Phi \in \mathcal{V} : \Phi \text{ is an invariant interaction with } \|\Phi\| < \infty\} .$$

Recall that $|\Lambda| \in \mathbb{N}$ stands for the number of points contained in $\Lambda \in \mathcal{P}_f$. $\mathcal{W}_1^{\circ} \subseteq \mathcal{W}_1$ denotes the subspace of its (globally) gauge-invariant interactions.

Note that

$$\mathcal{W}_1^f \doteq \{\Phi \in \mathcal{W}_1 \ : \ \Phi \text{ has finite range}\}$$

is a subspace of $\mathcal{W}_1$. In fact, $\mathcal{W}_1$ is a Banach space (i.e., it is complete with respect to the above-defined norm), and $\mathcal{W}_1^f \subseteq \mathcal{W}_1$ is a dense separable subspace. In particular, $\mathcal{W}_1$ is separable. It is easy to check that $\mathcal{W}_1^{\circ}$ is a separable Banach subspace of $\mathcal{W}_1$.

Example 5.7 (Heisenberg Interaction) Let $N \doteq 2, \mathcal{U} \doteq \mathrm{Spin}(2, \Gamma)$ and define the (Pauli) matrices $\sigma_1, \sigma_2, \sigma_3 \in \mathcal{L}(\mathbb{C}^2)^{\mathbb{R}}$, respectively, by

$$\sigma_1 \doteq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 \doteq \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 \doteq \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The "Heisenberg interaction" $\Phi_{\mathrm{Heis}} : \mathcal{P}_f \to \mathcal{U}$ is defined by $\Phi_{\mathrm{Heis}}(\Lambda) \doteq 0$ whenever $|\Lambda| \notin \{1, 2\}$ or $d(\Lambda) > 1$,

$$\Phi_{\mathrm{Heis}}(\{x\}) \doteq -\sum_{k=1}^{3} h_k a(x, \sigma_k)$$

for all $x \in \Gamma$, and

$$\Phi_{\mathrm{Heis}}(\{x, x'\}) \doteq -J \sum_{k=1}^{3} a(x, \sigma_k) a(x', \sigma_k)$$

for all $\{x, x'\} \subseteq \Gamma$ with $d(\{x, x'\}) = 1$. Here, $h = (h_1, h_2, h_3) \in \mathbb{R}^3$ and $J \in \mathbb{R}$ are parameters of the Heisenberg model: h represents a homogeneous external magnetic field and J the strength of the magnetic coupling of neighboring spins. The interaction Φ_{Heis} is said to be "ferromagnetic" if $J \geq 0$. It is called "antiferromagnetic," otherwise. Recall that the symbol a appearing above refers to a mapping from $\Gamma \times \mathcal{L}(\mathbb{C}^2)$ to the C^*-algebra $\mathcal{U}$, which satisfies the family of polynomial relations defining the universal C^*-algebra $\mathcal{U} \doteq \mathrm{Spin}(2, \Gamma)$.

Observe that the Heisenberg interaction is (globally) gauge-invariant if the external magnetic field h is zero:

Lemma 5.8 (Gauge Invariance of the Heisenberg Model) *If $h = 0$, then, for all $\Lambda \in \mathcal{P}_f$ and $\phi \in \mathrm{SU}(2)$, $\alpha_\phi(\Phi_{\mathrm{Heis}}(\Lambda)) = \Phi_{\mathrm{Heis}}(\Lambda)$.*

Proof If $h = 0$, then, for any $\Lambda \in \mathcal{P}_f$, up to a constant ($0$ or $-J$), $\Phi_{\mathrm{Heis}}(\Lambda)$ is equal to

$$\sum_{k=1}^{3} a(x, \sigma_k) a(x', \sigma_k)$$

for some $x, x' \in \Lambda \subseteq \Gamma$. Thus, it suffices to show that this type of element of $\mathrm{Spin}(2, \Gamma)$ belongs to (the gauge-invariant subalgebra) $\mathrm{Spin}_\circ(2, \Gamma)$. Recalling that the Pauli matrices are the infinitesimal generators of (the compact Lie group) $\mathrm{SU}(2)$, observe that, for any matrix $\phi \in \mathrm{SU}(2)$, there are real constants $c_1, c_2, c_3 \in \mathbb{R}$ such that

$$\phi = \exp\left(i \sum_{k=1}^{3} c_k \sigma_k\right) .$$

For all $s \in \mathbb{R}$, define

$$\phi(s) \doteq \exp\left(is \sum_{k=1}^{3} c_k \sigma_k\right) .$$

We prove now that

$$\frac{\mathrm{d}}{\mathrm{d}s}\alpha_{\phi(s)}\left(\sum_{k=1}^{3} a(x,\sigma_k)a(x',\sigma_k)\right) = 0 , \qquad s \in \mathbb{R} .$$

By direct computations, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}s}\alpha_{\phi(s)}\left(\sum_{l=1}^{3} a(x,\sigma_l)a(x',\sigma_l)\right)$$
$$= i\alpha_{\phi(s)}\left(\sum_{l,k=1}^{3} c_k(a(x,[\sigma_k,\sigma_l])a(x',\sigma_l) + a(x,\sigma_l)a(x',[\sigma_k,\sigma_l]))\right) .$$

Now, we use the well-known commutation relations for the Pauli matrices, i.e.,

$$[\sigma_{k,}\sigma_l] = 2i \sum_{m=1}^{3} \varepsilon_{klm}\sigma_m ,$$

where, as is usual, ε_{mlk} denotes the "completely antisymmetric tensor,"that is, $\varepsilon_{\pi(1)\pi(2)\pi(3)} \in \{-1, 1\}$ is the sign of the permutation π of three elements and it is zero, else. Hence,

$$\frac{\mathrm{d}}{\mathrm{d}s}\alpha_{\phi(s)}\left(\sum_{l=1}^{3} a(x,\sigma_l)a(x',\sigma_l)\right)$$
$$= 2 \sum_{l,k,m=1}^{3} \varepsilon_{klm}c_k\alpha_{\phi(s)}(a(x,\sigma_m)a(x',\sigma_l) + a(x,\sigma_l)a(x',\sigma_m)) = 0 .$$

Observe that the argument of the last sum is antisymmetric with respect to the permutation of the indices l and m, and, thus, the sum trivially vanishes. □

Example 5.9 (Hubbard Interaction) Let $\Omega \doteq \{\uparrow, \downarrow\}$ and $\mathcal{U} \doteq \mathrm{CAR}(\Omega, \Gamma)$. Define, for all $x \in \Gamma$ and $\mathrm{s} \in \Omega$, the functions $\mathrm{e}_{\mathrm{s},x} \in \ell^2(\Omega \times \Gamma)$ by $\mathrm{e}_{\mathrm{s},x}(\tilde{\mathrm{s}}, \tilde{x}) = 1$ if $(\mathrm{s}, x) = (\tilde{\mathrm{s}}, \tilde{x})$ and $\mathrm{e}_{\mathrm{s},x}(\tilde{\mathrm{s}}, \tilde{x}) = 0$, else. (Note that $\{\mathrm{e}_{\mathrm{s},x}\}_{(\mathrm{s},x)\in\Omega\times\Gamma}$ is a Hilbert basis of $\ell^2(\Omega\times\Gamma)$.) The "Hubbard interaction" $\Phi_{\mathrm{Hubb}} : \mathcal{P}_f \to \mathcal{U}$ is defined by $\Phi_{\mathrm{Hubb}} \doteq 0$ whenever $|\Lambda| \notin \{1, 2\}$ or $d(\Lambda) > 1$,

$$\Phi_{\mathrm{Hubb}}(\{x\}) \doteq U a(\mathrm{e}_{\uparrow,x})^* a(\mathrm{e}_{\downarrow,x})^* a(\mathrm{e}_{\downarrow,x}) a(\mathrm{e}_{\uparrow,x}) - \mu \sum_{\mathrm{s}\in\Omega} a(\mathrm{e}_{\mathrm{s},x})^* a(\mathrm{e}_{\mathrm{s},x}) ,$$

for all $x \in \Gamma$, and

$$\Phi_{\mathrm{Hubb}}(\{x, x'\}) \doteq -t \sum_{\mathrm{s}\in\Omega} (a(\mathrm{e}_{\mathrm{s},x})^* a(\mathrm{e}_{\mathrm{s},x'}) + a(\mathrm{e}_{\mathrm{s},x'})^* a(\mathrm{e}_{\mathrm{s},x}))$$

for all $\{x, x'\} \subseteq \Gamma$ with $d(\{x, x'\}) = 1$. Here, $t, \mu, U \in \mathbb{R}$ are parameters of the Hubbard model: μ represents the "chemical potential" of the electrons, t their "hopping amplitude," and U the strength of the electronic on-site repulsion. Recall that the symbol a appearing above refers to a mapping from $\ell^2(\Omega \times \Gamma)$ to the C^*-algebra $\mathcal{U}$ that satisfies the family of polynomial relations defining the universal C^*-algebra $\mathcal{U} \doteq \mathrm{CAR}(\Omega, \Gamma)$. Observe additionally that the Hubbard interaction is (globally) gauge-invariant for any choice of parameters.

The Heisenberg interaction Φ_{Heis} and the Hubbard interaction Φ_{Hubb} are both examples of (translation) invariant range-one interactions. As discussed above, any interaction determines an associated family of local energy observables (local Hamiltonians):

Definition 5.10 (Local Energy Observables) Let $\Phi \in \mathcal{V}$ be any (spin or fermion) interaction.

(i) For all $\Lambda \in \mathcal{P}_f$, let

$$H_\Lambda^\Phi \doteq \sum_{\Lambda'\in\mathcal{P}_f,\ \Lambda'\subseteq\Lambda} \Phi(\Lambda') \in \mathrm{Re}\{\mathcal{U}^e\} .$$

Recall that $\mathrm{Re}\{\mathcal{U}^e\}$ denotes the space of self-conjugate elements of $\mathcal{U}^\mathrm{e}$. $H_\Lambda^\Phi = (H_\Lambda^\Phi)^*$ is the "local Hamiltonian associated with the (finite) region Λ and the interaction Φ."

(ii) If $\Phi \in \mathcal{W}_1$, define the "energy density observable" associated with Φ by

$$e_\Phi \doteq \sum_{\Lambda\in\mathcal{P}_f,\ 0\in\Lambda} \frac{1}{|\Lambda|} \Phi(\Lambda) \in \mathrm{Re}\{\mathcal{U}^e\} .$$

Observe here that the set $\mathcal{P}_f$ of all finite subsets of the lattice $\Gamma \doteq \mathbb{Z}^d$ is countable.

If $\Phi \in \mathcal{W}_1$, notice that e_Φ is well-defined in the corresponding C^*-algebra, as the defining sum is absolutely convergent, by the finiteness of the norm $\|\Phi\|$. In this case, one has the estimate $\|e_\Phi\| \leq \|\Phi\|$. Additionally,

$$e_{\Phi+\Phi'} = e_\Phi + e_{\Phi'} , \qquad e_{\alpha\Phi} = \alpha e_\Phi$$

for all $\Phi, \Phi' \in \mathcal{W}_1$ and $\alpha \in \mathbb{R}$. The (spin or fermion) algebra element $e_\Phi = e_\Phi^*$ is called "energy density observable" associated with the interaction $\Phi \in \mathcal{W}_1$, because of the following result:

Proposition 5.11 *Fix any invariant interaction $\Phi \in \mathcal{W}_1$. Then, for any* invariant *state $\rho \in E_1$,*

$$\lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|} \rho(H_{\Lambda_\ell}^\Phi) = \rho(e_\Phi) ,$$

where Λ_ℓ, $\ell \in \mathbb{N}$, are the sequence of cubic boxes defined by (5.2).

Proof

1. For all invariant interactions $\Phi \in \mathcal{W}_1$,

$$\begin{aligned} H_{\Lambda_\ell}^\Phi &= \sum_{\tilde\Lambda \in \mathcal{P}_f} \mathbf{1}_{\left\{\tilde\Lambda \subseteq \Lambda_\ell\right\}} \Phi(\tilde\Lambda) \\ &= \sum_{x\in\Lambda_\ell} \sum_{\tilde\Lambda\in\mathcal{P}_f, x\in\tilde\Lambda} \frac{1}{|\tilde\Lambda|} \mathbf{1}_{\left\{\tilde\Lambda\subseteq\Lambda_\ell\right\}} \Phi(\tilde\Lambda) \\ &= \sum_{x\in\Lambda_\ell} \sum_{\tilde\Lambda\in\mathcal{P}_f, 0\in\tilde\Lambda} \mathbf{1}_{\left\{\tilde\Lambda\subseteq\Lambda_\ell - x\right\}} \frac{\Phi(x+\tilde\Lambda)}{|\tilde\Lambda|} \\ &= \sum_{x\in\Lambda_\ell} \sum_{\tilde\Lambda\in\mathcal{P}_f, 0\in\tilde\Lambda} \mathbf{1}_{\left\{\tilde\Lambda\subseteq\Lambda_\ell - x\right\}} \frac{\tau_x(\Phi(\tilde\Lambda))}{|\tilde\Lambda|} . \end{aligned}$$

2. Thus, if $\rho \in E_1$ is an invariant state, then

$$\frac{\rho\left(H_{\Lambda_\ell}^\Phi\right)}{|\Lambda_\ell|} = \sum_{\tilde\Lambda\in\mathcal{P}_f, 0\in\tilde\Lambda} \rho\left(\frac{\Phi(\tilde\Lambda)}{|\tilde\Lambda|}\right) \left(\sum_{x\in\Lambda_\ell} \frac{\mathbf{1}_{\left\{\tilde\Lambda\subseteq\Lambda_\ell - x\right\}}}{|\Lambda_\ell|}\right) .$$

3. For all $\tilde{\Lambda} \in \mathcal{P}_f$ and $\ell \in \mathbb{N}$,

$$\left| \rho \left(\frac{\Phi(\tilde{\Lambda})}{|\tilde{\Lambda}|} \right) \right| \left(\sum_{x \in \Lambda_\ell} \frac{\mathbf{1}_{\left\{\tilde{\Lambda} \subseteq \Lambda_\ell - x\right\}}}{|\Lambda_\ell|} \right) \leq \frac{\|\Phi(\tilde{\Lambda})\|}{|\tilde{\Lambda}|} .$$

Additionally, for any $\tilde{\Lambda} \in \mathcal{P}_f$,

$$\lim_{\ell \to \infty} \left(\sum_{x \in \Lambda_\ell} \frac{\mathbf{1}_{\left\{\tilde{\Lambda} \subseteq \Lambda_\ell - x\right\}}}{|\Lambda_\ell|} \right) = 1 .$$

4. By definition of the norm

$$\|\Phi\| \doteq \sum_{\Lambda \in \mathcal{P}_f,\, 0 \in \Lambda} \frac{\|\Phi(\Lambda)\|}{|\Lambda|} < \infty , \qquad \Phi \in \mathcal{W}_1 ,$$

it follows from Corollary 7.314 and the continuity of states that, for any $\Phi \in \mathcal{W}_1$ and $\rho \in E_1$,

$$\begin{aligned} \lim_{\ell \to \infty} \frac{1}{|\Lambda_\ell|} \rho(H_{\Lambda_\ell}^{\Phi}) &= \sum_{\Lambda \in \mathcal{P}_f,\, 0 \in \Lambda} \frac{1}{|\Lambda|} \rho(\Phi(\Lambda)) \\ &= \rho \left(\sum_{\Lambda \in \mathcal{P}_f,\, 0 \in \Lambda} \frac{1}{|\Lambda|} \Phi(\Lambda) \right) = \rho(e_\Phi) . \end{aligned}$$

□

The last proposition leads to a definition of energy density functionals on the space of invariant states:

Definition 5.12 (Energy Density Functionals) For any invariant interaction $\Phi \in \mathcal{W}_1$, we define the associated "energy density functional" $\mathfrak{e}_\Phi$ on the set E_1 of invariant states by

$$\mathfrak{e}_\Phi(\rho) \doteq \lim_{\ell \to \infty} \frac{1}{|\Lambda_\ell|} \rho(H_{\Lambda_\ell}^{\Phi}) = \rho(e_\Phi) .$$

Clearly, $\mathfrak{e}_\Phi$ is an affine functional on the convex set of invariant states. That is, for two invariant states $\rho, \rho' \in E_1$ and any constant $\lambda \in [0, 1]$,

$$\mathfrak{e}_\Phi(\lambda\rho + (1-\lambda)\rho') = \lambda\mathfrak{e}_\Phi(\rho) + (1-\lambda)\mathfrak{e}_\Phi(\rho') .$$

Additionally, for $\Phi, \Phi' \in \mathcal{W}_1$ and any invariant state $\rho \in E_1$,

$$|\mathfrak{e}_\Phi(\rho) - \mathfrak{e}_{\Phi'}(\rho)| \leq \|\Phi - \Phi'\| \ .$$

Observe that if the interaction $\Phi \in \mathcal{W}_1$ is gauge-invariant, i.e., $\Phi \in \mathcal{W}_1^\circ$, then $e_\Phi \in \mathcal{U}^\circ$. Thus, the energy density functional $\mathfrak{e}_\Phi$ is (globally) gauge-invariant in this case:

Lemma 5.13 (Gauge Invariance of the Energy Density) *For all gauge-invariant interactions $\Phi \in \mathcal{W}_1^\circ$, any global gauge automorphism α_ϕ of $\mathcal{U}$, with $\phi \in \mathbb{R}$ in the fermion case and $\phi \in \mathrm{SU}(N)$ in the quantum spin case, and any invariant state $\rho \in E_1$, one has that*

$$\mathfrak{e}_\Phi(\rho \circ \alpha_\phi) = \mathfrak{e}_\Phi(\rho) \ .$$

One important use of energy density functionals concerns the definition of a natural one-to-one-correspondence between the set E_1 of invariant states and a convex set of norm-one elements of $\mathcal{W}_1^{\mathrm{td}}$, the topological dual space of the separable Banach space $\mathcal{W}_1$:

Definition 5.14 (Invariant States as Linear Functionals on $\mathcal{W}_1$) With any invariant state $\rho \in E_1$, we associate the linear functional $\rho^{\mathcal{W}_1} \in \mathcal{W}_1^{\mathrm{td}}$ defined by

$$\rho^{\mathcal{W}_1}(\Phi) \doteq -\mathfrak{e}_\Phi(\rho) \ , \qquad \Phi \in \mathcal{W}_1 \ .$$

With this prescription, the space $E_1 \subseteq \mathcal{U}^*$ of invariant states can be canonically identified with a weak*-compact convex set of continuous linear functionals on $\mathcal{W}_1$: Observe first that the set

$$E_1^{\mathcal{W}_1} \doteq \{\rho^{\mathcal{W}_1} \ : \ \rho \in E_1\} \subseteq \mathcal{W}_1^{\mathrm{td}}$$

is norm-bounded and that $\mathcal{W}_1$ is a *separable* Banach space. Thus, exactly as in the case of states on a separable C^*-algebra, the weak* topology of $E_1^{\mathcal{W}_1}$ is given by a metric. See Sect. 4.5.1. In fact, the energy density functional $\mathfrak{e}_\Phi$ defines an affine weak*-homeomorphism $\rho \mapsto \rho^{\mathcal{W}_1}$ from E_1 to $\mathcal{W}_1^{\mathrm{td}}$ which is additionally a norm isometry:

Theorem 5.15 *The mapping $\rho \mapsto \rho^{\mathcal{W}_1}$ from E_1 to $\mathcal{W}_1^{\mathrm{td}}$ has the following properties:*

(i) *It is affine, i.e., for any invariant states $\rho_1, \rho_2 \in E_1$ and $\lambda \in [0, 1]$,*

$$\lambda\rho_1 + (1-\lambda)\rho_2 \mapsto \lambda\rho_1^{\mathcal{W}_1} + (1-\lambda)\rho_2^{\mathcal{W}_1} \ .$$

(ii) *It is isometric, i.e., for any invariant states* $\rho_1, \rho_2 \in E_1$,

$$\left\| \rho_1^{\mathcal{W}_1} - \rho_2^{\mathcal{W}_1} \right\|_{\mathrm{op}} = \|\rho_1 - \rho_2\|_{\mathrm{op}} \ .$$

(iii) *For every invariant state* $\rho \in E_1$, $\|\rho^{\mathcal{W}_1}\|_{\mathrm{op}} = 1$. *In particular, the image of the set* E_1 *of all invariant states under the above mapping is a norm-bounded convex subset of* $\mathcal{W}_1^{\mathrm{td}}$.

(iv) *It is a homeomorphism*[3] *from the set* E_1 *of invariant states to its image, with respect to the respective weak* topologies (of* E_1 *and its image under the mapping* $\rho \mapsto \rho^{\mathcal{W}_1}$*). In other words, the convex set* E_1 *can be canonically identified with a convex subset of* $\mathcal{W}_1^{\mathrm{td}}$.

Proof See [1, Section 4.5, in particular Lemma 4.18]. □

Because of the last theorem, by a slight abuse of terminology, we say that the continuous linear functional $\varphi \in \mathcal{W}_1^{\mathrm{td}}$ is an "invariant state," whenever $\varphi = \rho^{\mathcal{W}_1}$ for some (unique) invariant state of the fermion or spin C^*-algebra $\mathcal{U}$.

5.3 The Entropy Density of an Invariant State

In Sect. 3.2, we define the notion of von Neumann entropy of any state of the (C^*-)algebra of complex $n \times n$ matrices, $n \in \mathbb{N}$. See Definition 3.8. Although the fermion or spin C^*-algebra $\mathcal{U}$ is *infinite-dimensional*, we can still give a notion of entropy *density* for *invariant* states. This is done by considering the entropy in the increasing sequence $(\Lambda_\ell)_{\ell \in \mathbb{N}}$ of cubic boxes defined by (5.2), similar to what is done to define the energy density of Definition 5.12.

As the von Neumann entropy of Definition 3.8 is only defined for states of matrix (C^*-)algebras $\mathcal{L}(\mathbb{C}^n)$, $n \in \mathbb{N}$, we start by extending its definition to all C^*-algebras that are $*$-isomorphic to these special ones. In fact, there is such a natural extension, as a consequence of the following lemma:

Lemma 5.16 *Let* $\mathcal{A}$ *be any* C^**-algebra which is* $*$*-isomorphic to* $\mathcal{L}(\mathbb{C}^n)$, $n \in \mathbb{N}$, *and let* $\Theta, \Theta' : \mathcal{A} \to \mathcal{L}(\mathbb{C}^n)$ *be any two* $*$*-isomorphisms. Let* $\tilde{\rho} \in E(\mathcal{A})$ *and define the states* $\rho, \rho' \in E(\mathbb{C}^n)$ *by*

$$\rho \doteq \tilde{\rho} \circ \Theta^{-1} \qquad \text{and} \qquad \rho' \doteq \tilde{\rho} \circ \Theta'^{-1} \ .$$

There is a unitary $U \in \mathcal{L}(\mathbb{C}^n)$ *such that* $D_\rho = U^* D_{\rho'} U$, *where* $D_\rho, D_{\rho'} \in \mathcal{L}(\mathbb{C}^n)^+$ *are the density matrices associated with, respectively,* ρ *and* ρ'. *See Definition 2.41.*

[3] Recall that a homeomorphism is a bijection (one-to-one and onto) such that both the function and its inverse are continuous.

Proof

1. For any $\Omega \in \mathbb{C}^n$, $\|\Omega\| = 1$, there is $\Omega' \in \mathbb{C}^n$, $\|\Omega'\| = 1$, such that, for all $A \in \mathcal{L}(\mathbb{C}^n)$, one has

$$\langle \Omega, \Theta \circ \Theta'^{-1}(A)\Omega \rangle = \langle \Omega', A\Omega' \rangle \ .$$

In particular, for all $A \in \mathcal{A}$,

$$\langle \Omega, \Theta(A)\Omega \rangle = \langle \Omega', \Theta'(A)\Omega' \rangle \doteq \tilde{\rho}(A) \ ,$$

where $\tilde{\rho}$ is some state of $\mathcal{A}$.

2. In order to prove this fact, note that the composition of a pure state with a $*$-isomorphism is again a pure (or extreme) state, and recall that the pure (or extreme) states of $E(\mathbb{C}^n)$ are precisely the vector states. See Exercise 2.43. Observe also that every vector $x \neq 0 \in \mathbb{C}^n$ is cyclic with respect to the representations $(\mathbb{C}^n, \Theta)$ and $(\mathbb{C}^n, \Theta')$, as $\Theta(\mathcal{A}) = \Theta'(\mathcal{A}) = \mathcal{L}(\mathbb{C}^n)$. Hence, $(\mathbb{C}^n, \Theta, \Omega)$ and $(\mathbb{C}^n, \Theta', \Omega')$ are two cyclic representations associated with the state $\tilde{\rho} \in E(\mathcal{A})$. Consequently, by Lemma 4.112, they are unitarily equivalent:

$$\Theta(A) = U\Theta'(A)U^*$$

for some fixed unitary $U \in \mathcal{L}(\mathbb{C}^n)$ and all $A \in \mathcal{A}$. Hence, for all $A \in \mathcal{A}$,

$$\begin{aligned}
\mathrm{Tr}((U^* D_\rho U)A) &= \mathrm{Tr}(D_\rho(UAU^*)) \\
&= \mathrm{Tr}(D_\rho(U\Theta'(\Theta'^{-1}(A))U^*)) \\
&= \mathrm{Tr}(D_\rho\Theta(\Theta'^{-1}(A))) \\
&= \rho(\Theta(\Theta'^{-1}(A))) \\
&= \tilde{\rho}(\Theta^{-1} \circ \Theta(\Theta'^{-1}(A))) \\
&= \rho'(A) = \mathrm{Tr}(D_{\rho'}A) \ .
\end{aligned}$$

Thus, by the uniqueness of density matrices (Corollary 2.42), $U^* D_\rho U = D_{\rho'}$.

□

Recall from Exercise 3.3 that, for $n \in \mathbb{N}$ and any invariant state $\rho \in E(\mathbb{C}^n)$ with density matrix $D_\rho \in \mathcal{L}(\mathbb{C}^n)^+$,

$$\eta(UD_\rho U^*) = U\eta(D_\rho)U^* \ ,$$

where $\eta : [0, 1] \to \mathbb{R}$ is the function appearing in Definition 3.8 of the von Neumann entropy. Thus, from the above lemma and cyclicity of the trace, we obtain that

$$\begin{aligned}\mathrm{Tr}\ \eta(D_\rho) &= \mathrm{Tr}\ \eta(UD_{\rho'}U^*) \\ &= \mathrm{Tr}\ U\eta(D_{\rho'})U^* \\ &= \mathrm{Tr}\ U^*U\ \eta(D_{\rho'}) \\ &= \mathrm{Tr}\ \eta(D_{\rho'})\ .\end{aligned}$$

In other words, both states ρ and ρ' of the matrix algebra $\mathcal{L}(\mathbb{C}^n)$ in Lemma 5.16 have exactly the same von Neumann entropy. See also Lemma 3.29. This property motivates the following definition:

Definition 5.17 (von Neumann Entropy) Let $\mathcal{A}$ be a C^*-algebra which is $*$-isomorphic to $\mathcal{L}(\mathbb{C}^n)$, for some $n \in \mathbb{N}$. For every state $\rho \in E(\mathcal{A})$, we define its von Neumann entropy to be

$$S_{\mathcal{A}}(\rho) \doteq S(\rho \circ \Theta^{-1})\ ,$$

where $\Theta : \mathcal{A} \to \mathcal{L}(\mathbb{C}^n)$ is any $*$-isomorphism.

From the remark prior to the last definition, one also obtains the following import invariance property of the von Neumann entropy $S_{\mathcal{A}}$ just defined above:

Lemma 5.18 (Invariance of the von Neumann Entropy) *Let $\mathcal{A}$ be a C^*-algebra which is $*$-isomorphic to $\mathcal{L}(\mathbb{C}^n)$ for some $n \in \mathbb{N}$. For every state $\rho \in E(\mathcal{A})$ and any $*$-automorphism Θ of $\mathcal{A}$, $S_{\mathcal{A}}(\rho \circ \Theta) = S_{\mathcal{A}}(\rho)$.*

Proof This is an obvious consequence of the fact that, for two $*$-isomorphisms $\Theta, \Theta' : \mathcal{A} \to \mathcal{L}(\mathbb{C}^n)$, $\Theta \circ (\Theta')^{-1}$ is another $*$-isomorphism with inverse given by $\Theta'' \doteq \Theta' \circ \Theta^{-1}$. In particular, in this case, by definition of $S_{\mathcal{A}}$,

$$S_{\mathcal{A}}(\rho) = S(\rho \circ (\Theta'')^{-1})\ ,$$

while

$$S_{\mathcal{A}}(\rho \circ \Theta) = S(\rho \circ \Theta \circ (\Theta')^{-1}) = S(\rho \circ (\Theta' \circ \Theta^{-1})^{-1}) = S(\rho \circ (\Theta'')^{-1})\ .$$

□

For any finite region $\Lambda \in \mathcal{P}_f$, $\mathcal{U}_\Lambda$ is $*$-isomorphic to a matrix algebra in both (quantum spin and fermion) cases, thanks to Lemma 4.165. So, Lemma 5.16, along with Proposition 3.13, motivates the following definition for Gibbs states of such local algebras:

Definition 5.19 (Gibbs States for Finite Regions) Fix some nonempty $\Lambda \in \mathcal{P}_f$. For any (self-conjugate) element $H \in \mathrm{Re}\{\mathcal{U}_\Lambda\}$ and any inverse temperature $\beta \in$

$(0,\infty)$, we define the state $\omega_{H,\beta} \in E(\mathcal{U}_\Lambda)$ as being the unique minimizer of the free energy functional $F_{H,\beta} : E(\mathcal{U}_\Lambda) \to \mathbb{R}$ defined by

$$F_{H,\beta}(\rho) \doteq \rho(H) - \beta^{-1} S_{\mathcal{U}_\Lambda}(\rho)\ , \qquad \rho \in E(\mathcal{U}_\Lambda)\ .$$

By obvious reasons, this state is called the "Gibbs state of $\mathcal{U}_\Lambda$ associated with (the Hamiltonian) H at inverse temperature β." As in the matrix case, the quantity

$$P_{H,\beta} \doteq -\frac{1}{|\Lambda|} \inf F_{H,\beta}(E(\mathcal{U}_\Lambda))$$

is called the "pressure" at inverse temperature β associated with the Hamiltonian H.

Observe that the uniqueness of $\omega_{H,\beta}$ directly follows from Lemma 5.16 combined with Corollary 3.24. In the fermion case, note that if the Hamiltonian H is an even element, i.e., $H \in \mathrm{Re}\{\mathcal{U}_\Lambda^e\}$, then the Gibbs state $\omega_{H,\beta}$ is an even state, i.e., $\omega_{H,\beta} \in E_\mathrm{e}(\mathcal{U}_\Lambda)$. This follows from the uniqueness of the Gibbs states combined with the invariance of the entropy $S_{\mathcal{U}_\Lambda}$ with respect to $*$-automorphisms (Lemma 5.18).

Let ρ be any state of the (spin or fermion) C^*-algebra $\mathcal{U}$, and, for any $\ell \in \mathbb{N}$, let $\rho_\ell \in E(\mathcal{U}_{\Lambda_\ell})$ be the restriction of ρ to $\mathcal{U}_{\Lambda_\ell} \subsetneq \mathcal{U}$. Recall that, by Lemma 4.165, for any $\ell \in \mathbb{N}, \mathcal{U}_{\Lambda_\ell}$ is $*$-isomorphic to a matrix algebra. In particular, the von Neumann entropy of ρ_ℓ is well-defined. This entropy is denoted here by

$$S_\ell(\rho) \doteq S_{\mathcal{U}_{\Lambda_\ell}}(\rho_\ell)\ , \qquad \ell \in \mathbb{N}\ .$$

If $\rho \in E_1$ is an invariant state, it turns out that the finite volume entropy has a well-defined density in the infinite volume limit:

Theorem 5.20 (Existence of Entropy Densities) *For any invariant state $\rho \in E_1$,*

$$\mathfrak{s}(\rho) \doteq \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|} S_\ell(\rho) = \inf\left\{\frac{1}{|\Lambda_\ell|} S_\ell(\rho)\ :\ \ell \in \mathbb{N}\right\} \geq 0\ .$$

The entropy density functional $\mathfrak{s} : \rho \mapsto \mathfrak{s}(\rho)$ from E_1 to $\mathbb{R}_0^+$ is affine and bounded on the (convex) set E_1 of all invariant states.

The proof of the above theorem is quite involved and will not be presented here. For a modern approach, based on the theory of non-commutative conditional expectations, see [15, Theorem 10.3]. Note that [15] only considers the fermion case, but the case of quantum spin systems could be included in the proofs, by simple adaptations. For a proof for the specific case of quantum spins, see, for instance, [55, 70].

From Lemma 5.18 and the fact that global gauge automorphisms preserve the local algebras $\mathcal{U}_{\Lambda_\ell}$, $\ell \in \mathbb{N}$, similar to the energy density functionals for gauge-invariant interactions (see Lemma 5.13), the entropy density is invariant under global gauge transformations:

Corollary 5.21 (Gauge Invariance of the Entropy Density) *For any global gauge automorphism α_ϕ of $\mathcal{U}$, with $\phi \in \mathbb{R}$ in the fermion case and $\phi \in \mathrm{SU}(N)$ in the quantum spin case, and any invariant state $\rho \in E_1$, one has that*

$$\mathfrak{s}(\rho \circ \alpha_\phi) = \mathfrak{s}(\rho) \ .$$

We are now in a position to define free energy density functionals on invariant states, which represent the analogue of Definition 3.9 for infinitely extended quantum systems:

Definition 5.22 (Free Energy Density Functionals) For any invariant interaction $\Phi \in \mathcal{W}_1$ and inverse temperature $\beta \in (0, \infty)$ we define the "free energy density functional" $\mathfrak{f}_{\Phi,\beta}$ on the set E_1 of all (spin or fermion) invariant states by

$$\mathfrak{f}_{\Phi,\beta}(\rho) \doteq \mathfrak{e}_\Phi(\rho) - \beta^{-1}\mathfrak{s}(\rho) \ .$$

The set of all minimizers of $\mathfrak{f}_{\Phi,\beta}$ is denoted by $M_{\Phi,\beta} \subseteq E_1$. The elements of $M_{\Phi,\beta}$ are called the "globally stable equilibrium states" at temperature $T = \beta^{-1}$ associated with the invariant interaction $\Phi \in \mathcal{W}_1$.

We say that there is a "(first-order) phase transition" for $\Phi \in \mathcal{W}_1$ at temperature $T = \beta^{-1}$, if $M_{\Phi,\beta}$ contains more than one element. Notice that in physics the terminology "first order transition" means that when the temperature decreases below a critical temperature, the considered physical system undergoes a phase transition with discontinuous intensive thermodynamic quantities, like the magnetisation or the particle density. This situation is related to the transition from a thermodynamic phase with a unique equilibrium state to a phase with non-unique equilibrium states. We use the same terminology to refer to the non-uniqueness of the equilibrium state, only. This use is standard in mathematics-oriented literature.

By combining Lemma 5.13 and Corollary 5.21, we arrive at the following result on the gauge invariance of the free energy density:

Proposition 5.23 *Fix any inverse temperature $\beta \in (0, \infty)$. For all gauge-invariant interactions $\Phi \in \mathcal{W}_1^\circ$, any global gauge automorphism α_ϕ of $\mathcal{U}$, with $\phi \in \mathbb{R}$ in the fermion case and $\phi \in \mathrm{SU}(N)$ in the quantum spin case, and any invariant state $\rho \in E_1$, one has that*

$$\mathfrak{f}_{\Phi,\beta}(\rho \circ \alpha_\phi) = \mathfrak{f}_{\Phi,\beta}(\rho) \ .$$

The following definition refers to a physically very important case of first-order phase transition:

Definition 5.24 (Spontaneous Breaking of the Global Gauge Symmetry) Take a gauge-invariant interaction $\Phi \in \mathcal{W}_1^\circ$. We say that this interaction "spontaneously breaks the (global) gauge symmetry" at temperature $T = \beta^{-1}$, $\beta \in (0, \infty)$, if the corresponding set of (globally stable) equilibrium states $M_{\Phi,\beta}$ contains a state that is not invariant with respect to some (global) gauge automorphism, that is, there are

$\omega \in M_{\Phi,\beta}$ and a global gauge automorphism α_ϕ of $\mathcal{U}$ ($\phi \in \mathbb{R}$ in the fermion case and $\phi \in \mathrm{SU}(N)$ in the quantum spin case) such that $\omega \circ \alpha_\phi \neq \omega$.

By the last proposition, the spontaneous breaking of the gauge symmetry as defined above occurs only if a first-order phase transition takes place, i.e., the free energy density functional $\mathfrak{f}_{\Phi,\beta}$ has more than one minimizer. Thus, such a breaking is a special case of a first-order transition. In the case of fermions, the breaking of the global gauge symmetry by equilibrium states is related to superconductivity. In the quantum spin case, it is related to spontaneous magnetization.

Clearly, by the corresponding properties of $\mathfrak{e}_\Phi$ and $\mathfrak{s}$ (see discussion after Definition 5.12 as well as Theorem 5.20), the free energy density functional $\mathfrak{f}_{\Phi,\beta}$ is an affine and bounded mapping from the (convex) set E_1 of all invariant states to $\mathbb{R}$. In particular, $M_{\Phi,\beta}$ is convex. Recall that a function h on a convex set K is affine iff

$$h(\lambda x + (1-\lambda)\, y) = \lambda h(x) + (1-\lambda) h(y)$$

for all $x, y \in K$. Additionally, the set of minimizers of the free energy density functional is nonempty and has extreme points:

Theorem 5.25 *For all $\Phi \in \mathcal{W}_1$ and $\beta \in (0,\infty)$, the convex set $M_{\Phi,\beta}$ is nonempty and has extreme points.*

Proof The proof is provided in a more general setting in the next chapter. Basically, it uses the weak*-topology of Sect. 4.5.1 and we refrain from doing it in detail at this point, to minimize topological arguments in the present chapter. We only mention here that it is based on the lower weak*-semicontinuity of the functional $\mathfrak{f}_{\Phi,\beta}$ (Lemmata 6.7 and 6.8) as well as the weak*-compactness of the set E_1 of invariant states. With these observations, by Proposition 7.172, it follows that the convex set $M_{\Phi,\beta}$ is nonempty and weak*-closed (and thus weak*-compact). Observe also that $M_{\Phi,\beta}$ has extreme points, thanks to Proposition 7.334. □

The extreme points of the convex set $M_{\Phi,\beta}$ are called "pure globally stable equilibrium states" at temperature $T = \beta^{-1}$, associated with the invariant interaction $\Phi \in \mathcal{W}_1$. Observe meanwhile that pure globally stable equilibrium states are necessarily ergodic states:

Lemma 5.26 *For all $\Phi \in \mathcal{W}_1$ and $\beta \in (0,\infty)$, $\omega \in M_{\Phi,\beta}$ is a pure globally stable equilibrium state iff it is ergodic (i.e., it is extreme in the convex set of all invariant states).*

Proof If $\omega \in \mathcal{E}_1 \cap M_{\Phi,\beta}$, i.e., ω is ergodic (or extreme) in E_1, it cannot be decomposed within $M_{\Phi,\beta} \subseteq E_1$, and it is thus a pure globally stable equilibrium state. Conversely, let ω be an extreme point of $M_{\Phi,\beta}$, and assume, by contradiction, that there are invariant states $\omega', \omega'', \omega' \neq \omega''$, and a constant $\lambda \in (0,1)$ such that

$$\omega = \lambda\omega' + (1-\lambda)\omega'' .$$

In particular, by affineness of $\mathfrak{f}_{\Phi,\beta}$,

$$\mathfrak{f}_{\Phi,\beta}(\omega) = \lambda \mathfrak{f}_{\Phi,\beta}(\omega') + (1-\lambda)\mathfrak{f}_{\Phi,\beta}(\omega'') \ .$$

For ω is a minimizer of $\mathfrak{f}_{\Phi,\beta}$ and λ, $(1-\lambda) > 0$, this implies that also ω' and ω'' minimize $\mathfrak{f}_{\Phi,\beta}$, i.e., $\omega', \omega'' \in M_{\Phi,\beta}$. As ω is an extreme point of $M_{\Phi,\beta}$, one thus would have $\omega' = \omega''$, a contradiction. Hence, $\omega \in M_{\Phi,\beta}$ is ergodic whenever it is a pure globally stable equilibrium state. □

Because of Theorem 5.25 and Lemma 5.26, one has the following equivalent characterization of a first-order phase transition:

Proposition 5.27 *For any $\Phi \in \mathcal{W}_1$ and $\beta \in (0,\infty)$, there is a first-order phase transition iff $M_{\Phi,\beta}$ contains an invariant state which is non-ergodic.*

As the free energy density functional is bounded on the set E_1 of invariant states for any invariant interaction and inverse temperature, similar to Definition 3.16 for the case of finite quantum systems, we can define the pressure of infinitely extended systems as a function on the Banach space $\mathcal{W}_1$ of invariant interactions:

Definition 5.28 (Pressure Function on $\mathcal{W}_1$) For any fixed $\beta \in (0,\infty)$, the function $\mathfrak{p}_\beta : \mathcal{W}_1 \to \mathbb{R}$ defined by

$$\Phi \mapsto \mathfrak{p}_\beta(\Phi) \doteq -\inf \mathfrak{f}_{\Phi,\beta}(E_1) = \sup\{\rho^{\mathcal{W}_1}(\Phi) + \beta^{-1}\mathfrak{s}(\rho) \ : \ \rho \in E_1\}$$

is called "pressure function" at temperature $T = \beta^{-1}$.

Observe that

$$\sup\{\rho^{\mathcal{W}_1}(\Phi) + \beta^{-1}\mathfrak{s}(\rho) : \rho \in E_1\}$$

is the so-called "Legendre-Fenchel" transform at $\Phi \in \mathcal{W}_1$ of the entropy density functional $\mathfrak{s}$ times the constant $-\beta^{-1}$, seen as a function on the weak*-compact, norm-bounded, and convex set

$$E_1^{\mathcal{W}_1} \doteq \{\rho^{\mathcal{W}_1} : \rho \in E_1\} \subseteq \mathcal{W}_1^{\mathrm{td}}$$

of continuous linear functionals on $\mathcal{W}_1$. See Definition 7.348, as well as Definition 5.14 and remarks thereafter. By well-known results on the Legendre-Fenchel transform (see Sect. 7.5.5), the expression defining the pressure function in the last definition can be inverted to represent the entropy density as a function of the pressure:

Proposition 5.29 (Inverse Formula for the Entropy) *Fix $\beta \in (0,\infty)$. For any continuous linear functional $\varphi \in \mathcal{W}_1^{\mathrm{td}}$, the set of real numbers*

$$\{\varphi(\Phi) - \mathfrak{p}_\beta(\Phi) : \Phi \in \mathcal{W}_1\} \subseteq \mathbb{R}$$

is bounded from above iff $\varphi = \rho^{\mathcal{W}_1}$ *for some invariant state* $\rho \in E_1$*, and, in this case, one has*

$$\mathfrak{s}(\rho) = -\beta \sup_{\Phi \in \mathcal{W}_1} \{\rho^{\mathcal{W}_1}(\Phi) - \mathfrak{p}_\beta(\Phi)\} .$$

Proof With the remark prior to this proposition, the assertion directly follows from Theorem 7.353. See also Definition 7.351. □

Similar to the finite-dimensional case (see, e.g., Proposition 3.17), the pressure function at a given temperature is convex and continuous on the Banach space $\mathcal{W}_1$. In particular, it has tangent functionals (Definition 3.18), which are identified with invariant states, via the norm-isometric, affine homeomorphism $\rho \mapsto \rho^{\mathcal{W}_1}$ from E_1 to $\mathcal{W}_1^{\mathrm{td}}$ of Definition 5.14 and Theorem 5.15. This refers to the following proposition and theorem:

Proposition 5.30 *Given a fixed inverse temperature* $\beta \in (0, \infty)$*, the pressure function* $\mathfrak{p}_\beta : \mathcal{W}_1 \to \mathbb{R}$ *has the following properties:*

(i) *It is convex and continuous, with*

$$|\mathfrak{p}_\beta(\Phi) - \mathfrak{p}_\beta(\Phi')| \leq \left\| \Phi - \Phi' \right\| , \qquad \Phi, \Phi' \in \mathcal{W}_1 .$$

(ii) *For all* $\Phi \in \mathcal{W}_1$ *and* $\omega \in M_{\Phi,\beta}$*, the linear functional* $\omega^{\mathcal{W}_1} \in \mathcal{W}_1^{\mathrm{td}}$ *is tangent to* $\mathfrak{p}_\beta$ *at* Φ*. Additionally, if* $\varphi \in \mathcal{W}_1^{\mathrm{td}}$ *is tangent to* $\mathfrak{p}_\beta$ *at some* $\Phi \in \mathcal{W}_1$*, then* $\|\varphi\|_{\mathrm{op}} = 1$.

Proof Exercise. □

Observe from the above proposition and the Mazur theorem (Proposition 3.21) that, for all $\beta \in (0, \infty)$, the set

$$\widetilde{\mathcal{W}}_1(\beta) \doteq \{\Phi \in \mathcal{W}_1 \ : \ |M_{\Phi,\beta}| = 1\} \subseteq \mathcal{W}_1$$

of invariant (spin or fermion) interactions for which no phase transition appears is dense in $\mathcal{W}_1$. Thus, the absence of phase transitions is "typical," in a sense. In fact, this is just saying that if a physical system shows a mixed phase at equilibrium, then an arbitrarily small perturbation in the interaction of particles may force the system to "choose" one of the components of the mixture. This is in perfect accordance with our common intuition. What is more, as we will prove in the sequel, phase transitions do occur, at any fixed $\beta \in (0, \infty)$, for some invariant interaction $\Phi \in \mathcal{W}_1$ (depending on β). One of the main ingredients of the proof of phase transitions is a strengthening of Proposition 5.30 (ii), which implies the following results:

Theorem 5.31 (Globally Stable Equilibria as Tangent Functionals) *For any fixed* $\Phi \in \mathcal{W}_1$ *and* $\beta \in (0, \infty)$*, the mapping* $\rho \mapsto \rho^{\mathcal{W}_1}$ *establishes a one-to-*

one correspondence between linear functionals that are tangent to $\mathfrak{p}_\beta$ *at* Φ *and the elements of* $M_{\Phi,\beta}$.

Proof Observe from Proposition 5.30 (ii) that $\rho \mapsto \rho^{\mathcal{W}_1}$ maps elements of $M_{\Phi,\beta}$ to continuous linear functional on $\mathcal{W}_1$ that are tangent to $\mathfrak{p}_\beta$ at $\Phi \in \mathcal{W}_1$. By Theorem 5.15 (ii), this mapping is injective. Again by Proposition 5.30 (ii), any linear functional on $\mathcal{W}_1$ that is tangent to $\mathfrak{p}_\beta$ at $\Phi \in \mathcal{W}_1$ is norm-one. Thus, it remains to prove that every continuous linear functional on $\mathcal{W}_1$ that is tangent to $\mathfrak{p}_\beta$ at $\Phi \in \mathcal{W}_1$ has to be equal to $\rho^{\mathcal{W}_1}$ for some $\rho \in M_{\Phi,\beta}$. Let $\varphi \in \mathcal{W}_1^{\mathrm{td}}$ be such a tangent functional. Then, by definition of tangent functionals,

$$\mathfrak{p}_\beta(\Psi) - \mathfrak{p}_\beta(\Phi) \geq \varphi(\Psi - \Phi) , \qquad \Psi \in \mathcal{W}_1 .$$

In other words,

$$\mathfrak{p}_\beta(\Phi) - \varphi(\Phi) = \min\{\mathfrak{p}_\beta(\Psi) - \varphi(\Psi) \ : \ \Psi \in \mathcal{W}_1\} . \tag{5.12}$$

In fact, from this, φ has to be equal to $\rho^{\mathcal{W}_1}$ for some $\rho \in E_1$, because, otherwise,

$$\{\mathfrak{p}_\beta(\Psi) - \varphi(\Psi) : \Psi \in \mathcal{W}_1\}$$

would not be bounded from below, by Proposition 5.29. To conclude the proof, we now show that ρ minimizes the free energy density functional $\mathfrak{f}_{\Phi,\beta}$. This also directly follows from Proposition 5.29 and Eq. (5.12):

$$\begin{aligned} \mathfrak{f}_{\Phi,\beta}(\rho) &= -\rho^{\mathcal{W}_1}(\Phi) - \beta^{-1}\mathfrak{s}(\rho) = -\rho^{\mathcal{W}_1}(\Phi) + \sup_{\Psi\in\mathcal{W}_1} \{\rho^{\mathcal{W}_1}(\Psi) - \mathfrak{p}_\beta(\Psi)\} \\ &= -\rho^{\mathcal{W}_1}(\Phi) + \rho^{\mathcal{W}_1}(\Phi) - \mathfrak{p}_\beta(\Phi) = -\mathfrak{p}_\beta(\Phi) = \inf \mathfrak{f}_{\Phi,\beta}(E_1) . \end{aligned}$$

□

In fact, observe that the last theorem is an instance of a very general result related to the subgradients of Legendre-Fenchel transforms, which is given in Proposition 7.357.

5.4 Ergodic States

In this section, we discuss an equivalent characterization of extremality in the set E_1 of invariant states, which refers to vanishing dispersions of states at large volumes. This will be, later on, another one of the main arguments of the proof of existence of phase transitions for *infinitely extended* fermion and spin systems, which is one central purpose of the chapter.

To this end, we first need to exploit the invariance of states of E_1 in their cyclic representations, to be able to invoke the von Neumann ergodic theorem. Recall that cyclic representations of states are discussed in Sect. 4.6.2.

Proposition 5.32 *Let $\rho \in E_1$ be any invariant state with a cyclic representation denoted by $(H_\rho, \pi_\rho, \Omega_\rho)$. There is a unique family $\{U_x\}_{x\in\mathbb{Z}^d} \subseteq \mathcal{B}(H_\rho)$ of unitaries satisfying the following properties:*

(i) *For all $x \in \mathbb{Z}^d$, $U_x(\Omega_\rho) = \Omega_\rho$.*
(ii) *For all $x, x' \in \mathbb{Z}^d$, $U_0 = \mathrm{id}_{H_\rho}$ and $U_x U_{x'} = U_{x+x'}$.*
(iii) *For all $x \in \mathbb{Z}^d$ and algebra element $A \in \mathcal{U}$, $U_x \pi_\rho(A) U_x^* = \pi_\rho(\tau_x(A))$.*

Proof

1. Let

$$\tilde{H}_\rho \doteq \{\pi_\rho(A)\Omega_\rho \ : \ A \in \mathcal{U}\} \subseteq H_\rho \ .$$

Recall that, by cyclicity of Ω_ρ, $\tilde{H}_\rho$ is a dense subspace of H_ρ. Fix any $x \in \mathbb{Z}^d$. By invariance of the state $\rho \in E_1$, for any algebra element $A \in \mathcal{U}$,

$$\begin{aligned}\left\|\pi_\rho(\tau_x(A))\Omega_\rho\right\|^2 &= \rho(\tau_x(A)^*\tau_x(A)) \\ &= \rho(\tau_x(A^*A)) \\ &= \rho(A^*A) = \left\|\pi_\rho(A)\Omega_\rho\right\|^2 \ .\end{aligned}$$

From this, we conclude that, for all algebra elements $A, A' \in \mathcal{U}$ such that

$$\pi_\rho(A)\Omega_\rho = \pi_\rho(A')\Omega_\rho \ ,$$

i.e., $\pi_\rho(A - A')\Omega_\rho = 0$, one necessarily has $\pi_\rho(\tau_x(A - A'))\Omega_\rho = 0$, i.e.,

$$\pi_\rho(\tau_x(A))\Omega_\rho = \pi_\rho(\tau_x(A'))\Omega_\rho \ .$$

Additionally,

$$\left\|\pi_\rho(\tau_x(A))\Omega_\rho\right\| = \left\|\pi_\rho(A)\Omega_\rho\right\|$$

for every algebra element $A \in \mathcal{U}$.

2. From this, the linear mapping $\tilde{U}_x : \tilde{H}_\rho \to H_\rho$ defined by

$$\tilde{U}_x(\pi_\rho(A)\Omega_\rho) \doteq \pi_\rho(\tau_x(A))\Omega_\rho$$

for every algebra element $A \in \mathcal{U}$ is norm-preserving. Note that $\tilde{U}_x(\tilde{H}_\rho) = \tilde{H}_\rho$. As $\tilde{H}_\rho$ is a dense subspace of H_ρ, $\tilde{U}_x$ uniquely extends to a unitary $U_x \in \mathcal{B}(H_\rho)$.

Remark in particular that, for all $A, A' \in \mathcal{U}$,

$$\begin{aligned}\langle U_x \pi_\rho(A)\Omega_\rho, U_x \pi_\rho(A')\Omega_\rho\rangle &= \langle \pi_\rho(\tau_x(A))\Omega_\rho, \pi_\rho(\tau_x(A'))\Omega_\rho\rangle \\ &= \langle \Omega_\rho, \pi_\rho(\tau_x(A^* A'))\Omega_\rho\rangle \\ &= \rho(A^* A') \\ &= \langle \Omega_\rho, \pi_\rho(A^* A')\Omega_\rho\rangle \\ &= \langle \pi_\rho(A)\Omega_\rho, \pi_\rho(A')\Omega_\rho\rangle \ .\end{aligned}$$

3. Observe that

$$U_x(\Omega_\rho) = \tilde{U}_x(\pi_\rho(\mathbb{1})(\Omega_\rho)) = \pi_\rho(\tau_x(\mathbb{1}))(\Omega_\rho) = \pi_\rho(\mathbb{1})(\Omega_\rho) = \Omega_\rho \ .$$

For any algebra element $A \in \mathcal{U}$,

$$\begin{aligned}U_{-x}(U_x(\pi_\rho(A)(\Omega_\rho))) &= U_{-x}(\tilde{U}_x(\pi_\rho(A)(\Omega_\rho))) \\ &= U_{-x}(\pi_\rho(\tau_x(A))(\Omega_\rho)) \\ &= \tilde{U}_{-x}(\pi_\rho(\tau_x(A))(\Omega_\rho)) \\ &= \pi_\rho(\tau_{-x}(\tau_x(A)))(\Omega_\rho) \\ &= \pi_\rho(A)(\Omega_\rho) \ .\end{aligned}$$

By density of $\tilde{H}_\rho$ and continuity of U_x, this implies that

$$U_{-x} = U_x^{-1} = U_x^* \ .$$

4. Again by density of $\tilde{H}_\rho$ and continuity of U_x, in order to show that

$$U_x^* \pi_\rho(A) U_x = \pi_\rho(\tau_x(A)) \ ,$$

for $A \in \mathcal{U}$, it suffices to prove the equality

$$U_x \pi_\rho(A) U_x^*(\pi_\rho(A')(\Omega_\rho)) = \pi_\rho(\tau_x(A))(\pi_\rho(A')(\Omega_\rho))$$

for any algebra element $A' \in \mathcal{U}$. This is done as follows:

$$\begin{aligned}U_x \pi_\rho(A) U_x^*(\pi_\rho(A')(\Omega_\rho)) &= U_x \pi_\rho(A)(\pi_\rho(\tau_{-x}(A'))(\Omega_\rho)) \\ &= U_x(\pi_\rho(A\tau_{-x}(A'))(\Omega_\rho)) \\ &= \pi_\rho(\tau_x(A)A')(\Omega_\rho) \\ &= \pi_\rho(\tau_x(A))(\pi_\rho(A')(\Omega_\rho)) \ .\end{aligned}$$

5. Using similar arguments, one shows that $U_0 = \mathrm{id}_H$ and $U_x U_{x'} = U_{x+x'}$ for every $x, x' \in \mathbb{Z}^d$.

□

We are now in a position to state the celebrated von Neumann ergodic theorem, which is reminiscent of the law of large numbers, and use it in the sequel to study extreme invariant states:

Theorem 5.33 (von Neumann Ergodic Theorem) *Let H be any separable Hilbert space and $\{U_x\}_{x\in\mathbb{Z}^d} \subseteq \mathcal{B}(H)$, $d \in \mathbb{N}$, a family of unitary operators such that $U_0 = \mathrm{id}_H$ and $U_x U_{x'} = U_{x+x'}$ for every $x, x' \in \mathbb{Z}^d$. For all $\ell \in \mathbb{N}$, define the operator*

$$P_\ell \doteq \frac{1}{|\Lambda_\ell|} \sum_{x\in\Lambda_\ell} U_x \in \mathcal{B}(H) \ .$$

Then, the sequence $(P_\ell)_{\ell\in\mathbb{N}}$ converges in the strong operator topology to the orthogonal projector $P_\infty \in \mathcal{B}(H)$ whose range is the closed subspace

$$P_\infty(H) = \{v \in H \ : \ U_x(v) = v \text{ for all } x \in \mathbb{Z}^d\} \ .$$

Proof Define the unitaries

$$V_1^{\pm} = U_{(\pm 1,0,\ldots,0)} \ , \ V_2 = U_{(0,\pm 1,0,\ldots,0)} \ , \ldots , \ V_d = U_{(0,\ldots,0,\pm 1)} \ .$$

Then define the following set of vectors:

$$G \doteq \{v = V_k^{\pm}(w) - w \ : \ w \in H \ , \ k = 1, \ldots, d\} \subseteq H \ .$$

Observe, via simple computations, that, for all $v \in G^\perp$,

$$\left\langle V_k^{\pm}(v), w \right\rangle = \langle v, w \rangle \ , \qquad w \in H \ , \ k \in \{1, \ldots, d\} \ ,$$

which implies that, for all $v \in G^\perp$,

$$V_k^{\pm}(v) = v \ , \qquad k \in \{1, \ldots, d\} \ .$$

From this equality and the group property of $\{U_x\}_{x\in\mathbb{Z}^d}$, we conclude that $G^\perp \subseteq P_\infty(H)$. Conversely, one directly checks that any $v \in P_\infty(H)$ is orthogonal to G and, thus, $G^\perp = P_\infty(H)$. By Proposition 7.212, the Hilbert space H is the direct sum of $P_\infty(H)$ and $G^{\perp\perp}$, which, by Lemma 7.206 (iii) combined with Corollary 7.208, is the closure of the space of all linear combinations of vectors of G. Thus, by linearity and uniform continuity of the family $\{P_\ell\}_{\ell\in\mathbb{N}}$ of bounded linear operators on H, it suffices to prove the theorem for vectors of $P_\infty(H)$ and G. For $v \in P_\infty(H)$, the theorem is trivial. Now, if $v = V_k^{\pm}(w) - w$ for some $w \in H$

and $k \in \{1, \ldots, d\}$, i.e., $v \in G$,

$$P_\ell(v) = \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} (U_{x \pm (0,\ldots,0,1,0,\ldots,0)} - U_x)(w)$$
$$= \frac{1}{|\Lambda_\ell|} \left(\sum_{x \in \Lambda_\ell \pm (0,\ldots,0,1,0,\ldots,0)} U_x(w) - \sum_{x \in \Lambda_\ell} U_x(w) \right) ,$$

which clearly tends to zero, as $\ell \to \infty$, because most of the terms in the sums are cancelled. □

Proposition 5.32 together with the von Neumann ergodic theorem has the following important consequence:

Corollary 5.34 *Let $\rho \in E_1$ be any invariant state. For all $\ell \in \mathbb{N}$ and any (spin or fermion) algebra element $A \in \mathcal{U}$, define the space average*

$$A_\ell \doteq \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} \tau_x(A) .$$

Then, the sequence $(\rho(A_\ell^* A_\ell))_{\ell \in \mathbb{N}}$ converges in $\mathbb{C}$ and

$$\lim_{\ell \to \infty} \rho(A_\ell^* A_\ell) \in [|\rho(A)|^2, \|A\|^2] .$$

Proof

1. Let $(\pi_\rho, H_\rho, \Omega_\rho)$ be any cyclic representation associated with the invariant state $\rho \in E_1$. By Proposition 5.32, there is a family $\{U_x\}_{x \in \mathbb{Z}^d} \subseteq \mathcal{B}(H_\rho)$ of unitaries such that (i) $U_x(\Omega_\rho) = \Omega_\rho$ for all $x \in \mathbb{Z}^d$, (ii) $U_0 = \mathrm{id}_{H_\rho}$ and $U_x U_{x'} = U_{x+x'}$ for every $x, x' \in \mathbb{Z}^d$, and (iii) $U_x \pi_\rho(A) U_x^* = \pi_\rho(\tau_x(A))$ for any $x \in \mathbb{Z}^d$ and all algebra elements $A \in \mathcal{U}$.
2. Thus, for all $\ell \in \mathbb{N}$ and any algebra element $A \in \mathcal{U}$,

$$\rho(A_\ell^* A_\ell) = \frac{1}{|\Lambda_\ell|^2} \left\langle \Omega_\rho, \sum_{x \in \Lambda_\ell} \pi_\rho(\tau_x(A^*)) \sum_{x' \in \Lambda_\ell} \pi_\rho(\tau_{x'}(A))(\Omega_\rho) \right\rangle$$
$$= \frac{1}{|\Lambda_\ell|^2} \left\langle \sum_{x \in \Lambda_\ell} \pi_\rho(\tau_x(A))(\Omega_\rho), \sum_{x' \in \Lambda_\ell} \pi_\rho(\tau_{x'}(A))(\Omega_\rho) \right\rangle$$
$$= \left\| \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} \pi_\rho(\tau_x(A))(\Omega_\rho) \right\|^2$$

$$= \left\| \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} U_x \pi_\rho(A) U_x^*(\Omega_\rho) \right\|^2$$

$$= \left\| \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} U_x \pi_\rho(A) U_{-x}(\Omega_\rho) \right\|^2$$

$$= \left\| \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} U_x \pi_\rho(A)(\Omega_\rho) \right\|^2 = \left\| P_\ell(\pi_\rho(A)(\Omega_\rho)) \right\|^2 ,$$

where

$$P_\ell \doteq \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} U_x \ .$$

3. By the von Neumann ergodic theorem,

$$\lim_{\ell \to \infty} P_\ell(\pi_\rho(A)(\Omega_\rho)) = P_\infty(\pi_\rho(A)(\Omega_\rho)) \ ,$$

where $P_\infty \in \mathcal{B}(H)$ is the orthogonal projector whose range is the closed subspace

$$P_\infty(H_\rho) = \{v \in H_\rho \ : \ U_x(v) = v \text{ for all } x \in \mathbb{Z}^d\} \ .$$

From this, we conclude that $\rho(A_\ell^* A_\ell)$ converges, as $\ell \to \infty$, and

$$\lim_{\ell \to \infty} \rho(A_\ell^* A_\ell) = \| P_\infty(\pi_\rho(A)(\Omega_\rho)) \|^2 \leq \|A\|^2.$$

4. As $U_x(\Omega_\rho) = \Omega_\rho$ for all $x \in \mathbb{Z}^d$, $P_\infty(\Omega_\rho) = \Omega_\rho$. Using the orthogonal projection P_ρ, whose range is $P_\rho(H_\rho) = \mathbb{C}\Omega_\rho$, we thus compute that

$$\begin{aligned}
|\rho(A)|^2 &= \overline{\rho(A)}\rho(A) \\
&= \rho(A^*)\rho(A) \\
&= \langle \Omega_\rho, \pi_\rho(A^*)\Omega_\rho \rangle \langle \Omega_\rho, \pi_\rho(A)\Omega_\rho \rangle \\
&= \langle \Omega_\rho, \pi_\rho(A^*) P_\infty \Omega_\rho \rangle \langle P_\infty \Omega_\rho, \pi_\rho(A)\Omega_\rho \rangle \\
&= \langle \Omega_\rho, \pi_\rho(A^*) P_\infty \Omega_\rho \rangle \langle \Omega_\rho, P_\infty \pi_\rho(A)\Omega_\rho \rangle \\
&= \langle \Omega_\rho, \pi_\rho(A^*) P_\infty P_\rho P_\rho P_\infty \pi_\rho(A)\Omega_\rho \rangle \\
&= \langle P_\rho P_\infty \pi_\rho(A)\Omega_\rho, P_\rho P_\infty \pi_\rho(A)\Omega_\rho \rangle
\end{aligned}$$

$$= \| P_\rho P_\infty \pi_\rho(A)\Omega_\rho \|^2 \, .$$

In particular,

$$|\rho(A)|^2 \leq \| P_\infty(\pi_\rho(A)(\Omega_\rho))\|^2 \, .$$

From this estimate and the preceding steps, one arrives at

$$\lim_{\ell\to\infty} \rho(A_\ell^* A_\ell) \geq |\rho(A)|^2 \, .$$

□

Corollary 5.34 motivates the following definition, which turn out to characterize all ergodic states, i.e., extreme states of the convex set E_1 of invariant states:

Definition 5.35 (Dispersionless-at-Infinity States) We say that an invariant state $\rho \in E_1$ is "dispersionless at infinity" if

$$\lim_{\ell\to\infty} \rho(A_\ell^* A_\ell) = |\rho(A)|^2$$

for every algebra element $A \in \mathcal{U}$.

Let ρ be any state of a C^*-algebra $\mathcal{A}$. For any element $A \in \mathcal{A}$, the non-negative number

$$\sqrt{\rho((A - \rho(A)\mathfrak{1})^*(A - \rho(A)\mathfrak{1}))}$$

is called the (statistical) "dispersion" of A with respect to the state ρ. This number measures the (quantum) uncertainty of the quantity represented by A, with respect to its expected value $\rho(A)$, when the corresponding quantum system is found in the state ρ. Suppose that $\rho \in E_1$ is an invariant state (of the fermion or spin C^*-algebra $\mathcal{U}$), which is dispersionless at infinity. Then, for any algebra element $A \in \mathcal{U}$,

$$\begin{aligned}
&\lim_{\ell\to\infty} \sqrt{\rho((A_\ell - \rho(A_\ell)\mathfrak{1})^*(A_\ell - \rho(A_\ell)\mathfrak{1}))} \\
&\quad = \lim_{\ell\to\infty} \sqrt{\rho((A_\ell - \rho(A)\mathfrak{1})^*(A_\ell - \rho(A)\mathfrak{1}))} \\
&\quad = \lim_{\ell\to\infty} \sqrt{\rho(A_\ell^* A_\ell) - \overline{\rho(A)}\rho(A_\ell) - \overline{\rho(A_\ell)}\rho(A) + \overline{\rho(A)}\rho(A)} \\
&\quad = \lim_{\ell\to\infty} \sqrt{\rho(A_\ell^* A_\ell) - |\rho(A)|^2} \\
&\quad = \sqrt{\lim_{\ell\to\infty} \rho(A_\ell^* A_\ell) - |\rho(A)|^2} = 0 \, ,
\end{aligned}$$

noting that $\rho(A_\ell) = \rho(A)$ for any $\ell \in \mathbb{N}$. This property motivates the term "dispersionless at infinity." Compare this situation with the case of extreme states of the C^*-algebra of continuous functions on a compact space, as discussed after Corollary 4.73. In fact, also in the case of invariant states, extremality can be equivalently seen as vanishing statistical dispersions for the considered state, in an appropriate sense:

Lemma 5.36 *Every invariant state $\rho \in E_1$ (of the fermion or spin C^*-algebra $\mathcal{U}$) that is dispersionless at infinity is ergodic, i.e., $\rho \in \mathcal{E}_1$ is an extreme point of the convex set E_1 of all invariant states.*

Proof Assume that the invariant state $\rho \in E_1$ is dispersionless at infinity but not ergodic, i.e., $\rho \in E_1 \backslash \mathcal{E}_1$. Then, there are two invariant states $\rho', \rho'' \in E_1$ and some algebra element $A = A^* \in \mathcal{U}$ such that $\rho'(A) \neq \rho''(A)$ and $\rho = (\rho' + \rho'')/2$. By the strict convexity of the function $x \mapsto |x|^2$ on $\mathbb{R}$,

$$\begin{aligned} |\rho(A)|^2 &= \left|\frac{1}{2}\rho'(A) + \frac{1}{2}\rho''(A)\right|^2 < \frac{1}{2}|\rho'(A)|^2 + \frac{1}{2}|\rho''(A)|^2 \\ &\leq \frac{1}{2}\lim_{\ell\to\infty}(\rho'(A_\ell^2) + \rho''(A_\ell^2)) = \lim_{\ell\to\infty}\rho(A_\ell^2) \, . \end{aligned}$$

But this would contradict the fact that ρ is dispersionless at infinity. Thus, an invariant state $\rho \in E_1$ that is dispersionless at infinity has to be ergodic. □

The converse of the above lemma holds also true, and one arrives at the following important result:

Theorem 5.37 *An invariant state $\rho \in \mathcal{E}_1$ is ergodic iff it is dispersionless at infinity.*

Idea of Proof Lemma 5.36 already says that an invariant state that is dispersionless at infinity is ergodic. The proof that all ergodic states are dispersionless at infinity is quite involved. In fact, invariant states are even and can be canonically seen as states of the (even) subalgebra $\mathcal{U}^{\mathrm{e}} \subseteq \mathcal{U}$, thanks to Theorem 5.3 (fermion case; in the quantum spin case, this is trivial, since $\mathcal{U}^{\mathrm{e}} = \mathcal{U} = \mathrm{Spin}(N, \Gamma)$). $\mathcal{U}^{\mathrm{e}}$ is a non-commutative C^*-algebra, but it is "commutative at infinity." This term refers to the following property:

$$\lim_{|x|\to\infty}[A, \tau_x(B)] = \lim_{|x|\to\infty}(A\tau_x(B) - \tau_x(B)A) = 0$$

for all (even) algebra elements $A, B \in \mathcal{U}^e$. The property is also known as the "asymptotic abelianess" of $\mathcal{U}^{\mathrm{e}}$. It is crucial in order to get the equivalence between ergodicity and dispersionless at infinity. For the detailed proof, see [70, Theorem IV.2.7 in Section IV] in the quantum spin case or [1, Theorem 1.16 and Section 4.2] in the fermion case. □

5.5 Bishop-Phelps Theorem and the Existence of Phase Transitions

In this section, we prove the existence of (first-order) phase transitions in quantum lattice systems. In view of all results discussed above, the proof will consist in showing that the pressure function has tangent functionals that are non-ergodic states. See in particular Proposition 5.27 and Theorem 5.31. In order to achieve this, we use an important result of convex analysis in Banach spaces, the Bishop-Phelps theorem, via one corollary of it that is pertinent in our setting, namely, Corollary 7.365. Notice that the approach via the Bishop-Phelps theorem for proving the existence of phase transitions was proposed by Israel [70]. Here, we give a version of his original arguments.

For convenience, we reproduce below Corollary 7.365, for which a complete proof is provided in Sect. 7.5.7:

Corollary 5.38 (Bishop-Phelps Theorem—Version for Convex Functions I) *Let X be any real Banach space and $f : X \to \mathbb{R}$ a convex continuous function. Let $\varphi_0 \in X^{\mathrm{td}}$ be any continuous linear functional such that, for some constant $c < \infty$ and all $x \in X$,*

$$f(x) \geq \varphi_0(x) - c\ .$$

Then, for all $\varepsilon > 0$, there are $x_\varepsilon \in X$ and a linear functional $\varphi_\varepsilon \in X^{\mathrm{td}}$, which is tangent to f at x_ε and satisfies $\|\varphi_\varepsilon - \varphi_0\|_{\mathrm{op}} \leq \varepsilon$.

Now we are in a position to prove the existence of phase transitions, by combining the above corollary (i.e., Corollary 7.365) with Proposition 5.27 and Theorem 5.31:

Proposition 5.39 (Existence of First-Order Phase Transitions I) *For every inverse temperature $\beta \in (0, \infty)$, there is an invariant interaction $\Phi \in \mathcal{W}_1$ such that $|M_{\Phi,\beta}| > 1$, where $M_{\Phi,\beta} \subseteq E_1$ is the convex set of all minimizers of the free energy density functional $\mathfrak{f}_{\Phi,\beta}$ (Definition 5.22).*

Proof

1. Let $\rho', \rho'' \in E_1$, $\rho' \neq \rho''$, be two different invariant states. Observe that such states may be even constructed explicitly, for instance, by means of product states. Pick some self-adjoint element $A = A^* \in \mathcal{U}$ such that $\|A\| \leq 1$ and $\rho'(A) \neq \rho''(A)$, which must exist, by virtue of Exercise 4.64 (i). Define the convex combination $\rho \doteq (\rho' + \rho'')/2 \in E_1$. Then, by Corollary 5.34,

$$\begin{aligned}\lim_{\ell\to\infty} \rho(A_\ell^2) &= \frac{1}{2}\lim_{\ell\to\infty} \rho'(A_\ell^2) + \frac{1}{2}\lim_{\ell\to\infty} \rho''(A_\ell^2) \\ &\geq \frac{1}{2}\rho'(A)^2 + \frac{1}{2}\rho''(A)^2\end{aligned}$$

$$> \left(\frac{1}{2}\rho'(A) + \frac{1}{2}\rho''(A)\right)^2$$
$$= \rho(A)^2 \,. \tag{5.13}$$

2. Choose $\varepsilon > 0$ sufficiently small so that

$$3\varepsilon < \lim_{\ell\to\infty} \rho(A_\ell^2) - \rho(A)^2 \,. \tag{5.14}$$

   Recall Definitions 5.22 and 5.28: For fixed $\beta \in (0, \infty)$, the pressure function $\mathfrak{p}_\beta : \mathcal{W}_1 \to \mathbb{R}$ is defined by

$$\Phi \mapsto \mathfrak{p}_\beta(\Phi) \doteq -\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \doteq -\inf \left\{ \mathfrak{e}_\Phi(\rho) - \beta^{-1}\mathfrak{s}(\rho) : \rho \in E_1 \right\} .$$

   By Definition 5.14, one thus has the inequality

$$\mathfrak{p}_\beta(\Phi) \geq \rho^{\mathcal{W}_1}(\Phi) + \beta^{-1}\mathfrak{s}(\rho)$$

   for all $\Phi \in \mathcal{W}_1$ and every fixed invariant state $\rho \in E_1$. By Theorem 5.15 and Proposition 5.30, the assumptions of Corollary 5.38 are satisfied for $f = \mathfrak{p}_\beta$, $\varphi_0 = \rho^{\mathcal{W}_1}$, and $c = -\beta^{-1}\mathfrak{s}(\rho)$. Thus, for all $\varepsilon > 0$, there is an invariant interaction $\Psi \in \mathcal{W}_1$, as well as a tangent functional $\omega^{\mathcal{W}_1} \in \mathcal{W}_1^{\mathrm{td}}$ at Ψ, such that

$$\|\rho^{\mathcal{W}_1} - \omega^{\mathcal{W}_1}\|_{\mathrm{op}} = \|\rho - \omega\|_{\mathrm{op}} \leq \varepsilon \,, \tag{5.15}$$

   thanks to Theorem 5.15 (ii). Note that the (unique) preimage $\omega \in E_1$ of $\omega^{\mathcal{W}_1} \in \mathcal{W}_1^{\mathrm{td}}$ belongs to $M_{\Psi,\beta}$, by Theorem 5.31.
3. From Inequalities (5.13)–(5.15) and $\|A\|, \|A_\ell\| \leq 1$ (see Theorem 5.34 for the definition of A_ℓ), we finally compute for the above state $\omega \in M_{\Psi,\beta}$ that

$$\begin{aligned}\lim_{\ell\to\infty} \omega(A_\ell^2) - \omega(A)^2 &= \lim_{\ell\to\infty} \rho(A_\ell^2) - \rho(A)^2 + \lim_{\ell\to\infty} (\omega - \rho)(A_\ell^2) \\ &\quad -(\omega(A) - \rho(A))(\omega(A) + \rho(A)) \\ &\geq \left(\lim_{\ell\to\infty} \rho(A_\ell^2) - \rho(A)^2\right) - \varepsilon - 2\varepsilon > 0 \,.\end{aligned}$$

   Hence, $\omega \in M_{\Psi,\beta}$ is non-ergodic. As a consequence, by Proposition 5.27, $|M_{\Psi,\beta}| > 1$.

□

As demonstrated above, the Bishop-Phelps theorem provides a very elegant argument for the existence of phase transitions, that is, for the existence of invariant interactions $\Psi \in \mathcal{W}_1$ whose (globally stable) equilibrium state $\omega \in M_{\Psi,\beta}$ is not unique. However, the above proof gives absolutely no hint about the actual form of such an interaction. It turns out that, by a simple variation of the above argument, one can show that even so-called "two-body" interactions, which are particularly simple interactions, are able to produce phase transitions. With this aim, one uses Proposition 7.366, which is a stronger version of Corollary 7.365. Again, for convenience, we reproduce this proposition below:

Proposition 5.40 (Bishop-Phelps Theorem—Version for Convex Functions II) *Let X be any real Banach space, $f : X \to \mathbb{R}$ a* convex continuous *function, $C \subseteq X$ a* closed convex cone*, a linear functional such that*

$$f(x) \geq \varphi_0(x) - c$$

for some $c \in \mathbb{R}$ and all $x \in X$. Then there are a vector $\tilde{x} \in C$ and a linear functional $\tilde{\varphi} \in X^{\mathrm{td}}$ that is tangent to f at $\tilde{x}$, such that

$$\tilde{\varphi}(y) \geq \varphi_0(y) - \varepsilon \|y\|$$

for all $y \in C$.

As explained after Proposition 7.366, Corollary 5.38 is a special case of this proposition: It suffices to apply Proposition 5.40 to $C = X$ in order to arrive at Corollary 5.38.

Lemma 5.41 *Let $A \in \mathcal{U}_{\{0\}}^{\mathrm{e}}$ be any fixed* self-conjugate even *element of the "one-site algebra" $\mathcal{U}_{\{0\}}$. Take any function (two-body potential)* $\mathrm{v} : \Gamma \to \mathbb{R}$ *that is* reflection invariant and absolutely summable, *i.e.,* $\mathrm{v}(-x) = \mathrm{v}(x)$ *and*

$$\|\mathrm{v}\|_{\Gamma} \doteq \sum_{x \in \Gamma} |\mathrm{v}(x)| < \infty \, .$$

Then, define the invariant interaction $\Phi_{\mathrm{v}} \in \mathcal{W}_1$ by

$$\Phi_{\mathrm{v}}(\{x, x'\}) \doteq \mathrm{v}(x - x')\tau_x(A)\tau_{x'}(A) \, , \qquad x, x' \in \Gamma \equiv \mathbb{Z}^d \, ,$$

and $\Phi_{\mathrm{v}}(\Lambda) \doteq 0$ for all $\Lambda \in \mathcal{P}_f$ with $|\Lambda| \neq \{1, 2\}$. The set

$$\mathcal{W}_1^{2\mathrm{b}}(A) \doteq \{\Phi_{\mathrm{v}} \text{ with } \mathrm{v} : \Gamma \to \mathbb{R} \, , \; \mathrm{v}(-x) = \mathrm{v}(x) \, , \; \|\mathrm{v}\|_{\Gamma} < \infty\} \subseteq \mathcal{W}_1$$

is a closed subspace *of the Banach space $\mathcal{W}_1$.*

Proof Exercise. □

The elements of $\mathcal{W}_1^{2\mathrm{b}}(A)$ are called "two-body" interactions associated with the (one-site even) observable $A \in \mathrm{Re}\{\mathcal{U}_{\{0\}}^{\mathrm{e}}\}$. With this definition, from Proposition 5.40 (i.e., Proposition 7.366), we obtain the following stronger, and physically more relevant, version of Proposition 5.39:

Proposition 5.42 (Existence of First-Order Phase Transitions II) *For every inverse temperature* $\beta \in (0, \infty)$ *and every* $A \in \mathrm{Re}\{\mathcal{U}^{\mathrm{e}}_{\{0\}}\}\backslash\{0\}$*, there is a two-body interaction* $\Phi \in \mathcal{W}_1^{2\mathrm{b}}(A)$ *such that* $|M_{\Phi,\beta}| > 1$.

Proof Having Lemma 5.41 in mind ($\mathcal{W}_1^{2\mathrm{b}}(A)$ is a closed convex cone, being a closed subspace), by using Proposition 5.40 instead of Corollary 5.38, the proof of the present proposition is a straightforward adaptation of the proof of Proposition 5.39. □

Finally, we can go one step further and ask whether the sign of the two-body interaction (i.e., whether the interactions is purely attractive, purely repulsive, or without a definite sign) is important to have a phase transition, or not. It turns out that we can use Proposition 5.40 again, in order to prove that phase transitions can be produced by purely attractive two-body interactions. In fact, we say that the two-body interaction $\Phi \in \mathcal{W}_1^{2\mathrm{b}}(A)$ is "purely attractive (repulsive)" if, for all $\Lambda \in \mathcal{P}_f$, $-\mathcal{W}_1^{2\mathrm{b}}(A)(\Lambda)$ ($\mathcal{W}_1^{2\mathrm{b}}(A)(\Lambda)$) is a positive element of $\mathcal{U}_\Lambda$.

For any fixed $A \in \mathrm{Re}\{\mathcal{U}^{\mathrm{e}}_{\{0\}}\}$, let $\mathcal{W}_1^{2\mathrm{b},-}(A)$ ($\mathcal{W}_1^{2\mathrm{b},+}(A)$) denote that set of all purely attractive (repulsive) interactions of $\mathcal{W}_1^{2\mathrm{b}}(A)$. It is easy to see that these sets are *closed convex cones* of the Banach space $\mathcal{W}_1$. With this remark, we have the following result:

Proposition 5.43 (Existence of First-Order Phase Transitions III) *For every inverse temperature* $\beta \in (0, \infty)$ *and every* $A \in \mathrm{Re}\{\mathcal{U}^{\mathrm{e}}_{\{0\}}\}\backslash\{0\}$*, there is an* attractive *two-body interaction* $\Phi \in \mathcal{W}_1^{2\mathrm{b},-}(A)$ *such that* $|M_{\Phi,\beta}| > 1$.

Proof Having in mind that $\mathcal{W}_1^{2\mathrm{b},-}(A)$ is a closed cone of the Banach space $\mathcal{W}_1$, the proof uses again Proposition 5.40 via an adaptation of the proof of Propositions 5.39 and 5.42. We omit the details and only remark that the sign of the interaction matters in this proof, that is, the same type of argument would not yield a similar result for purely repulsive two-body interactions. □

To conclude, we notice that the last proposition should not give the impression that there are no phase transitions for (two-body) purely repulsive interactions: Phase transitions *do occur* also in this situation. However, the strategy of proof is different from the attractive case. In fact, in the repulsive case, one uses arguments (based on Choquet's theorem; see Theorem 7.339) similar to those presented in the next subsection for the proof of existence of spontaneous symmetry breaking, and we thus refrain from considering this case here.

5.6 Choquet's Theorem and Existence of Spontaneous Symmetry Breaking

In this section, we prove the existence of spontaneous symmetry breaking in quantum lattice systems. To keep the discussions as simple as possible, we only consider the particular case of spontaneous breaking of *gauge symmetry in* *fermion systems*, but the same arguments can be easily adapted to the quantum spin case. Thus, in this section, we set $\mathcal{U} = \mathrm{CAR}(\Omega, \Gamma)$. Recall that (gauge) symmetry breaking is a special (physically very important) form of (first-order) phase transition. See Definition 5.24 and remarks thereafter. The Bishop-Phelps theorem indeed plays again an important role in the proof (via Corollary 5.45 below), but the main argument of the proof is now a version of Choquet's theorem for states (see Definition 7.335 and Theorem 7.339), which we reproduce below, for convenience:

Theorem 5.44 (Choquet's Theorem—Version for States) *Let $\mathcal{A}$ be a separable unital C^*-algebra and $E \subseteq E(\mathcal{A})$ a nonempty closed* convex *set of states on $\mathcal{A}$. Every $\rho \in E$ is the barycenter of some $\mu \in E(C(E;\mathbb{C}))$, that is, $\rho \doteq \mu \circ \Xi$, whose unique extension to a probability measure is supported in the extreme boundary $\mathcal{E}(E)$ of the convex set E. Here, Ξ is the Gelfand transform*[4] *of Definition 4.79.*

For more details on this theorem, see Sect. 7.5.3.

As we show below, in the presence of non-ergodic equilibrium states (which is guaranteed by Bishop-Phelps-type arguments, as done in the last section), Choquet's theorem allows us to detect *ergodic* equilibrium states that break the gauge symmetry.

Note that the space $\mathcal{W}_1^{\circ} \subseteq \mathcal{W}_1$ of gauge-invariant is a closed subspace of the Banach space $\mathcal{W}_1$ of (invariant) interactions. Thus, from Proposition 5.40, we arrive at the following corollary:

Corollary 5.45 *Assume that the spin set Ω contains more than one element and let*

$$A \doteq a(\mathrm{e}_{0,s})a(\mathrm{e}_{0,s'}) \in \mathcal{U}_{\{0\}}^{\mathrm{e}}$$

for some $s, s' \in \Omega$ with $s \neq s'$, where $\mathrm{e}_{0,\tilde{s}} \in \ell^2(\Omega \times \Gamma)$, $\tilde{s} \in \Omega$, is the canonical basis element defined by $\mathrm{e}_{0,\tilde{s}}(x, \tilde{s}') \doteq 1$ if $(x, \tilde{s}') = (0, \tilde{s})$ and $\mathrm{e}_{0,\tilde{s}}(x, \tilde{s}') \doteq 0$, else. For every $\beta \in (0, \infty)$, there is a gauge-invariant interaction $\Phi \in \mathcal{W}_1^{\circ}$, as well as an equilibrium state $\omega \in M_{\Phi,\beta}$ and a strictly positive constant $C > 0$, such that

$$\omega(A_\ell^* A_\ell) \geq C$$

for all $\ell \in \mathbb{N}$, with A_ℓ defined as in Theorem 5.34.

Proof As $A \neq 0 \in \mathcal{U}_{\{0\}}^{\mathrm{e}}$, there is some (even) state $\rho_0 \in E(\mathcal{U}_{\{0\}}^{\mathrm{e}})$ such that $\rho_0(A) \neq 0$. See Exercise 4.64 (i). Then, consider the product state $\rho \in E_1$ defined by

$$\rho \doteq \bigotimes_{x \in \mathbb{Z}^d} \rho_0 \circ \tau_{-x}|_{\mathcal{U}_{\{x\}}} \, .$$

[4] By definition, Ξ is a mapping $\mathcal{A} \to C(E(\mathcal{A});\mathbb{C})$. It is naturally seen here as a mapping $\mathcal{A} \to C(E;\mathbb{C})$, by restriction of continuous functions on $E(\mathcal{A})$ to E.

See Proposition 4.193. By construction, we have

$$\rho(A_\ell^* A_\ell) = \frac{1}{|\Lambda_\ell|^2} \sum_{x,y\in\Lambda_\ell} \rho\left(\tau_x(A)\tau_y(A)\right) = |\rho_0(A)|^2 \geq \frac{1}{2}|\rho_0(A)|^2 \doteq 2C > 0 \, .$$

Observe that, for every $\ell \in \mathbb{N}$, there is a gauge-invariant interaction $\Phi_\ell \in \mathcal{W}_1^\circ$ such that $\|\Phi_\ell\| \leq \|A\|^2$ and $\tilde{\rho}(A_\ell^* A_\ell) = -\mathfrak{e}_{\Phi_\ell}(\tilde{\rho})$ for all $\tilde{\rho} \in E_1$. (We let the proof of this claim as an exercise.) Take $\varepsilon \doteq C\|A\|^{-2}$. Then, by Proposition 5.40 (combined with Proposition 5.27 and Theorem 5.31, as before), there is a gauge-invariant interaction $\Phi \in \mathcal{W}_1^\circ$, as well as an equilibrium state $\omega \in M_\Phi$, such that

$$\omega(A_\ell^* A_\ell) \geq \rho(A_\ell^* A_\ell) - C\|A\|^{-2}\|\Phi_\ell\| \geq C > 0$$

for all $\ell \in \mathbb{N}$. □

Now we are in a position to prove the existence of spontaneous symmetry breaking:

Proposition 5.46 (Existence of Spontaneous Symmetry Breaking) *Assume that the spin set Ω contains more than one element and let*

$$A \doteq a(\mathrm{e}_{0,s})a(\mathrm{e}_{0,s'}) \in \mathcal{U}_{\{0\}}^{\mathrm{e}}$$

for some $s, s' \in \Omega$ with $s \neq s'$, as in Corollary 5.45. For every $\beta \in (0,\infty)$, there is a gauge-invariant interaction $\Phi \in \mathcal{W}_1^\circ$, as well as an equilibrium state $\omega \in M_{\Phi,\beta}$, for which $\omega(A) \neq 0$. In particular, for all $\phi \in (0,\pi)$, one has that

$$\omega \circ \alpha_\phi(a(\mathrm{e}_{0,s})a(\mathrm{e}_{0,s'})) = e^{i2\phi}\omega(a(\mathrm{e}_{0,s})a(\mathrm{e}_{0,s'})) \neq \omega(a(\mathrm{e}_{0,s})a(\mathrm{e}_{0,s'})).$$

Proof

1. Let $\Phi \in \mathcal{W}_1^\circ$ and $\omega \in M_{\Phi,\beta}$ be given as in Corollary 5.45. Then,

$$\lim_{\ell\to\infty} \omega(A_\ell^* A_\ell) \geq C > 0 \, .$$

By Choquet's theorem (Theorem 5.44), there is a probability measure μ_ω on (the compact metric space) $M_{\Phi,\beta}$ that is supported (see Definition 4.11) in $\mathcal{E}(M_{\Phi,\beta})$, such that

$$\omega(A_\ell^* A_\ell) = \mu_\omega \circ \Xi(A_\ell^* A_\ell)$$

for all $\ell \in \mathbb{N}$. Now, by Proposition 4.12 and Corollary 5.34 (coming from the von Neumann ergodic theorem),

$$\lim_{\ell\to\infty} \omega(A_\ell^* A_\ell) = \mu_\omega\left(\lim_{\ell\to\infty} \omega(A_\ell^* A_\ell)\right) = \mu_\omega(\Delta_A) \, ,$$

where $\Delta_A \in \mathfrak{M}_\mathrm{b}(M_{\Phi,\beta};\mathbb{C})$ is defined by

$$\Delta_A(\rho) \doteq \lim_{\ell\to\infty} \rho(A_\ell^* A_\ell)\ , \qquad \rho \in M_{\Phi,\beta}\ .$$

See again Corollary 5.34, as well as Definition 6.9.

2. Now, as the probability measure is supported in $\mathcal{E}(M_{\Phi,\beta})$, which is a subset of $\mathcal{E}(E_1)$, $M_{\Phi,\beta}$ being a face of E_1, we again infer from Corollary 5.34 that

$$\lim_{\ell\to\infty} \omega(A_\ell^* A_\ell) = \mu_\omega(\delta_A)\ ,$$

where $\delta_A \in C_\mathrm{b}(M_{\Phi,\beta};\mathbb{C}) \subseteq \mathfrak{M}_\mathrm{b}(M_{\Phi,\beta};\mathbb{C})$ is defined by

$$\delta_A(\rho) \doteq |\rho(A)|^2\ , \qquad \rho \in M_{\Phi,\beta}\ .$$

See below Eq. (6.9). Thus, as

$$\lim_{\ell\to\infty} \omega(A_\ell^* A_\ell) \geq C > 0\ ,$$

there must be some $\omega' \in \mathcal{E}(M_{\Phi,\beta})$ such that $\omega'(A) \neq 0$.

□

5.7 A Brief Introduction to Hartree-Fock Theory

The study of equilibrium states of fermion or spin systems on the lattice amounts to solve the variational problem

$$\inf \mathfrak{f}_{\Phi,\beta}(E_1) = \inf\{\mathfrak{e}_\Phi(\rho) - \beta^{-1}\mathfrak{s}(\rho)\ :\ \rho \in E_1\} \tag{5.16}$$

for invariant interactions $\Phi \in \mathcal{W}_1$ and inverse temperatures $\beta \in (0,\infty)$. In other words, the aim is to determine the set $M_{\Phi,\beta} \subseteq E_1$ of minimizers of the free energy density $\mathfrak{f}_{\Phi,\beta}$. These minimizers are, by definition, the (infinite-volume globally stable) equilibrium states of the quantum lattice system, which is fixed by the choice of an interaction $\Phi \in \mathcal{W}_1$ and an inverse temperature $\beta \in (0,\infty)$. See Definition 5.22. This variational problem is by far nontrivial, in general, and concrete computations of correlation functions of the corresponding minimizers, which are of physical interest, can be a very hard task, if not impossible, even for apparently simple models like the celebrated Hubbard model (see Example 5.9).

This kind of problem appeared already at the beginning of quantum mechanics, as the number of degrees of freedom dramatically increases with the (unthinkable) number of quantum bodies that compose most of the physical systems. In the 1920s and 1930s, Hartree, Fock, and Slater set up a general method to simplify the analysis of *fermion* systems. This is known as the Hartree-Fock method, or theory, which is based on the assumption that the many-fermion wave function is a Slater determinant. Since then, this method has been largely used in theoretical and computational physics, as well as in quantum chemistry. In mathematics, important developments on the Hartree-Fock theory applied to Coulomb systems, i.e., for atoms and molecules, were done without interruption from the 1970s, in particular with Lieb and Simon's seminal works [73–75], along with the mathematical foundations of the Thomas-Fermi theory and the stability of matter. See [76]. In these works, among other things, the Hartree-Fock ground state energy has been progressively approximated in the limit of large nuclei charges for neutral atoms or molecules. In this series of approximations, the main contribution is given by the Thomas-Fermi energy [75], and various corrections to the Hartree-Fock ground state energy have been then explicitly derived and studied in the relativistic and nonrelativistic cases by many researchers (e.g., Bach, Ivrii, Hughes, Siedentop, Sigal, Solovej, Sørensen, Spitzer, Weikard). See [78, Section I] and references therein.

The (possibly unrestricted) Hartree-Fock theory and its variants, like the HFz approximation [80], are all based on the n-fermion Hilbert space and the use of Slater determinants to approximate the true ground state of the system. In fact, it refers to a *gauge-invariant Ansatz* to approximate the ground states of n fixed, possibly interacting, fermions. A natural extension of the Hartree-Fock theory to positive temperatures is to minimize the corresponding free energy in the set of simple quasi-free states, i.e., (globally) gauge-invariant quasi-free states. See Definition 4.190 and Proposition 4.221. Indeed, the Hartree-Fock theory can be rephrased in terms of simple quasi-free states as follows:

Let $(\mathcal{H}_{\mathrm{Fock}}, \pi_{\mathrm{Fock}}, \Omega_{\mathrm{Fock}})$ be any cyclic representation of $\mathrm{CAR}(\Omega, \Gamma)$ associated with the Fock state $\rho_{\mathrm{Fock}} \in E(\mathrm{CAR}(\Omega, \Gamma))$, as explained in Definition 4.177, for $\mathcal{G} = \ell^2(\Omega \times \Gamma)$. Recall from Definition 4.177 (ii)–(iii) that

$$x_1 \wedge \cdots \wedge x_n \doteq A(x_1)^* \cdots A(x_n)^* \Omega_{\mathrm{Fock}} \doteq \pi_{\mathrm{Fock}}(a(x_1)^*) \cdots \pi_{\mathrm{Fock}}(a(x_n)^*) \Omega_{\mathrm{Fock}}$$

for any $n \in \mathbb{N}$ and all $x_1, \ldots, x_n \in \ell^2(\Omega \times \Gamma)$, while we remind from Proposition 4.178 that for any $n, m \in \mathbb{N}$ and all $x_1, \ldots, x_n, x_1', \ldots, x_m' \in \ell^2(\Omega \times \Gamma)$,

$$\left\langle x_1 \wedge \cdots \wedge x_m, x_1' \wedge \cdots \wedge x_n' \right\rangle_{\mathrm{Fock}} = \begin{cases} \det\left[\left\langle x_i, x_j' \right\rangle\right]_{i,j=1}^n & \text{if} \quad m = n \\ 0 & \text{else} \end{cases},$$

with $\langle \cdot, \cdot \rangle_{\mathrm{Fock}}$ being the scalar product of the Hilbert space $\mathcal{H}_{\mathrm{Fock}}$, while $\langle \cdot, \cdot \rangle$ is the scalar product of $\ell^2(\Omega \times \Gamma)$. In particular, using the (canonical) Hilbert basis[5] $\{\mathfrak{e}_{\mathrm{s},y}\}_{(\mathrm{s},y) \in \Omega \times \Gamma}$ of $\ell^2(\Omega \times \Gamma)$, one observes that, for any $n \in \mathbb{N}$ and $(\mathrm{s}_1, y_1), \ldots, (\mathrm{s}_n, y_n) \in \Omega \times \Gamma$,

[5] That is, $\mathfrak{e}_{\mathrm{s},x}(\tilde{\mathrm{s}}, \tilde{x}) = 1$ if $(\mathrm{s}, x) = (\tilde{\mathrm{s}}, \tilde{x})$ and $\mathfrak{e}_{\mathrm{s},x}(\tilde{\mathrm{s}}, \tilde{x}) = 0$, else.

$$x_1 \wedge \cdots \wedge x_m \left((\mathrm{s}_1, y_1), \ldots, (\mathrm{s}_n, y_n)\right) \doteq \left\langle x_1 \wedge \cdots \wedge x_m, \mathrm{e}_{\mathrm{s}_1, y_1} \wedge \cdots \wedge \mathrm{e}_{\mathrm{s}_n, y_n} \right\rangle_{\mathrm{Fock}}$$
$$= \det \left[x_i \left(\mathrm{s}_j, y_j \right) \right]_{i,j=1}^n .$$

This shows that the "n-fermion wave function" $x_1 \wedge \cdots \wedge x_n \in H_{\mathrm{Fock}}$ is nothing else than a usual Slater determinant for $n \in \mathbb{N}$ fixed fermions, whose "one-body wave functions are" $x_1, \ldots, x_n \in \ell^2(\Omega \times \Gamma)$, respectively. See, for instance, [77, Equation (2a.l)]. What is more, these Slater determinants define simple quasi-free states:

Lemma 5.47 (From Slater Determinants to Simple Quasi-Free States) *For any $n \in \mathbb{N}$ and each non-zero vector*

$$\varphi \in \{\Omega_{\mathrm{Fock}}\} \cup \left\{ x_1 \wedge \cdots \wedge x_n \ : \ x_1, \ldots, x_n \in \ell^2(\Omega \times \Gamma) \right\}$$

in the Fock space H_{Fock}, the state ρ_{HF} defined on $\mathrm{CAR}(\Omega, \Gamma)$ *by*

$$\rho_{\mathrm{HF}}(A) = \|\varphi\|^{-2} \langle \varphi, \pi_{\mathrm{Fock}}(A)\varphi \rangle_{\mathrm{Fock}} , \qquad A \in \mathrm{CAR}(\Omega, \Gamma) ,$$

is a simple quasi-free state.

Proof If $\varphi = \Omega_{\mathrm{Fock}}$, then the assertion directly follows from Propositions 4.178 and 4.221 together with elementary properties of Pfaffians in relation with determinants. Now, fix $n, m \in \mathbb{N}$ and non-zero vectors $x_1, \ldots, x_n \in \ell^2(\Omega \times \Gamma)$. Let P be here the orthogonal projection on the vector space $\mathrm{span}\{x_1, \ldots, x_n\}^{\perp} \subseteq \ell^2(\Omega \times \Gamma)$ and $\varphi = x_1 \wedge \cdots \wedge x_n$. For any $y \in \ell^2(\Omega \times \Gamma)$, observe that

$$A(y)^* \varphi = A(P(y) + \left(\mathrm{id}_{H_{\mathrm{Fock}}} - P\right)(y))^* x_1 \wedge \cdots \wedge x_n$$
$$= A(P(y))^* x_1 \wedge \cdots \wedge x_n = A(P(y))^* \varphi = P(y) \wedge \varphi ,$$

by linearity of the mapping $y \mapsto A(y)^* \doteq \pi_{\mathrm{Fock}}(a(y)^*)$ (see the CAR relations given by Definition 4.163). Therefore, using Proposition 4.178, for any $m, p \in \mathbb{N}$ and $y_1, \ldots, y_m, y_1', \ldots, y_p' \in \ell^2(\Omega \times \Gamma)$, we obtain that, if $m = p$,

$$\begin{aligned}
&\rho_{\mathrm{HF}}(a(y_1) \cdots a(y_m) a(y_m')^* \cdots a(y_1')^*) \\
&= \frac{1}{\langle x_1 \wedge \cdots \wedge x_m, x_1 \wedge \cdots \wedge x_n \rangle_{\mathrm{Fock}}} \\
&\quad \times \left\langle P(y_1) \wedge \cdots \wedge P(y_m) \wedge \varphi, P\left(y_1'\right) \wedge \cdots \wedge P\left(y_m'\right) \wedge \varphi \right\rangle_{\mathrm{Fock}} \\
&= \frac{1}{\det \left[\langle x_i, x_j \rangle \right]_{i,j=1}^n} \det \begin{pmatrix} \left[\left\langle y_i, P y_j' \right\rangle \right]_{i,j=1}^m & 0 \\ 0 & \left[\langle x_i, x_j \rangle \right]_{i,j=1}^n \end{pmatrix} \\
&= \det \left[\left\langle y_i, P y_j' \right\rangle \right]_{i,j=1}^m ,
\end{aligned}$$

while, for any $p \neq m$,

$$\rho(a(y_1) \cdots a(y_m) a(y'_p)^* \cdots a(y'_1)^*) = 0 = \rho(a(y_1) \cdots a(y_m)) ,$$

thanks to the orthogonality of vectors $z_1 \wedge \cdots \wedge z_k \wedge \varphi$ and φ for $k \in \mathbb{N}$ and $z_1, \ldots, z_k \in \ell^2(\Omega \times \Gamma)$. By using properties of Pfaffians in relation with determinants, one deduces that ρ_{HF} is a quasi-free state, which, by Proposition 4.221, is simple with simple symbol given by the orthogonal projector $P \in \mathcal{B}(\ell^2(\Omega \times \Gamma))$ whose range is the orthogonal complement $\{x_1, \ldots, x_n\}^{\perp}$. □

Thus, minimizing free energies on simple quasi-free states represents a natural extension of the Hartree-Fock theory to positive temperatures. In other words, following the Hartree-Fock-Slater's lines of reasoning, as an Ansatz, one can try to minimize the free energy density functional in the set of simple quasi-free states. That is, instead of (5.16), one considers the variational problem

$$\inf\{\mathfrak{e}_{\Phi}(\rho) - \beta^{-1}\mathfrak{s}(\rho) \ : \ \rho \text{ is a simple quasi-free state of } E_1\}$$

for any invariant interaction $\Phi \in \mathcal{W}_1$ and inverse temperature $\beta \in (0, \infty)$.

Recall that simple quasi-free states are (globally) gauge-invariant quasi-free states, by Definition 4.190 and Proposition 4.221. It physically means that they conserve the number of particles. As the celebrated BCS theory of superconductivity demonstrates, we cannot expect this symmetry to be preserved by the equilibrium states of fermion systems, that is, the global gauge invariance symmetry can be spontaneously broken in realistic fermion systems. See Sect. 5.6.

In fact, Araki defines in [69, Definition 3.1] a general notion of quasi-free states which only imposes these states to be even, which is a strictly weaker property than being gauge-invariant. This refers to the "extended quasi-free states" discussed at the end of Sect. 4.8.3. Therefore, as strongly advocated by Bach, Lieb, and Solovej in [77], we shall minimize the free energy density functional in the set of *all* (not necessarily simple) quasi-free states (see Definition 4.217 and discussions after Proposition 4.219), and, instead of (5.16), we should consider the variational problem

$$\inf\{\mathfrak{e}_{\Phi}(\rho) - \beta^{-1}\mathfrak{s}(\rho) \ : \ \rho \text{ is a (general) quasi-free state of } E_1\}$$

for any invariant interaction $\Phi \in \mathcal{W}_1$ and inverse temperature $\beta \in (0, \infty)$. This leads to the *generalized* Hartree-Fock theory, first studied in [77] for the special case of the Hubbard model (Example 5.9). Here, we call the minimizers of the above variational problem the "Hartree-Fock equilibrium states" of the interaction Φ.

The use of quasi-free states to study many-fermion systems has been very popular, even recently. They are used to study not only the minimization of free energies but also many-fermion dynamics. For instance, many rigorous studies [82–

92] from the 2000s on fermion mean-field dynamics use approximating *quasi-free*[6] states (or a mixture of them) as initial states. Quoting [82, p. 79]:

> Slater determinants are relevant at zero temperature because they provide (or at least they are expected to provide) a good approximation to the fermionic ground state of Hamilton operators like (6.1) in the mean-field limit. At positive temperature, equilibrium states are mixed; in the mean-field regime, they are expected to be approximately quasi-free mixed states, like the Gibbs state of a non-interacting gas.

These arguments are probably true in the mean-field regimes considered in these studies because the non-mean-field part of the model always corresponds to a one-particle Hamiltonian. In fact, we unveil in Sect. 6.10 the affinity of the generalized Hartree-Fock theory with mean-field theories.

However, one cannot expect the above-quoted property to hold true for general fermion models. For instance, the Hartree-Fock theory applied to the Hubbard model becomes a good approximation to the true ground state energy only in the limit of *vanishing* coupling constants [79, 81]. Additionally, adding a Hubbard interaction term (Example 5.9) to a quadratic fermionic models directly destroys the quasi-free property of the corresponding equilibrium states, even in the sense of a mixture. In fact, a general many-fermion wave function cannot be represented as a single determinant, even at zero temperature. Consequently, even if the Hartree-Fock method can provide important inputs to our understanding of many-fermion systems with interactions, in particular for the Hubbard model [77], we cannot expect the Hartree-Fock theory to be rigorously correct, in general. For instance, this method usually overestimates the full (ground state) energy. In the same spirit, also other methods are used in computational chemistry and condensed matter physics. They refer to the density-functional theory (e.g., the so-called local-density approximations (LDA), the Kohn-Sham LDA, the generalized gradient approximation). They are not necessarily more accurate, but computationally more efficient, approximations than (extensions of) the Hartree-Fock theory. These approximations are based on effective energy functionals on the electron density only, in line with the Hohenberg-Kohn theorem. See, for instance, [93] and also [78, Section I], which provides a concise review on recent mathematical results related to such methods for Coulomb systems.

[6] In some papers, only (approximating) Slater determinants are considered.

Chapter 6
Equilibrium States of Mean-Field Models and Bogoliubov's Approximation Method

6.1 Topological Framework

Recall that $\mathcal{P}_f$ is the set of all finite subsets of the (cubic) lattice $\Gamma \doteq \mathbb{Z}^d$, for some (space dimension) $d \in \mathbb{N}$. Fix once and for all $N \in \mathbb{N}$ (spin number for quantum spins) or a finite set Ω (the spin set for fermions). These parameters define two different (separable, unital) C^*-algebras, $\mathrm{Spin}(N, \Gamma)$ and $\mathrm{CAR}(\Omega, \Gamma)$, which are always denoted by $\mathcal{U}$, as explained above. In the fermion case, one has additionally to consider the even (CAR) C^*-subalgebra $\mathrm{CAR}(\Omega, \Gamma)^e$, which is denoted by $\mathcal{U}^{\mathrm{e}}$. If one considers the quantum spin case, $\mathcal{U}^{\mathrm{e}}$ is just the original algebra, i.e., $\mathcal{U}^{\mathrm{e}} \doteq \mathcal{U} = \mathrm{Spin}(N, \Gamma)$. Recall that Sect. 5.1 presents the notation in more detail.

For simplicity of notation, as there is no risk of confusion with other objects, the (topological) dual space $\mathcal{U}^{\mathrm{td}}$ of $\mathcal{U}$ is denoted here by $\mathcal{U}^*$, as is usual. The space $\mathcal{U}^*$ is a Banach space when it is endowed with the usual norm for linear functionals on a normed space, that is,

$$\|\rho\|_{\mathrm{op}} \doteq \sup_{A \in \mathcal{U}} \frac{|\rho(A)|}{\|A\|}$$

for all continuous linear functionals $\rho \in \mathcal{U}^*$. However, the norm topology is too strong in practice. The natural topology in the study of infinite systems is given by the $\sigma(\mathcal{U}^*, \mathcal{U})$-topology, usually called the weak* topology of $\mathcal{U}^*$. It is the initial topology of the family of linear mappings $\rho \mapsto \rho(A)$ from $\mathcal{U}^*$ to $\mathbb{C}$ for all algebra elements $A \in \mathcal{U}$. It is, by definition, the coarsest topology on $\mathcal{U}^*$ that makes the mapping $\rho \mapsto \rho(A)$ continuous for every $A \in \mathcal{U}$. See [18, Section 3.8]. The topology of the dual space $\mathcal{U}^*$ is, by default, the weak* topology. In this case, $\mathcal{U}^*$ is a (Hausdorff) locally convex space, and its (topological) dual space is $\mathcal{U}$: Any element of $\mathcal{U}^{**} \equiv \mathcal{U}$ is of the form $\rho \mapsto \rho(A)$ for some algebra element $A \in \mathcal{U}$. See, e.g., [18, Theorem 3.10]. In fact, recall that in Sect. 4.5.1, we define the weak*

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_6

topology for *states* of any separable C^*-algebra (like $\mathcal{U}$) in a more concrete way, via an explicit metric. See Definition 4.80 and Exercise 4.82. It is important to notice that this metric does not reproduce the weak* topology in the whole space $\mathcal{U}^*$, but only in its norm-bounded subsets, like any set of states of $\mathcal{U}$.

The convex subset of invariant states on $\mathcal{U}$ is denoted by $E_1 \subseteq \mathcal{U}^*$. See again Sect. 5.1 for more details. One easily verifies that E_1 is a weak*-closed set. In addition, recall that any continuous linear functional $\rho \in \mathcal{U}^*$ is a state iff $\rho(\mathfrak{1}) = 1$ and $\|\rho\|_{\mathrm{op}} = 1$, $\mathfrak{1} \in \mathcal{U}$ being the unit of $\mathcal{U}$. Hence, from the Banach-Alaoglu theorem [18, Theorem 3.15] and the closedness of E_1, the set E_1 of invariant states is a weak*-compact subset of the unit ball of $\mathcal{U}^*$. See Proposition 4.84 for a direct proof of compactness of the set of states of separable unital algebras, keeping in mind that the (spin or fermion) algebra $\mathcal{U}$ is of this type.

Proposition 7.334 tells us then that the convex weak*-compact space E_1 of invariant states is the weak* closure of the convex hull of the (nonempty) set $\mathcal{E}_1$ of its extreme points:

$$E_1 = \overline{\mathrm{co}\mathcal{E}_1} \ .$$

The set $\mathcal{E}_1 \subseteq E_1 \subseteq \mathcal{U}^*$ also refers in the literature to the extreme boundary of E_1. Here, recall that extreme points of the convex set E_1 are called here *ergodic*, because of the formal analogy to the classical case.

As discussed above, since the (spin or fermion) algebra $\mathcal{U}$ is separable, the weak* topology is metrizable on any weak*-compact subset of $\mathcal{U}^*$. See, e.g., Proposition 4.84 or [18, Theorem 3.16]. In particular, the space E_1 is metrizable, in this case. This is an important property, which strongly simplifies the study of E_1, in particular because it allows for Choquet decompositions of invariant states as barycenters of ergodic ones.

Nonetheless, in spite of the metrizability of the weak* topology in E_1, the space E_1 of all invariant states has still a *fairly complicated* geometrical structure: In 1961, E. T. Poulsen [16] constructed an example of a metrizable simplex with dense set of extreme points. This simplex is now known as *the* Poulsen simplex because it is unique [17, Theorem 2.3], up to an affine homeomorphism. One can show that the set E_1 of invariant states is also a simplex and, in fact, the Poulsen simplex. In particular, the following assertion holds true:

Theorem 6.1 (Density of Ergodic States) *The set $\mathcal{E}_1$ of ergodic states is a weak*-dense subset of the set E_1 of all invariant states.*

Proof Recall that, for all $n \in \mathbb{N}$,

$$\Lambda_n \doteq \{(x_1, \ldots, x_d) \in \Gamma : |x_i| \leq n\} \in \mathcal{P}_f \ .$$

For any invariant state $\rho \in E_1$ and $n \in \mathbb{N}$, let $\tilde{\rho}_n$ be the product state defined by

$$\tilde{\rho}_n = \bigotimes_{x \in \mathbb{Z}^d} \rho|_{\mathcal{U}_{\Lambda_n + (2n+1)x}} \ .$$

It is a periodic state, whose period is $(2n+1, \ldots, 2n+1) \in \mathbb{Z}^d$, and satisfies

$$\tilde{\rho}_n(A) \doteq \rho(A) \ , \qquad A \in \mathcal{U}_{\Lambda_n} \ .$$

Here, $\mathcal{U}_\Lambda \doteq \mathrm{Spin}(N, \Lambda)$ (quantum spin case) or $\mathcal{U}_\Lambda \doteq \mathrm{CAR}(\Omega, \Lambda)$ (fermion case) for any $\Lambda \in \mathcal{P}_f$ (see Sect. 5.1). Note that this construction is possible also in the fermion case because any invariant state $\rho \in E_1$ is even, thanks to Theorem 5.3. See Proposition 4.193. Then,

$$\hat{\rho}_n \doteq \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \tilde{\rho}_n \circ \tau_x \tag{6.1}$$

is a well-defined invariant state, where $\tau_x : \mathcal{U} \to \mathcal{U}$, $x \in \Gamma$, are the translation automorphisms, that is, the unique unital $*$-homomorphisms defined by (5.4), in the quantum spin case, and (5.9), in the fermion case. Fix $A \in \mathcal{U}_{\mathrm{loc}}$, where we recall that $\mathcal{U}_{\mathrm{loc}} \subseteq \mathcal{U}$ is the $*$-algebra of local elements, defined as the (countable) union of $\mathcal{U}_\Lambda$ for all $\Lambda \in \mathcal{P}_f$ (see again Sect. 5.1). Then,

$$\lim_{n \to \infty} \hat{\rho}_n(A) = \rho(A)$$

and so, $\hat{\rho}_n$ converges in the weak* topology to the invariant state $\rho \in E_1$, by density of the $*$-algebra $\mathcal{U}_{\mathrm{loc}} \subseteq \mathcal{U}$ of local elements. Moreover, $\tilde{\rho}_n$ being a product state, there is a constant $C > 0$ (depending on $\Lambda \in \mathcal{P}_f$) such that

$$\tilde{\rho}_n\left(\tau_x(A^*)\tau_y(A)\right) = \tilde{\rho}_n\left(\tau_x(A^*)\right)\tilde{\rho}_n\left(\tau_y(A)\right) \ ,$$

whenever $|x - y| \geq C$. Then, using the notation $|B|^2 = B^* B$,

$$\begin{aligned} \tilde{\rho}_n(|A_\ell|^2) &= \frac{1}{|\Lambda_\ell|^2} \sum_{x \in \Lambda_\ell} \sum_{y \in \Lambda_\ell} \tilde{\rho}_n\left(\tau_x(A)\tau_y(A)\right) \\ &= \frac{1}{|\Lambda_\ell|^2} \sum_{x \in \Lambda_\ell} \sum_{y \in \Lambda_\ell : |x-y| \geq C} \tilde{\rho}_n\left(\tau_x(A)\tau_y(A)\right) \\ &\quad + \underbrace{\frac{1}{|\Lambda_\ell|^2} \sum_{x \in \Lambda_\ell} \sum_{y \in \Lambda_\ell : |x-y| < C} \tilde{\rho}_n\left(\tau_x(A)\tau_y(A)\right)}_{\mathcal{O}(\ell^{-d})} \ . \end{aligned}$$

Thus,

$$\tilde{\rho}_n(|A_\ell|^2) = \frac{1}{|\Lambda_\ell|^2} \sum_{x,y \in \Lambda_\ell} \tilde{\rho}_n\left(\tau_x(A^*)\right)\tilde{\rho}_n\left(\tau_y(A)\right) + \mathcal{O}(\ell^{-d}) \ . \tag{6.2}$$

Since $\hat{\rho}_n$ is an invariant state, for any $\Lambda \in \mathcal{P}_f$ and $A \in \mathcal{U}_\Lambda$,

$$\frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} \tilde{\rho}_n \circ \tau_x(A) = \underbrace{\frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} \hat{\rho}_n \circ \tau_x(A)}_{=\hat{\rho}_n(A)} + \mathcal{O}(\ell^{-1}) ,$$

which, combined with (6.2), yields

$$\lim_{\ell \to \infty} \tilde{\rho}_n(|A_\ell|^2) = \left|\hat{\rho}_n (A)\right|^2 .$$

Using this last equality and Equation (6.1),

$$\begin{aligned} \lim_{\ell \to \infty} \hat{\rho}_n(|A_\ell|^2) &= \lim_{\ell \to \infty} \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \tilde{\rho}_n \circ \tau_x(|A_\ell|^2) \\ &= \lim_{\ell \to \infty} \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \tilde{\rho}_n \left(\left|(\tau_x (A))_\ell\right|^2\right) \\ &= \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \left|\hat{\rho}_n \circ \tau_x (A)\right|^2 = \left|\hat{\rho}_n (A)\right|^2 . \end{aligned} \tag{6.3}$$

By density of the $*$-algebra $\mathcal{U}_{\text{loc}} \subseteq \mathcal{U}$ of local elements, (6.3) holds true for any (spin or fermion) algebra element $A \in \mathcal{U}$, i.e., $\hat{\rho}_n$ is dispersionless at infinity (Definition 5.35). By Theorem 5.37, $\hat{\rho}_n \in \mathcal{E}_1$ is therefore ergodic for each $n \in \mathbb{N}$.

□

In fact, it turns out that also the *full* set of states of the unital C^*-algebra $\mathcal{U}$ associated with an infinitely extended (quantum spin or fermion) system has the property proven above for that set of *invariant* states: $\mathcal{U}$ is a so-called approximately finite-dimensional (AF) C^*-algebra, i.e., it is generated by an increasing family of *finite-dimensional* C^*-subalgebras. In this case, by [22, Lemma 11.2.4], the set $\mathcal{E}$ of extreme points of the set E of all states of $\mathcal{U}$ is weak*-dense in E, i.e.,

$$E = \overline{\text{co}\mathcal{E}} = \overline{\mathcal{E}} . \tag{6.4}$$

For more details, we recommend [21, Section 8]. Note that, astonishingly, (6.4) do not prevent E from having a unique center [24] (i.e., a sort of maximally mixed point).

The property of having a dense extreme boundary should however not be so surprising for mathematicians. The existence of such convex sets is well-known in infinite-dimensional vector spaces. For instance, the unit ball of any infinite-dimensional Hilbert space has a dense extreme boundary in the weak topology. It turns out that this situation is not accidental, but *generic* for weak*-compact convex sets in infinite dimension. See [21, Section 2.3] for more details, which has been extended in [23] for the dual space $\mathcal{X}^*$, endowed with its weak* topology, of any infinite-dimensional, separable topological vector space $\mathcal{X}$.

In the sequel, we will show that such a property is not just a mathematical curiosity, but has important consequences in terms of thermodynamic properties of infinitely extended (quantum spin or fermion) systems.

6.2 Spin and Fermion Mean-Field Models

In Definition 5.5, we introduce spin and fermion interactions on the cubic lattice $\Gamma \doteq \mathbb{Z}^d$ ($d \in \mathbb{N}$). Here, it is convenient to remove from this definition the self-conjugate property of interactions and use the vector space

$$\mathcal{V}^{\mathbb{C}} \doteq \left\{\Phi + i\Phi' : \Phi, \Phi' \text{ interactions in the sense of Definition 5.5}\right\}$$

of *complex* interactions, where, for all $\Psi, \Psi' \in \mathcal{V}^{\mathbb{C}}$ and $\alpha \in \mathbb{C}$, $\Psi + \Psi' \in \mathcal{V}^{\mathbb{C}}$ and $\alpha\Psi \in \mathcal{V}^{\mathbb{C}}$ are, respectively, defined by

$$(\Psi + \Psi')(\Lambda) \doteq \Psi(\Lambda) + \Psi'(\Lambda) , \qquad (\alpha\Psi)(\Lambda) \doteq \alpha(\Psi(\Lambda)) , \qquad \Lambda \in \mathcal{P}_f .$$

Cf. Eq. (5.11). Invariant (with respect to space translations) complex interactions are defined exactly as in the real case. See Definition 5.5 (iii). In Definition 5.6, we introduce a real Banach space $\mathcal{W}_1$ of invariant (spin or fermion) interactions which is now embedded in a complex Banach space of (complex, invariant, spin, or fermion) interactions:

Definition 6.2 (A Banach Space of Invariant Complex Interactions) The Banach space of (short-range) invariant complex interactions is defined by

$$\mathcal{W}_1^{\mathbb{C}} \doteq \{\Phi \in \mathcal{V}^{\mathbb{C}} : \Phi \text{ is an invariant interaction for which } \|\Phi\| < \infty\} ,$$

where the norm of $\mathcal{W}_1^{\mathbb{C}}$ is defined like in Definition 5.6, that is,

$$\|\Phi\| \doteq \sum_{\Lambda\in\mathcal{P}_f,\, 0\in\Lambda} \frac{1}{|\Lambda|} \|\Phi(\Lambda)\| \in \mathbb{R}_0^+ \cup \{\infty\} , \qquad \Phi \in \mathcal{V}^{\mathbb{C}} .$$

This space serves to define a much more general Banach space of mean-field models:

Definition 6.3 (A Banach Space of Mean-Field Models) The space of mean-field models is the *real* Banach space $\mathcal{M}_1 \doteq \mathcal{W}_1 \times \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})^2$, where

$$\ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}}) \doteq \left\{\Psi \equiv (\Psi_n)_{n\in\mathbb{N}} \subseteq \mathcal{W}_1^{\mathbb{C}} : \|\Psi\|_2^2 \doteq \sum_{n\in\mathbb{N}} \|\Psi_n\|^2 < \infty\right\} ,$$

whose norm is defined by

$$\|\mathfrak{m}\| \doteq \|\Phi\| + \|\Psi_-\|_2 + \|\Psi_+\|_2 , \qquad \mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1 .$$

Here, Ψ_- represents the mean-field attraction of the model, while Ψ_+ refers to its mean-field repulsion.

Note that $\mathcal{W}_1 \subseteq \mathcal{M}_1$, using the identification $\Phi \equiv (\Phi, 0, 0)$ for $\Phi \in \mathcal{W}_1$.

Similar to Definition 5.10, local energy observables, or Hamiltonians, are defined for all complex interactions as follows: For all $\Phi \in \mathcal{V}^{\mathbb{C}}$ and $\Lambda \in \mathcal{P}_f$,

$$H_\Lambda^\Phi \doteq \sum_{\Lambda' \in \mathcal{P}_f,\ \Lambda' \subseteq \Lambda} \Phi(\Lambda') \in \mathcal{U}^e .$$

These complex local Hamiltonians are then used to define local Hamiltonians for any mean-field model in $\mathcal{M}_1$:

Definition 6.4 (Local Energy Observables) For any $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and finite subset $\Lambda \in \mathcal{P}_f$,

$$H_\Lambda^{\mathfrak{m}} \doteq H_\Lambda^\Phi + \frac{1}{|\Lambda|} \sum_{n \in \mathbb{N}} \left(|H_\Lambda^{\Psi_{+,n}}|^2 - |H_\Lambda^{\Psi_{-,n}}|^2 \right) \in \mathrm{Re}\{\mathcal{U}_\Lambda^e\} ,$$

where, as is usual, $|A|^2 \doteq A^* A$. The self-conjugate element $H_\Lambda^{\mathfrak{m}} = (H_\Lambda^{\mathfrak{m}})^*$ is the (local) "Hamiltonian associated with the (finite) region Λ and the mean-field model $\mathfrak{m}$."

Note that the identification $\Phi \equiv (\Phi, 0, 0)$ for $\Phi \in \mathcal{W}_1$ is coherent with Definitions 5.10 and 6.4, since $H_\Lambda^{(\Phi,0,0)} = H_\Lambda^\Phi$ for any $\Lambda \in \mathcal{P}_f$.

By Definition 6.4, the Hamiltonian associated with a mean-field model (Φ, Ψ_-, Ψ_+) has a mean-field attraction term, and a repulsion one, respectively, defined from the components Ψ_- and Ψ_+ of $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$. The mean-field model $\mathfrak{m}$ is said to be "purely attractive" iff $\Psi_+ = 0$, while it is "purely repulsive" iff $\Psi_- = 0$. Distinguishing between these two special types of models is important because the effects of mean-field attractions and repulsions on the structure of the corresponding sets of (globally stable) equilibrium states can be very different. For instance, by [1, Theorem 2.25], mean-field attractions have no particular effect on the structure of the set of (generalized) equilibrium states. By contrast, mean-field repulsions have a geometrical effect, by possibly preventing the set of equilibrium states of being a face of the set of all invariant states. See [1, Lemma 9.8].

Exercise 6.5 Show that, for $\Lambda \in \mathcal{P}_f$ and $\mathfrak{m} \in \mathcal{M}_1$,

$$\left\| H_\Lambda^{\mathfrak{m}} \right\| \leq |\Lambda| \, \|\mathfrak{m}\| . \tag{6.5}$$

Example 6.6 Like in Example 5.9, let $\eta \in \mathbb{R}^+$, $\Omega \doteq \{\uparrow, \downarrow\}$, $\mathcal{U} \doteq \mathrm{CAR}(\{\uparrow, \downarrow\}, \Gamma)$, and take the (canonical) Hilbert basis $\{\mathrm{e}_{\mathrm{s},x}\}_{(\mathrm{s},x)\in\{\uparrow,\downarrow\}\times\Gamma}$ of $\ell^2(\{\uparrow, \downarrow\} \times \Gamma)$. The "BCS interaction" $\Psi_{\mathrm{BCS}} \in \mathcal{W}_1^{\mathbb{C}}$ is defined by $\Psi_{\mathrm{BCS}}(\Lambda) \doteq 0$ whenever $|\Lambda| \notin \{1\}$ and $\Psi_{\mathrm{BCS}}(\{x\}) \doteq \eta^{1/2} a\left(\mathrm{e}_{x,\downarrow}\right) a\left(\mathrm{e}_{x,\uparrow}\right)$ for every $x \in \Gamma$. Then, the (reduced) BCS model of superconductivity refers to the purely attractive mean-field model $\mathfrak{n} = (\Phi, (\Psi_{\mathrm{BCS}}, 0, 0, \ldots), 0)$, where $\Phi = \Phi_{\mathrm{Hubb}}$ for $U = 0$ (see Example 5.9). In this case, we get as local Hamiltonians the usual (reduced) BCS Hamiltonians:

$$H_\Lambda^{\mathfrak{n}} \doteq -t \sum_{s\in\{\uparrow,\downarrow\}} \sum_{x,y\in\Lambda,|x-y|=1} a\left(\mathrm{e}_{x,s}\right)^* a\left(\mathrm{e}_{y,s}\right) - \mu \sum_{s\in\{\uparrow,\downarrow\}} \sum_{x\in\Lambda,} a\left(\mathrm{e}_{x,s}\right)^* a\left(\mathrm{e}_{x,s}\right)$$
$$-\frac{\eta}{|\Lambda|} \sum_{x,y\in\Lambda} a\left(\mathrm{e}_{x,\uparrow}\right)^* a\left(\mathrm{e}_{x,\downarrow}\right)^* a\left(\mathrm{e}_{y,\downarrow}\right) a\left(\mathrm{e}_{y,\uparrow}\right) .$$

Here, $\eta \geq 0$ is the "BCS interaction strength." If we take $\Phi = \Phi_{\mathrm{Hubb}}$ for $U \neq 0$, then we obtain the so-called BCS-Hubbard model.

See Sect. 6.6 for more details. Another, more general, example is given in Sect. 6.9.

6.3 Free Energy Density of Mean-Field Models

Recall that the entropy density functional $\mathfrak{s} : E_1 \to \mathbb{R}_0^+$ is the thermodynamic limit of the von Neumann entropy per unit volume:

$$\mathfrak{s}(\rho) \doteq \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|} S_\ell(\rho) .$$

See Theorem 5.20, which states that this functional is affine[1] and bounded on the convex weak*-compact space E_1 of all invariant states. We show next its continuity properties with respect to the weak* topology:

Lemma 6.7 (Ergodic Abundance) *The entropy density functional* $\mathfrak{s} : E_1 \to \mathbb{R}_0^+$ *is affine and weak*-upper semicontinuous. Additionally, for any invariant state* $\rho \in E_1$*, there is a sequence* $(\hat{\rho}_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}_1$ *of ergodic states converging to* ρ *and such that*

$$\mathfrak{s}(\rho) = \lim_{n\to\infty} \mathfrak{s}(\hat{\rho}_n) .$$

[1] Recall that a function h on a convex set K is affine iff $h(\lambda x + (1-\lambda) y) = \lambda h(x) + (1-\lambda)h(y)$ for all $x, y \in K$.

Proof Theorem 5.20 tells us that, for any invariant state $\rho \in E_1$,

$$\mathfrak{s}(\rho) = \inf \left\{ \frac{1}{|\Lambda_\ell|} S_\ell(\rho) \ : \ \ell \in \mathbb{N} \right\} .$$

In other words, $\mathfrak{s}$ is given by the infimum of weak*-continuous functionals $S_\ell : E_1 \to \mathbb{R}_0^+$. It is therefore weak*-upper semicontinuous, by Lemma 7.144. Now, it is shown in the proof of Theorem 6.1 that the states

$$\hat{\rho}_n \doteq \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \tilde{\rho}_n \circ \tau_x$$

for $n \in \mathbb{N}$ are not only invariant but also ergodic and moreover, as $n \to \infty$, they converge to ρ in the weak* topology. Recall that $\tilde{\rho}_n$ is a periodic (product) state, whose period is $(2n+1, \ldots, 2n+1) \in \mathbb{Z}^d$, for which

$$\tilde{\rho}_n (A) \doteq \rho (A)$$

for any $A \in \mathcal{U}_{\Lambda_n}$. If $\mathfrak{s}$ can be defined for invariant states, thanks to Theorem 5.20, then it can also be defined for periodic states by redefining the parameter $N \in \mathbb{N}$ (defining the algebra $\mathrm{Spin}(N, \Gamma)$), in the quantum spin case, or the spin set Ω (defining the algebra $\mathrm{CAR}(\Omega, \Gamma)$), in the fermion case, in order to see any periodic state as an invariant state. In particular, $\mathfrak{s}$ can be defined as an affine functional on periodic states, and in this case, for any fixed $n \in \mathbb{N}$,

$$\mathfrak{s}(\hat{\rho}_n) = \frac{1}{|\Lambda_n|} \sum_{x \in \Lambda_n} \mathfrak{s}\left(\tilde{\rho}_n \circ \tau_x\right) = s(\tilde{\rho}_n) = \frac{1}{|\Lambda_n|} S_n(\rho) .$$

The above sequence $(\hat{\rho}_n)_{n \in \mathbb{N}} \subseteq \mathcal{E}_1$ of ergodic states thus satisfies all the desired properties. For more details, see [1, Lemma 1.29]. □

Notice that the convergence of the entropy density along sequences of pure invariant states, referring to the second part of the lemma, has a classical analogue called "ergodic abundance" [11, Section 2.1]. Important applications of this property have been recently found (see [11] and references therein) to the so-called "nonlinear thermodynamic formalism" of classical dynamical systems.

Recall that the energy density observable associated with an invariant interaction $\Phi \in \mathcal{W}_1$ refers to Definition 5.10 (ii). Extended to all complex interactions, it corresponds to

$$e_\Phi \doteq \sum_{\Lambda \in \mathcal{P}_f,\ 0 \in \Lambda} \frac{1}{|\Lambda|} \Phi(\Lambda) \in \mathcal{U}^e \tag{6.6}$$

for any $\Phi \in \mathcal{W}_1^{\mathbb{C}}$. From Proposition 5.11, it defines an energy density functional $\mathfrak{e}_\Phi : E_1 \to \mathbb{R}$ for any interaction $\Phi \in \mathcal{W}_1$. See Definition 5.12. This definition is also extended to all complex interactions: For any invariant state $\rho \in E_1$ and $\Phi \in \mathcal{W}_1^{\mathbb{C}}$,

$$\mathfrak{e}_\Phi(\rho) \doteq \rho(e_\Phi) . \tag{6.7}$$

It is clearly an affine functional on the convex weak*-compact space E_1 of all invariant states. Its main basic properties are gathered in the following lemma:

Lemma 6.8 *For any complex interaction $\Phi \in \mathcal{W}_1^{\mathbb{C}}$, the energy density functional $\mathfrak{e}_\Phi : E_1 \to \mathbb{R}$ is affine and weak*-continuous. Moreover, for any $\Phi, \Phi' \in \mathcal{W}_1^{\mathbb{C}}$ and invariant state $\rho \in E_1$,*

$$|\mathfrak{e}_\Phi(\rho) - \mathfrak{e}_{\Phi'}(\rho)| \leq \|\Phi - \Phi'\|.$$

Proof The properties directly follow from Eq. (6.7). Note that the last inequality is already mentioned after Definition 5.12. Its proof results from direct computations using the bound

$$|\mathfrak{e}_\Phi(\rho) - \mathfrak{e}_{\Phi'}(\rho)| = |\mathfrak{e}_{\Phi-\Phi'}(\rho)| \leq \|e_{\Phi-\Phi'}\|$$

and the explicit expression for the algebra element e_Φ, as well as the definition of the norm of interactions given in Definition 6.2. □

In addition to the energy and entropy density functionals, we need the so-called space-averaging functionals, in order to study the thermodynamic properties of mean-field models. This new functionals are defined on the convex weak*-compact space E_1 of all invariant states as follows: Recall that, for any (spin or fermion) algebra element $A \in \mathcal{U}$,

$$A_\ell \doteq \frac{1}{|\Lambda_\ell|} \sum_{x \in \Lambda_\ell} \tau_x(A) , \tag{6.8}$$

where the unital $*$-homomorphisms $\tau_x : \mathcal{U} \to \mathcal{U}$, $x \in \Gamma$, are the above-defined translation automorphisms (see Eqs. (5.4), for the quantum spin case, or (5.9), for the fermion case), while, for any natural number $\ell \in \mathbb{N}$, $\Lambda_\ell \in \mathcal{P}_f$ is defined by (5.2).

Then, we use Corollary 5.34 to define a the space-averaging functionals on invariant states:

Definition 6.9 (Space-Averaging Functional) Fix a fixed (spin or fermion) algebra element $A \in \mathcal{U}$. Then, the "space-averaging functional" associated with this algebra element is the mapping Δ_A from the space E_1 of invariant states to $\mathbb{R}$ defined by

$$\rho \mapsto \Delta_A(\rho) \doteq \lim_{\ell \to \infty} \rho\left(A_\ell^* A_\ell\right) \in \left[|\rho(A)|^2, \|A\|^2\right] .$$

Observe from Definitions 5.35 and 6.9 and Theorem 5.37 that an invariant state is ergodic iff it is dispersionless at infinity, i.e., $\rho \in \mathcal{E}_1$ iff

$$\Delta_A(\rho) = |\rho(A)|^2 \doteq \delta_A(\rho) , \qquad A \in \mathcal{U} . \tag{6.9}$$

The space-averaging functional is therefore explicitly given on the (dense) set $\mathcal{E}_1$ of ergodic states.

Lemma 6.10 *The space-averaging functional has the following properties:*

(i) *At fixed (spin or fermion) algebra element $A \in \mathcal{U}$, Δ_A is weak*-upper semicontinuous and affine.*
(ii) *At fixed invariant state $\rho \in E_1$ and for all algebra elements $A, B \in \mathcal{U}$,*

$$|\Delta_A(\rho) - \Delta_B(\rho)| \leq (\|A\| + \|B\|)\|A - B\| .$$

Proof Except for the upper semicontinuity property, all the assertions directly follow from the definition. The upper semicontinuity of Δ_A, $A \in \mathcal{U}$, follows by combining Lemma 7.144 with the fact that Δ_A is the infimum over a family of continuous functionals:

$$\Delta_A(\rho) = \inf_{\ell \in \mathbb{N}} \left\{ \rho(|A_\ell|^2) \right\} .$$

This property is proven by using the von Neumann ergodic theorem and the GNS representation of states (Theorem 4.113). See proof of Corollary 5.34 or [1, Section 1.3] for more details. □

Note that the space-averaging functionals cannot be generally weak*-continuous. This is a consequence of the density of the set $\mathcal{E}_1 \subseteq E_1$ of ergodic states: By Theorem 6.1, if Δ_A is weak*-continuous, then (6.9) holds true for all invariant states, i.e., $\Delta_A = \delta_A$. Therefore, it must exist an algebra element $A \in \mathcal{U}$ such that Δ_A is not weak*-continuous; otherwise, all invariant states would be ergodic, thanks to Theorem 5.37. In fact, we have the following general statement concerning the continuity of Δ_A:

Theorem 6.11 *Fix a (spin or fermion) algebra element $A \in \mathcal{U}$ and let δ_A be the weak*-continuous convex function defined by $\delta_A(\rho) \doteq |\rho(A)|^2$ on the convex weak*-compact convex space E_1 of invariant states. Then, one has:*

(i) *Δ_A is weak*-continuous iff δ_A is a constant function.*
(ii) *Δ_A is weak*-discontinuous on a weak*-dense subset of invariant states, unless δ_A is a constant function.*
(iii) *Δ_A is weak*-continuous on the dense subset $\mathcal{E}_1$ of ergodic states.*
(iv) *For all invariant states, $\rho \in E_1$, $\Delta_A(\rho) = \mu_\rho(\delta_A)$ with μ_ρ being the positive linear functional of Theorem 7.339, on weak*-continuous complex-valued functions on E_1. (See also Theorem 4.68 and related remarks.)*

(v) *We have* $\gamma(\Delta_A) = \delta_A$, *where* $\gamma(\Delta_A)$ *is the so-called* γ*-regularization of* Δ_A *on* E_1, *defined by*

$$\gamma(\Delta_A)(\rho) \doteq \sup\{\rho(B) : B \in \mathrm{Re}\{\mathcal{U}\} \text{ such that } \forall \varpi \in E_1,\ \varpi(B) \leq \Delta_A(\varpi)\} \ .$$

See Definition 7.340 and Proposition 7.347.

Proof (i)–(iii) result partially from Theorems 7.339 and 5.37. (iv) follows from Lemma 6.10 (i) combined with Theorems 7.339 and 5.37: By affineness and upper semicontinuity of Δ_A (see Lemma 6.10 (i) and [1, Lemma 10.17]) as well as from Theorems 5.37 and 7.339,

$$\Delta_A(\rho) = \mu_\rho(\Delta_A) = \mu_\rho(\delta_A) \ .$$

It remains to prove (v): By Corollary 7.342, the γ-regularization $\gamma(\Delta_A)$ on E_1 is the largest weak*-lower semicontinuous and convex minorant of Δ_A on E_1. Since

$$\Delta_A(\rho) \doteq \lim_{\ell\to\infty} \rho\left(A_\ell^* A_\ell\right) \in \left[|\rho(A)|^2, \|A\|^2\right]$$

for any invariant state $\rho \in E_1$, the function δ_A is a weak*-continuous convex minorant of Δ_A on E_1. Therefore, for any invariant state, $\rho \in E_1$, $\delta_A \leq \gamma(\Delta_A) \leq \Delta_A$, and it follows that $\gamma(\Delta_A)(\rho) = \delta_A(\rho)$ for any ergodic state $\rho \in \mathcal{E}_1$. By weak*-density of $\mathcal{E}_1 \subseteq E_1$ (Theorem 6.1), (v) follows. □

We are now in a position to define the free energy density associated with mean-field models at fixed (non-zero) temperatures:

Definition 6.12 (Free Energy Density) For any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$, the "free energy density functional" $\mathfrak{f}_{\mathfrak{m},\beta} : E_1 \to \mathbb{R}$ on the space E_1 of all invariant states is defined by

$$\mathfrak{f}_{\mathfrak{m},\beta} \doteq \Delta_{\Psi_+} - \Delta_{\Psi_-} + \mathfrak{e}_\Phi - \beta^{-1}\mathfrak{s} \doteq \Delta_{\Psi_+} - \Delta_{\Psi_-} + \mathfrak{f}_{\Phi,\beta}$$

(see Definitions 5.22 and 6.9, Theorem 5.20, and Eq. (6.7)), where, for any sequence $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$ of complex interactions,

$$\Delta_\Psi \doteq \sum_{n\in\mathbb{N}} \Delta_{\mathfrak{e}_{\Psi_n}} \ .$$

The free energy density is clearly the same as the one of Definition 5.22 for any $\Phi \in \mathcal{W}_1 \subseteq \mathcal{M}_1$ and $\beta \in (0, \infty)$. Note that Δ_Ψ is well-defined, because, for any sequence $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$ of complex (invariant) interactions,

$$\sum_{n\in\mathbb{N}} \sup_{\rho\in E_1} \left|\Delta_{\mathfrak{e}_{\Psi_n}}(\rho)\right| \leq \sum_{n\in\mathbb{N}} \|\mathfrak{e}_{\Psi_n}\|^2 \leq \sum_{n\in\mathbb{N}} \|\Psi_n\|^2 < \infty , \tag{6.10}$$

thanks to Lemma 6.8 and Definition 6.9.

The (previous) free energy density functional $\mathfrak{f}_{\Phi,\beta}$ of Definition 6.12 looks natural, as the energy and entropy per unit volume associated with the invariant interaction Φ in a given invariant state ρ. Nevertheless, the mean-field terms in the new free energy density functional $\mathfrak{f}_{\mathfrak{m},\beta}$, defined above by means of the functionals space-averaging $\Delta_{\Psi_\pm}$, may look more intriguing. To explain the origin of these new terms, we come back to finite-volume systems:

Recall that the Gibbs states of Definition 5.19, i.e., equilibrium states at finite volume, are minimizers of the finite-volume free energy, which leads to the concept of the pressure. See Proposition 3.13 and Definition 3.16. In particular, by considering the local Hamiltonians $H_\Lambda^{\mathfrak{m}}$ of Definition 6.4, given a fixed mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$, we can define, for any finite (nonempty) region $\Lambda \in \mathcal{P}_f$, the pressure

$$P_{H_\Lambda^{\mathfrak{m}},\beta} \doteq -\frac{1}{|\Lambda|} \inf \left\{ F_{H_\Lambda^{\mathfrak{m}},\beta}(\rho) : \rho \in E(\mathcal{U}_\Lambda) \right\} , \tag{6.11}$$

where the free energy functional $F_{H_\Lambda^{\mathfrak{m}},\beta}$ is the one of Definition 5.19, for $H = H_\Lambda^{\mathfrak{m}}$. Then, by taking, for instance, the sequence (5.2) of cubic boxes in Γ, one may ask about the limit $\ell \to \infty$ of the sequence $(P_{H_{\Lambda_\ell}^{\mathfrak{m}},\beta})_{\ell\in\mathbb{N}}$, as well as the corresponding Gibbs states. Such a limit is known in statistical mechanics as the "thermodynamic limit." Answering such a question naturally yields the free energy density functional of Definition 6.12:

Theorem 6.13 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$,

$$\mathfrak{p}_\beta(\mathfrak{m}) \doteq -\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = \lim_{\ell\to\infty} P_{H_{\Lambda_\ell}^{\mathfrak{m}},\beta} \in \mathbb{R} .$$

Idea of Proof Any state $\rho \in E(\mathcal{U})$ on $\mathcal{U}$ can be seen, by restriction, as a state $\rho|_{\mathcal{U}_{\Lambda_\ell}} \in E(\mathcal{U}_{\Lambda_\ell})$ on $\mathcal{U}_{\Lambda_\ell} \subseteq \mathcal{U}$ for any $\ell \in \mathbb{N}$. Using Definition 6.4 and Proposition 3.13, we thus deduce that, for any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$, $\beta \in (0, \infty)$ and all states $\rho \in E(\mathcal{U})$,

$$\begin{aligned} P_{H_{\Lambda_\ell}^{\mathfrak{m}},\beta} \geq{} & \frac{1}{|\Lambda_\ell|^2} \sum_{n\in\mathbb{N}} \rho(|H_{\Lambda_\ell}^{\Psi_{-,n}}|^2) - \frac{1}{|\Lambda_\ell|^2} \sum_{n\in\mathbb{N}} \rho(|H_{\Lambda_\ell}^{\Psi_{+,n}}|^2) - \frac{1}{|\Lambda_\ell|} \rho\left(H_{\Lambda_\ell}^{\Phi}\right) \\ & + \frac{1}{\beta|\Lambda_\ell|} S(\rho|_{\mathcal{U}_{\Lambda_\ell}}) \end{aligned}$$

with equality when $\rho|_{\mathcal{U}_{\Lambda_\ell}}$ is the Gibbs states of Definition 5.19 for $H = H_{\Lambda_\ell}^{\mathfrak{m}}$. When the state ρ is invariant, i.e., $\rho \in E_1$,

$$\lim_{\ell\to\infty} \left\{ \frac{1}{|\Lambda_\ell|} \rho\left(H_{\Lambda_\ell}^{\Phi}\right) - \frac{1}{\beta|\Lambda_\ell|} S(\rho|_{\mathcal{U}_{\Lambda_\ell}}) \right\} = \mathfrak{e}_\Phi(\rho) - \beta^{-1}\mathfrak{s}(\rho) \doteq \mathfrak{f}_{\Phi,\beta}(\rho) \ .$$

See Theorem 5.20 and Definition 5.12. Moreover, for any invariant state $\rho \in E_1$ and any complex (invariant) interaction $\Psi \in \mathcal{W}_1^{\mathbb{C}}$, one checks from direct estimates that

$$\lim_{\ell\to\infty} \left(\frac{1}{|\Lambda_\ell|^2} \rho(|H_{\Lambda_\ell}^{\Psi}|^2) - \rho(\left|(e_\Psi)_\ell\right|^2) \right) = 0$$

with $(e_\Psi)_\ell \in \mathcal{U}$ being given by

$$(e_\Psi)_\ell \doteq \frac{1}{|\Lambda_\ell|} \sum_{x\in\Lambda_\ell} \tau_x(e_\Psi) \ ,$$

the algebra element $e_\Psi \in \mathcal{U}$ being the energy density observable (6.6). See also Eq. (6.8). By Definition 6.9 of $\Delta_{\mathfrak{e}_\Psi}$, it follows that

$$\lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|^2} \rho(|H_{\Lambda_\ell}^{\Psi}|^2) = \Delta_{\mathfrak{e}_\Psi}(\rho) \ .$$

Using Corollary 7.314, we deduce that

$$\lim_{\ell\to\infty} P_{H_{\Lambda_\ell}^{\mathfrak{m}},\beta} \geq -\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \doteq -\inf \left\{ \mathfrak{f}_{\mathfrak{m},\beta}(\rho) : \rho \in E_1 \right\} \ .$$

The upper bound is more difficult to derive, in particular for non-zero mean-field attraction $\Psi_- \neq 0$. See [1, Chapter 6] for more details. □

Like in Definition 5.28, we define the pressure as follows:

Definition 6.14 (Pressure Function on $\mathcal{M}_1$) For $\beta \in (0, \infty)$, the function $\mathfrak{p}_\beta : \mathcal{M}_1 \to \mathbb{R}$ defined by

$$\mathfrak{m} \mapsto \mathfrak{p}_\beta(\mathfrak{m}) \doteq -\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1)$$

is called "pressure function" at temperature $T = \beta^{-1}$.

Similar to Proposition 5.30, for two fixed sequences $\Psi_\pm \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$ of complex (invariant) interactions, the pressure function is a continuous convex real-valued function

$$\Phi \mapsto \mathfrak{p}_\beta(\Phi, \Psi_-, \Psi_+)$$

on the real Banach space $\mathcal{W}_1$ of invariant interactions. In addition, for all Φ, $\Phi' \in \mathcal{W}_1$,

$$|\mathfrak{p}_\beta(\Phi, \Psi_-, \Psi_+) - \mathfrak{p}_\beta(\Phi', \Psi_-, \Psi_+)| \leq \left\| \Phi - \Phi' \right\| .$$

The arguments are the same as those proving Proposition 5.30. We can therefore study tangent functionals to this function, as discussed from Proposition 5.30. However, in the sequel, we perform, instead, a more direct study of the minimizers of the free energy density functional, which are naturally viewed as equilibrium states of the corresponding mean-field model. In fact, this study becomes quite interesting, and highly non-trivial, in the presence of non-zero mean-field terms $\Psi_\pm$.

6.4 Equilibrium States of Mean-Field Models

The free energy density functional on the set E_1 of invariant states is in general **not** weak*-lower semicontinuous: By Lemmata 6.7, 6.8, and 6.10, observe from Definition 6.12 that, for any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and $\beta \in (0, \infty)$,

$$\mathfrak{f}_{\mathfrak{m},\beta} = \underbrace{\Delta_{\Psi_+}}_{\text{upper semicont.}} + \underbrace{\left(-\Delta_{\Psi_-} + \mathfrak{e}_\Phi - \beta^{-1}\mathfrak{s}\right)}_{\text{lower semicont.}} .$$

The free energy density functional $\mathfrak{f}_{\mathfrak{m},\beta} : E_1 \to \mathbb{R}$, which is an affine functional on the convex weak*-compact space E_1 of invariant states, has thus a **topological** drawback. In particular, it is not clear from the beginning whether there are solutions to the variational problem

$$\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) ,$$

or not. The situation is much simpler in the absence of mean-field terms $\Psi_\pm$: When $\Psi_\pm = 0$, the set $M_{\Phi,\beta} \subseteq E_1$ of all minimizers of $\mathfrak{f}_{\Phi,\beta}$, named the globally stable equilibrium states for the interaction $\Phi \in \mathcal{W}_1$ at inverse temperature $\beta \in (0, \infty)$, appearing in Definition 5.22 is always nonempty, $\mathfrak{f}_{\Phi,\beta}$ being lower semicontinuous on a compact set. See Proposition 7.172. The generalization of the notion of globally stable equilibrium states to the mean-field case is done, as is usual, via the (weak*) limits of approximating minimizers:

Definition 6.15 (Equilibrium States) For any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$,

$$\Omega_{\mathfrak{m},\beta} \doteq \Big\{ \omega \in E_1 : \exists (\rho_n)_{n\in\mathbb{N}} \subseteq E_1 \text{ weak}^* \text{ converging to } \omega \text{ such that } \lim_{n\to\infty} \mathfrak{f}_{\mathfrak{m},\beta}(\rho_n) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \Big\} .$$

The set $\Omega_{\mathfrak{m},\beta}$ is clearly convex, $\mathfrak{f}_{\mathfrak{m},\beta}$ being affine (Lemmata 6.7, 6.8, and 6.10) on a convex set, i.e., E_1. Note also that it is not empty, since any sequence of invariant states has weak*-convergent subsequences, the space E_1 of invariant states being weak*-compact.

Elements of the set $\Omega_{\mathfrak{m},\beta} \subseteq E_1$ of all weak* limits of approximating minimizers of $\mathfrak{f}_{\mathfrak{m},\beta}$ are named again "globally stable equilibrium states" at temperature $T = \beta^{-1}$, associated with the mean-field model $\mathfrak{m}$. The extreme elements of the convex set $\Omega_{\mathfrak{m},\beta}$ are called "pure globally stable equilibrium states." As before, we say that there is a "(first-order) phase transition" for $\mathfrak{m} \in \mathcal{M}_1$ at temperature $T = \beta^{-1}$ if $\Omega_{\mathfrak{m},\beta}$ contains more than one element. Recall that, in contrast with the finite-volume situation, there are possibly many globally stable equilibrium states, even in the absence of mean-field terms. See, for instance, Corollary 5.39. This is reminiscent of the non-uniqueness of irreducible representations of the infinite-dimensional unital C^*-algebra $\mathcal{U}$.

Globally stable equilibrium states in the above sense are directly related with the thermodynamic limit of Gibbs states associated with local Hamiltonians of Definition 6.4. We shortly explain this fact: For any cubic box $\Lambda_\ell \subseteq \Gamma$, $\ell \in \mathbb{N}$, let $\omega_{H^{\mathfrak{m}}_{\Lambda_\ell},\beta} \in E(\mathcal{U}_{\Lambda_\ell})$ be the Gibbs state of Definition 5.19, which is periodically extended (with period $(2\ell + 1)$ in each direction of $\Gamma \doteq \mathbb{Z}^d$), and define

$$\hat{\rho}_{\ell,\mathfrak{m},\beta} \doteq \frac{1}{|\Lambda_\ell|} \sum_{x\in\Lambda_\ell} \omega_{H^{\mathfrak{m}}_{\Lambda_\ell},\beta} \circ \tau_x \in \mathcal{E}_1 \subseteq E_1$$

These invariant states are particular cases of the ones used in the proofs of Theorems 6.1 and Lemma 6.7. They are in particular ergodic. Then, one can prove the following statement:

Theorem 6.16 (Limit of Space-Averaged Gibbs States) *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0,\infty)$, *the weak** *accumulation points of* $(\hat{\rho}_{\ell,\mathfrak{m},\beta})_{\ell\in\mathbb{N}}$ *belong to* $\Omega_{\mathfrak{m},\beta}$.

Idea of Proof One uses the notion of tangent functionals (Definition 3.18), as explained in Proposition 5.30. See [1, Section 2.6] for more details. □

Observe that Theorem 6.16 does not exactly refer to the limits of Gibbs states. In fact, the set $E(\mathcal{U})$ of all states on $\mathcal{U}$ being weak*-compact, Gibbs states, seen as periodic states on $\mathcal{U}$, have weak*-convergent subsequences, but it is not clear that such limits always belong to the set E_1 of invariant states, as for the sequence $\{\hat{\rho}_{\ell,\mathfrak{m},\beta}\}_{\ell\in\mathbb{N}} \subseteq E_1$. If a weak*-convergent sequence of Gibbs states has an invariant state as limit, then it must belong to $\Omega_{\mathfrak{m},\beta}$. This condition can be ensured by taking periodic boundary conditions, as explained in [1, Chapter 3]. In particular, in this case, the weak*-accumulation points of Gibbs states $\{\omega_{H^{\mathfrak{m}}_{\Lambda_\ell},\beta}\}_{\ell\in\mathbb{N}}$ belong to $\Omega_{\mathfrak{m},\beta}$.

Apart from the fact that E_1 is convex and weak*-compact, recall that it has a weak*−dense set of extreme points, i.e., the set $\mathcal{E}_1$ of ergodic states is dense in E_1. See Theorem 6.1. Moreover, the space-averaging functional of Definition 6.9 takes a simple (explicit) form on this dense set:

$$\Delta_A(\rho) = |\rho(A)|^2 ,$$

for all ergodic states $\rho \in \mathcal{E}_1$, thanks to Theorem 5.37. In particular, the free energy density functional of Definition 6.12 equals the following function on the dense set of ergodic states:

Definition 6.17 (Nonlinear Free Energy Density) For any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and $\beta \in (0, \infty)$, the "nonlinear[2] free energy density functional" $\mathfrak{g}_{\mathfrak{m},\beta} : E_1 \to \mathbb{R}$ on the space E_1 of all invariant states is defined by

$$\begin{aligned}\mathfrak{g}_{\mathfrak{m},\beta}(\rho) &\doteq \|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2 - \|\mathfrak{e}_{\Psi_-}(\rho)\|_2^2 + \mathfrak{e}_\Phi - \beta^{-1}\mathfrak{s} \\ &= \|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2 - \|\mathfrak{e}_{\Psi_-}(\rho)\|_2^2 + \mathfrak{f}_{\Phi,\beta}(\rho) , \qquad \rho \in E_1 ,\end{aligned}$$

(see Definition 5.22, Theorem 5.20, and Eq. (6.7)), where, for any sequence $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$ of complex interactions,

$$\mathfrak{e}_\Psi(\rho) \doteq \left(\mathfrak{e}_{\Psi_n}(\rho)\right)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}) .$$

Note that, for any sequence $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$ of complex (invariant) interactions,

$$\|\mathfrak{e}_\Psi(\rho)\|_2^2 \le \sum_{n\in\mathbb{N}} \sup_{\rho\in E_1} |\mathfrak{e}_{\Psi_n}(\rho)|^2 \le \sum_{n\in\mathbb{N}} \|\Psi_n\|^2 < \infty , \tag{6.12}$$

thanks to Lemma 6.8. The nonlinear free energy density functional $\mathfrak{g}_{\mathfrak{m},\beta}$ is not affine anymore, but has, instead, the following important properties:

Lemma 6.18 *For every mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$, $\mathfrak{g}_{\mathfrak{m},\beta}$ is weak*-lower semicontinuous. Additionally, for any invariant state $\rho \in E_1$, there is a sequence $(\hat{\rho}_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}_1$ of ergodic states weak*-converging to ρ, such that*

$$\mathfrak{g}_{\mathfrak{m},\beta}(\rho) = \lim_{n\to\infty} \mathfrak{g}_{\mathfrak{m},\beta}(\hat{\rho}_n) .$$

Proof To prove the lower semicontinuity, combine Lemmata 6.7 and 6.8 together with the weak*-continuity of the functional $\rho \mapsto \mathfrak{e}_\Psi(\rho)$ from E_1 to $\ell^2(\mathbb{N})$, which is deduced from (6.12) and Corollary 7.314. The second part of the lemma directly follows from the corresponding property of the entropy density (see Lemma 6.7) combined with the previously proven continuity of the mapping $\rho \mapsto \mathfrak{e}_\Psi(\rho)$. ◻

[2] We adopt this terminology, because of the formal analogy to the classical "nonlinear thermodynamic formalism," as, for instance, described in [11].

As already explained above, the nonlinear free energy density functionals equal the usual ones on the dense set of ergodic states:

$$\mathfrak{f}_{\mathfrak{m},\beta}(\rho) = \mathfrak{g}_{\mathfrak{m},\beta}(\rho) \ , \qquad \rho \in \mathcal{E}_1 \ ,$$

thanks to Theorem 5.37. More generally, for (possibly non-ergodic) invariant states, both functionals are related to each other via the following assertion:

Lemma 6.19 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$, $\beta \in (0,\infty)$ *and every invariant states* $\rho \in E_1$, $\mathfrak{f}_{\mathfrak{m},\beta}(\rho) = \mu_\rho\left(\mathfrak{g}_{\mathfrak{m},\beta}\right)$ *with* μ_ρ *being the positive linear functional defined by Theorem 7.339 on the Borel-measurable*[3] *functions on* E_1.

Proof Combine Lemmata 6.7 and 6.8 with Theorems 7.339 and 6.11 (iv). □

The nonlinear free energy density functional is clearly not affine, in contrast with the free energy density functional, but it is, at least, weak*-lower semicontinuous. In fact, being not affine, $\mathfrak{g}_{\mathfrak{m},\beta}$ has a **geometrical** drawback, whereas, being **not** weak*-lower semicontinuous, $\mathfrak{f}_{\mathfrak{m},\beta}$ has a **topological** drawback. Nonetheless, interestingly, both functionals lead to the pressure function of Definition 6.14:

Theorem 6.20 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0,\infty)$,

$$\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(E_1) > -\infty \ ,$$

with $\mathcal{E}_1$ *being the weak*-dense set of ergodic (or extreme) states of* E_1 *(Theorem 6.1).*

Proof We apply the extension of the Bauer maximum principle (Lemma 7.344) to the weak*-compact and convex space $K = E_1$ and the functional

$$\mathfrak{f}_{\mathfrak{m},\beta} = \underbrace{\Delta_{\Psi_+}}_{\text{upper semicont.}} + \underbrace{\left(-\Delta_{\Psi_-} + \mathfrak{e}_\Phi - \beta^{-1}\mathfrak{s}\right)}_{\text{lower semicont.}}$$

for any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and $\beta \in (0,\infty)$. See Lemmata 6.7, 6.8, and 6.10. In fact, using Lemma 7.344, we conclude that

$$\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(\mathcal{E}_1) \ , \tag{6.13}$$

[3] Semicontinuous functions on a metric space, like $\mathfrak{g}_{\mathfrak{m},\beta}$, are special cases of Borel-measurable ones. Moreover, $\mathfrak{g}_{\mathfrak{m},\beta}$ is the supremum of a countable family of continuous functions, because (up to a sign) the entropy density functional has this property. See Lemma 7.144 and related discussions.

keeping in mind that the free energy density functional $\mathfrak{f}_{\mathfrak{m},\beta}$ equals the nonlinear one, $\mathfrak{g}_{\mathfrak{m},\beta}$, on the (dense) set $\mathcal{E}_1$ of ergodic states. Additionally, using Lemma 6.18, we deduce the following facts:

- $\mathfrak{g}_{\mathfrak{m},\beta}$ is weak*-lower semicontinuous, and, thus, there is a minimizer $\omega \in E_1$ for the variational problem

$$\inf \mathfrak{g}_{\mathfrak{m},\beta}(E_1) = \mathfrak{g}_{\mathfrak{m},\beta}(\omega) \ .$$

- There is a sequence $(\hat{\rho}_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}_1$ of ergodic states weak*-converging to ω and such that

$$\mathfrak{g}_{\mathfrak{m},\beta}(\omega) = \lim_{n\to\infty} \mathfrak{g}_{\mathfrak{m},\beta}(\hat{\rho}_n) \ .$$

It follows that

$$\inf \mathfrak{g}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(E_1) \ ,$$

which combined with (6.13), in turn, yields the assertion. □

This theorem opens the door to a new definition of equilibrium states, which can now be defined as minimizers of the weak*-lower semicontinuous nonlinear free energy functional:

Definition 6.21 (Nonlinear Equilibrium States) For any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$,

$$\hat{M}_{\mathfrak{m},\beta} \doteq \left\{ \omega \in E_1 : \mathfrak{g}_{\mathfrak{m},\beta}(\omega) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(E_1) \right\} \ .$$

The elements of $\hat{M}_{\mathfrak{m},\beta}$ are called here "nonlinear (globally stable) equilibrium states" of the mean-field model $\mathfrak{m}$ at inverse temperature β.

Clearly, $\hat{M}_{\mathfrak{m},\beta}$ is nonempty, since $\mathfrak{g}_{\mathfrak{m},\beta}$ is weak*-lower semicontinuous (Lemma 6.18). See Proposition 7.172. In general, this set of minimizers differs from the set $\Omega_{\mathfrak{m},\beta}$ of (usual) globally stable equilibrium states of Definition 6.15, i.e., $\hat{M}_{\mathfrak{m},\beta} \neq \Omega_{\mathfrak{m},\beta}$. Recall that

$$\Omega_{\mathfrak{m},\beta} \doteq \Big\{ \omega \in E_1 : \exists (\rho_n)_{n\in\mathbb{N}} \subseteq E_1 \text{ weak}^* \text{ converging to } \omega \text{ so that } \lim_{n\to\infty} \mathfrak{f}_{\mathfrak{m},\beta}(\rho_n) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \Big\}$$

for any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$. In fact, even if they are generally different sets, there is a strong relation between both notions of equilibrium states: It turns out that $\hat{M}_{\mathfrak{m},\beta} \subseteq \Omega_{\mathfrak{m},\beta}$, i.e., nonlinear equilibrium states are special cases of globally stable equilibrium states of mean-field models. What is more, the nonlinear equilibrium states *generate* the convex set of all equilibrium states, for *all* mean-field models. These properties are precisely stated in the following lemma and Theorem 6.25:

Lemma 6.22 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$*, the following properties hold true:*

(i) $\Omega_{\mathfrak{m},\beta}$ *is a (nonempty) convex weak***-compact subset of* E_1*.*
(ii) $\hat{M}_{\mathfrak{m},\beta}$ *is a (nonempty) weak***-compact subset of* E_1*.*
(iii) *The weak***-closed convex hull of* $\hat{M}_{\mathfrak{m},\beta}$ *belong to* $\Omega_{\mathfrak{m},\beta}$*, i.e.,* $\overline{\mathrm{co}(\hat{M}_{\mathfrak{m},\beta})} \subseteq \Omega_{\mathfrak{m},\beta}$.

Proof

(i) The set $\Omega_{\mathfrak{m},\beta}$ is convex, $\mathfrak{f}_{\mathfrak{m},\beta}$ being affine (Lemmata 6.7, 6.8, and 6.10) on the convex set E_1. Since the (spin or fermion) algebra $\mathcal{U}$ is separable, the weak* topology is metrizable on any weak*-compact subset of $\mathcal{U}^*$; see, e.g., Proposition 4.84 or [18, Theorem 3.16]. As E_1 is weak*-compact, one uses the metric generating the weak* topology on E_1 in order to show that $\Omega_{\mathfrak{m},\beta} \subseteq E_1$ is weak*-closed and therefore weak*-compact.
(ii) It is a direct consequence of the weak*-lower semicontinuity of $\mathfrak{g}_{\mathfrak{m},\beta}$ (Lemma 6.18) together with the weak*-compactness of E_1. See Proposition 7.172.
(iii) By Lemma 6.18, for any $\omega \in \hat{M}_{\mathfrak{m},\beta}$, there is a sequence $(\hat{\rho}_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}_1$ weak*-converging to ω such that $\mathfrak{g}_{\mathfrak{m},\beta}(\hat{\rho}_n) = \mathfrak{f}_{\mathfrak{m},\beta}(\hat{\rho}_n)$ converges to $\mathfrak{g}_{\mathfrak{m},\beta}(\omega)$, as $n \to \infty$. Since, by Theorem 6.20,

$$\mathfrak{g}_{\mathfrak{m},\beta}(\omega) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \ ,$$

we obtain that $\omega \in \Omega_{\mathfrak{m}}$. As a consequence, the assertion holds true because $\Omega_{\mathfrak{m}}$ is convex and weak*-compact.

□

In fact, we can strengthen Lemma 6.22 (iii) by showing that

$$\overline{\mathrm{co}(\hat{M}_{\mathfrak{m},\beta})} = \Omega_{\mathfrak{m},\beta}$$

for any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$. To prove this equality, we use a relatively recent result of convex analysis [25, Theorem 1.4], which corresponds in our (less general) setting to Theorem 7.345. More precisely, we apply this theorem to the γ-regularization of the free energy density functionals $\mathfrak{g}_{\mathfrak{m},\beta}$ and $\mathfrak{f}_{\mathfrak{m},\beta}$ on the convex weak*-compact space E_1 of invariant states, defined by

$$\gamma\left(\mathfrak{g}_{\mathfrak{m},\beta}\right)(\rho) \doteq \sup\left\{\rho\left(B\right) : B \in \mathrm{Re}\{\mathcal{U}\} \text{ so that } \forall \varpi \in E_1, \varpi\left(B\right) \leq \mathfrak{g}_{\mathfrak{m},\beta}\left(\varpi\right)\right\} \ ,$$
$$\gamma\left(\mathfrak{f}_{\mathfrak{m},\beta}\right)(\rho) \doteq \sup\left\{\rho\left(B\right) : B \in \mathrm{Re}\{\mathcal{U}\} \text{ so that } \forall \varpi \in E_1, \varpi\left(B\right) \leq \mathfrak{f}_{\mathfrak{m},\beta}\left(\varpi\right)\right\} \ ,$$

for any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$. See Definition 7.340 and Proposition 7.347.

Corollary 6.23 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$,

$$\begin{aligned}\inf \gamma(\mathfrak{f}_{\mathfrak{m},\beta})(E_1) &= \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(\mathcal{E}_1) \\ &= \inf \mathfrak{g}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(E_1) = \inf \gamma(\mathfrak{g}_{\mathfrak{m},\beta})(E_1) \, .\end{aligned}$$

Proof Combine Theorem 6.20 with Theorem 7.345. □

The next question is the following: How are the γ-regularizations $\gamma(\mathfrak{f}_{\mathfrak{m},\beta})$ and $\gamma(\mathfrak{g}_{\mathfrak{m},\beta})$ on the convex weak*-compact space E_1 of invariant states (i.e., the largest weak*-lower semicontinuous and convex minorants of, respectively, $\mathfrak{f}_{\mathfrak{m},\beta}$ and $\mathfrak{g}_{\mathfrak{m},\beta}$ on E_1, by Corollary 7.342) related to each other? A simple and satisfying answer to this question is given by the following lemma:

Lemma 6.24 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$, *we have* $\gamma(\mathfrak{f}_{\mathfrak{m},\beta}) = \gamma(\mathfrak{g}_{\mathfrak{m},\beta})$ *on the space* E_1 *of invariant states.*

Proof
Lower bound: As $\Delta_A(\hat{\rho}) = |\hat{\rho}(A)|^2$ on $\mathcal{E}_1$ (see, e.g., Theorem 6.11 (iv)), for any ergodic state $\hat{\rho} \in \mathcal{E}_1$,

$$\mathfrak{f}_{\mathfrak{m},\beta}(\hat{\rho}) = \mathfrak{f}^{\flat}_{\mathfrak{m},\beta}(\hat{\rho}) = \mathfrak{g}_{\mathfrak{m},\beta}(\hat{\rho}) \, ,$$

where, for any invariant state $\rho \in E_1$,

$$\mathfrak{f}^{\flat}_{\mathfrak{m},\beta}(\rho) \doteq \underbrace{\|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2}_{\text{convex semicont.}} + \underbrace{\left(-\Delta_{\Psi_-}(\rho) + \mathfrak{e}_{\Phi}(\rho) - \beta^{-1}\mathfrak{s}(\rho)\right)}_{\text{affine lower semicont.}} \, . \tag{6.14}$$

See Lemmata 6.7, 6.8, and 6.10. Therefore, for any ergodic state $\hat{\rho} \in \mathcal{E}_1$,

$$\mathfrak{f}_{\mathfrak{m},\beta}(\hat{\rho}) = \gamma(\mathfrak{f}_{\mathfrak{m},\beta})(\hat{\rho}) = \mathfrak{g}_{\mathfrak{m},\beta}(\hat{\rho}) \, . \tag{6.15}$$

By Lemma 6.18, for any invariant state $\rho \in E_1$, there is a sequence $(\hat{\rho}_n)_{n\in\mathbb{N}} \subseteq \mathcal{E}_1$ of ergodic states converging to ρ and such that

$$\lim_{n\to\infty} \mathfrak{g}_{\mathfrak{m},\beta}(\hat{\rho}_n) = \mathfrak{g}_{\mathfrak{m},\beta}(\rho) \, .$$

By (6.15) and weak*-lower semicontinuity of $\gamma(\mathfrak{f}_{\mathfrak{m},\beta})$, for any invariant state $\rho \in E_1$,

$$\lim_{n\to\infty} \gamma(\mathfrak{f}_{\mathfrak{m},\beta})(\hat{\rho}_n) = \mathfrak{g}_{\mathfrak{m}}(\rho) \geq \gamma(\mathfrak{f}_{\mathfrak{m},\beta})(\rho),$$

implying $\gamma(\mathfrak{f}_{\mathfrak{m},\beta}) \leq \gamma(\mathfrak{g}_{\mathfrak{m},\beta})$.

Upper bound: By Theorem 7.339 and Jensen's inequality (Lemma 7.330; see also [1, Lemma 10.33]),[4] for any invariant state $\rho \in E_1$, there is a (unique) positive linear functional μ_ρ such that

$$h(\rho) \leq \mu_\rho(h)$$

for any convex weak*-lower semicontinuous complex-valued functions h on E_1. By convexity and weak*-lower semicontinuity of $\gamma(\mathfrak{g}_{\mathfrak{m},\beta})$, it follows that

$$\gamma(\mathfrak{g}_{\mathfrak{m},\beta})(\rho) \leq \mu_\rho\left(\gamma(\mathfrak{g}_{\mathfrak{m},\beta})\right) \underset{(=)}{\leq} \mu_\rho\left(\mathfrak{g}_{\mathfrak{m},\beta}\right)$$

which combined with Lemma 6.19 yields $\gamma(\mathfrak{g}_{\mathfrak{m},\beta}) \leq \mathfrak{f}_{\mathfrak{m},\beta}$ and therefore $\gamma(\mathfrak{g}_{\mathfrak{m},\beta}) \leq \gamma(\mathfrak{f}_{\mathfrak{m},\beta})$. □

We are now in a position to prove that the weak*-closed convex hull of the set nonlinear globally stable equilibrium states (see Definition 6.21) is precisely the set of all (usual) globally stable equilibrium states (Definition 6.15):

Theorem 6.25 *For any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$,*

$$\Omega_{\mathfrak{m},\beta} = \overline{\mathrm{co}(\hat{M}_{\mathfrak{m},\beta})}\ .$$

Moreover, if $\Psi_- = 0$, then $\Omega_{\mathfrak{m},\beta} = \hat{M}_{\mathfrak{m},\beta}$.

Proof Apply Theorem 7.345 to the convex and weak*-compact space $K = E_1$ of invariant states and the real functional $\varphi = \mathfrak{f}_{\mathfrak{m},\beta}$ to show that the set M of minimizers of $\gamma(\mathfrak{f}_{\mathfrak{m},\beta})$ over E_1 is

$$M = \overline{\mathrm{co}\left(\Omega_{\mathfrak{m},\beta}\right)}\ .$$

As $\gamma(\mathfrak{g}_{\mathfrak{m},\beta}) = \gamma(\mathfrak{f}_{\mathfrak{m},\beta})$ (Lemma 6.24), we also deduce from Theorem 7.345 that

$$M = \overline{\mathrm{co}(\hat{M}_{\mathfrak{m},\beta})}\ .$$

[4] For any $\rho \in E$, the positive linear functional μ_ρ is associated with a probability measure on the set $\mathcal{E}_1$ of ergodic states such that

$$\rho = \int_{\mathcal{E}_1} \mathrm{d}\mu_\rho(\hat{\rho})\ \hat{\rho}$$

(in the weak sense). This is reminiscent of the Riesz-Markov theorem. This observation highlights the use of Jensen's inequality, which states that the image of an expectation value of a random variable by a convex function is less than or equal to the expectation value of the image of the random variable by the same function.

By Lemma 6.22, it follows that

$$\Omega_{\mathfrak{m},\beta} = \overline{\mathrm{co}\left(\Omega_{\mathfrak{m},\beta}\right)} = \overline{\mathrm{co}(\hat{M}_{\mathfrak{m},\beta})} \ .$$

Assume now $\Psi_- = 0$. Then, $\mathfrak{g}_{\mathfrak{m},\beta}$ becomes convex. So, $\hat{M}_{\mathfrak{m}}$ is also convex and weak*-compact, because of Lemma 6.22 (ii) and the equality

$$\overline{\mathrm{co}(\hat{M}_{\mathfrak{m},\beta})} = \hat{M}_{\mathfrak{m},\beta} \ . \qquad \square$$

6.5 Approximating Invariant Interactions

In the previous sections, we describe the set of globally stable equilibrium states of mean-field models by means of different variational problems. However, it is a priori not clear how useful these variational formulae are to study phase transitions. To answer to this question, it is convenient to consider the so-called Bogoliubov approximations of mean-field models, which are reminiscent of the "approximating Hamiltonian method" used in the past to compute the pressure associated with particular mean-field models, as explained in [1, Section 2.10]. In [1], we generalize this method in such a way that it can be applied to all elements of the Banach space of mean-field models, as well as to the corresponding equilibrium states. We use the viewpoint of game theory by interpreting the mean-field attractions Ψ_- and repulsions Ψ_+ of any model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ as attractive and repulsive players, respectively. This leads to a two-person zero-sum game named in [1] the "thermodynamic game," which is defined as follows:

Using the Hilbert space of square-integrable sequences

$$\ell^2(\mathbb{N}) \equiv \ell^2(\mathbb{N};\mathbb{C}) \doteq \left\{ c \equiv (c_n)_{n\in\mathbb{N}} \subseteq \mathbb{C} : \|c\|_2^2 \doteq \sum_{n\in\mathbb{N}} |c_n|^2 < \infty \right\} ,$$

we first define approximating (short-range) invariant interactions associated with mean-field models:

Definition 6.26 (Approximating Interactions) For any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and sequences $c_-, c^+ \in \ell^2(\mathbb{N})$, we define the corresponding "approximating interaction" to be

$$\Phi_{\mathfrak{m}}(c_-, c_+) \doteq \Phi + 2 \sum_{n\in\mathbb{N}} \left(\mathrm{Re}\left\{ \overline{c_+} \Psi_{+,n} \right\} - \mathrm{Re}\left\{ \overline{c_-} \Psi_{-,n} \right\} \right) \in \mathcal{W}_1 \ .$$

This interaction is a well-defined element of the Banach space $\mathcal{W}_1$ because, for any $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and $c_-, c_+ \in \ell^2(\mathbb{N})$,

$$\begin{aligned}\|\Phi_{\mathfrak{m}}(c_-,c_+)\| &\leq \|\Phi\| + 2\sum_{n\in\mathbb{N}}\left(|c_{+,n}|\,\|\Psi_{+,n}\| + |c_{-,n}|\,\|\Psi_{-,n}\|\right)\\ &\leq \|\Phi\| + 2\,\|c_+\|_2\,\|\Psi_+\|_2 + 2\,\|c_-\|_2\,\|\Psi_-\|_2\\ &\leq \max\left\{1, 2\,\|c_+\|_2\,, 2\,\|c_-\|_2\right\}\|\mathfrak{m}\| < \infty\,,\end{aligned}$$

thanks to the triangle and Cauchy-Schwarz inequalities. See Definition 6.3.

For each mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and $c_-, c_+ \in \ell^2(\mathbb{N})$, observe from Definition 5.10 that, for any finite subset $\Lambda \in \mathcal{P}_f$,

$$H_\Lambda^{\Phi_{\mathfrak{m}}(c_-,c_+)} \doteq \sum_{\Lambda'\in\mathcal{P}_f,\ \Lambda'\subseteq\Lambda} \Phi(\Lambda') = H_\Lambda^{\Phi} + 2\sum_{n\in\mathbb{N}}\left(\mathrm{Re}\{\overline{c_+}H_\Lambda^{\Psi_{+,n}}\} - \mathrm{Re}\{\overline{c_-}H_\Lambda^{\Psi_{-,n}}\}\right).$$

Compare this expression with the full Hamiltonian

$$H_\Lambda^{\mathfrak{m}} \doteq H_\Lambda^{\Phi} + \frac{1}{|\Lambda|}\sum_{n\in\mathbb{N}}\left(|H_\Lambda^{\Psi_{+,n}}|^2 - |H_\Lambda^{\Psi_{-,n}}|^2\right)$$

associated with the mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ for $\Lambda \in \mathcal{P}_f$. See Definition 6.4. (Recall that the identification $\Phi \equiv (\Phi, 0, 0)$ for $\Phi \in \mathcal{W}_1$ is coherent with Definitions 5.10 and 6.4.) In particular,

$$\begin{aligned}|\Lambda|^{-1}(H_\Lambda^{\mathfrak{m}} - H_\Lambda^{\Phi_{\mathfrak{m}}(c_-,c_+)}) + \|c_+\|_2^2 - \|c_-\|_2^2 = \sum_{n\in\mathbb{N}}\Big(&|(|\Lambda|^{-1}H_\Lambda^{\Psi_{+,n}} - c_{+,n})|^2\\ &-|(|\Lambda|^{-1}H_\Lambda^{\Psi_{-,n}} - c_{-,n})|^2\Big)\,.\end{aligned} \tag{6.16}$$

This last expression shall be considered in infinite volume limit $\Lambda \uparrow \Gamma$: If (6.16) would vanish as $\Lambda \uparrow \Gamma$, then one could replace the mean-field model by the simpler model given by the corresponding approximating interaction. Note, however, that this argument is only heuristic, since we compare in the left-hand side of (6.16) a sum over non-commuting elements of the spin or fermion algebra with complex numbers. In fact, the relation between mean-field models and its approximating interactions can be more properly understood via their respective pressures (Definition 6.14). See also Theorem 6.13, which links the pressure function with local Hamiltonians.

Using Definition 6.14 and Theorem 6.20, we first recall the pressures associated with mean-field models $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ and their approximating interactions $\Phi_{\mathfrak{m}}(c_-, c_+)$ for $c_-, c_+ \in \ell^2(\mathbb{N})$, at a given inverse temperature $\beta \in (0, \infty)$:

- Pressure of mean-field models:

$$-\mathrm{p}_{\beta}(\mathfrak{m}) \doteq \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(E_1) \ , \tag{6.17}$$

  where, by Definition 6.17, for any invariant state $\rho \in E_1$,

$$\mathfrak{g}_{\mathfrak{m},\beta}(\rho) \doteq \|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2 - \|\mathfrak{e}_{\Psi_-}(\rho)\|_2^2 + \mathfrak{f}_{\Phi,\beta}(\rho) \, , \tag{6.18}$$

  with $\mathfrak{e}_{\Psi}(\rho) \doteq \left(\mathfrak{e}_{\Psi_n}(\rho)\right)_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$ for any $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$.
- Pressure of approximating interactions:

$$-\mathrm{p}_{\beta}(\Phi_{\mathfrak{m}}(c_-, c_+)) \doteq \inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}(E_1) = \inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}(\mathcal{E}_1) \ , \tag{6.19}$$

  where, by Definition 6.12, for any invariant state $\rho \in E_1$,

$$\mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}(\rho) = 2\,\mathrm{Re}\left\langle c_+, \mathfrak{e}_{\Psi_+}(\rho)\right\rangle - 2\,\mathrm{Re}\left\langle c_-, \mathfrak{e}_{\Psi_-}(\rho)\right\rangle + \mathfrak{f}_{\Phi,\beta}(\rho) \tag{6.20}$$

  with $\langle \cdot, \cdot \rangle$ being the usual scalar product in the Hilbert space $\ell^2(\mathbb{N})$.

Keeping in mind (6.17) and (6.19), the question we shall answer is whether one can find particular sequences $d_+, d_- \in \ell^2(\mathbb{N})$ such that

$$\mathrm{p}_{\beta}(\mathfrak{m}) = \mathrm{p}_{\beta}(\Phi_{\mathfrak{m}}(d_+, d_-)) \ .$$

In fact, we construct such sequences via the so-called thermodynamic game associated with the given mean-field model. However, before explaining (later, in Sect. 6.7) in detail this game and the related construction of sequences, we make a simple observation, leading us to the appropriate payoff function for the thermodynamic game. In fact, one should compare (6.19) with (6.18) in light of the following equality:

Lemma 6.27 *For any invariant state $\rho \in E_1$ and every sequence $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$,*

$$\sup_{c\in\ell^2(\mathbb{N})} \left\{ - \|c\|_2^2 + 2\,\mathrm{Re}\,\langle c, \mathfrak{e}_{\Psi}(\rho)\rangle \right\} = \|\mathfrak{e}_{\Psi}(\rho)\|_2^2$$

with unique maximizer $d(\rho) = \mathfrak{e}_{\Psi}(\rho) \doteq (\mathfrak{e}_{\Psi_n}(\rho))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$.

Proof Obviously, for any invariant state $\rho \in E_1$, complex number $c \in \mathbb{C}$, and algebra element $A \in \mathcal{U}$,

$$|\rho(A - c)|^2 = |\rho(A)|^2 - 2\,\mathrm{Re}\,\{\rho(A)\,\bar{c}\} + |c|^2 \geq 0 \, , \tag{6.21}$$

which in turn implies that

$$|\rho(A)|^2 = \sup_{c \in \mathbb{C}} \left\{ -|c|^2 + 2\mathrm{Re}\{\rho(A)\bar{c}\} \right\}$$

with unique maximizer $d = \rho(A)$. This assertion yields the lemma, keeping in mind that $\mathfrak{e}_\Phi(\rho) = \rho(e_\Phi)$ for any invariant state $\rho \in E_1$ and complex interaction $\Phi \in \mathcal{W}_1^{\mathbb{C}}$. Here, $e_\Phi \in \mathcal{U}^e$ is defined by (6.7). □

Keeping in mind Eqs. (6.18) and (6.19) and Lemma 6.27, we define the following approximating free energy density for mean-field models:

Definition 6.28 (Approximating Free Energy Density) For any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$, the corresponding "approximating free energy density" is the function $h_{\mathfrak{m},\beta} : \ell^2(\mathbb{N}) \times \ell^2(\mathbb{N}) \to \mathbb{R}$ defined by

$$h_{\mathfrak{m},\beta}(c_-, c_+) \doteq -\|c_+\|_2^2 + \|c_-\|_2^2 + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}(E_1) .$$

The thermodynamic game will be the two-person zero-sum game whose payoff function is nothing else than the above-defined approximating free energy density for the given mean-field model. Before explaining this game in Sect. 6.7, as well as its consequences for the theory of equilibrium states of general mean-field models, we first study the special case of purely attractive mean-field models. In fact, considering the special attractive case gives some insight in how to tackle the above-explained problem for general mean-field models.

6.6 Purely Attractive Mean-Field Models and Application to the BCS Theory

6.6.1 Purely Attractive Mean-Field Models

Recall that mean-field models are elements $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$, where $\mathcal{M}_1 \doteq \mathcal{W}_1 \times \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})^2$. See Definition 6.3. Recall that, for any such a mean-field model $\mathfrak{m}$, the component Ψ_- represents its mean-field attraction, while Ψ_+ is its mean-field repulsion, and, consequently, a mean-field model (Φ, Ψ_-, Ψ_+) is said to be purely attractive if $\Psi_+ = 0$, while it is purely repulsive if $\Psi_- = 0$. In this section, we are interested in the study of purely *attractive* mean-field models. This is the easiest mean-field case to study. Moreover, (partial) results referring to this particular case are pivotal to analyze the general case, later on. We start by proving a relation between the pressure function (Definition 6.14) of purely attractive mean-field models and the (payoff) function of Definition 6.28:

Proposition 6.29 *For any purely attractive mean-field model* $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$,

$$\mathfrak{p}_\beta(\mathfrak{m}) \doteq -\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = -\inf_{c_- \in \overline{B}_R(0)} h_{\mathfrak{m},\beta}(c_-, 0)$$

with $h_{\mathfrak{m},\beta}(c_-, 0)$ *defined as in Definition 6.28 and* $\overline{B}_R(0) \subseteq \ell^2(\mathbb{N})$ *being a closed ball of sufficiently large radius* $R > 0$, *centered at* 0.

Proof Fix $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$ and $\beta \in (0, \infty)$. By Theorem 6.20 and Definition 6.17,

$$\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) = \inf \mathfrak{g}_{\mathfrak{m},\beta}(\mathcal{E}_1) = \inf_{\rho \in \mathcal{E}_1} \left\{ -\|\mathfrak{e}_{\Psi_-}(\rho)\|_2^2 + \mathfrak{f}_{\Phi,\beta}(\rho) \right\} .$$

From Lemma 6.27, Eqs. (6.19)–(6.20) and Definition 6.28, it follows that

$$\begin{aligned}
\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) &= \inf_{\rho \in \mathcal{E}_1} \inf_{c_- \in \ell^2(\mathbb{N})} \left\{ \|c_-\|_2^2 - 2\,\mathrm{Re}\,\langle c_-, \mathfrak{e}_\Psi(\rho) \rangle + \mathfrak{f}_{\Phi,\beta}(\rho) \right\} \\
&= \inf_{\rho \in \mathcal{E}_1} \inf_{c_- \in \ell^2(\mathbb{N})} \left\{ \|c_-\|_2^2 + \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}(\rho) \right\} \\
&= \inf_{c_- \in \ell^2(\mathbb{N})} \left\{ \|c_-\|_2^2 + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}(\mathcal{E}_1) \right\} \\
&= \inf_{c_- \in \ell^2(\mathbb{N})} \left\{ \|c_-\|_2^2 + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}(E_1) \right\} \\
&= \inf_{c_- \in \ell^2(\mathbb{N})} h_{\mathfrak{m},\beta}(c_-, 0) \ .
\end{aligned} \tag{6.22}$$

Finally, the existence of a radius $R > 0$ such that

$$\inf_{c_- \in \ell^2(\mathbb{N})} h_{\mathfrak{m}}(c_-, 0) = \inf_{c_- \in \overline{B}_R(0)} h_{\mathfrak{m},\beta}(c_-, 0)$$

directly follows from the fact that, for all sequences $c_- \in \ell^2(\mathbb{N})$,

$$\begin{aligned}
|\inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}(E_1)| &\leq 2 \sup_{\rho \in E_1} \left|\langle c_-, \mathfrak{e}_{\Psi_-}(\rho)\rangle\right| + \left|\inf \mathfrak{f}_{\Phi,\beta}(E_1)\right| \\
&\leq 2\|c\|_2 \|\Psi_-\| + \left|\inf \mathfrak{f}_{\Phi,\beta}(E_1)\right| \ ,
\end{aligned}$$

by the Cauchy-Schwarz inequality, as well as the bound $|\mathfrak{e}_{\Psi_-}(\rho)| \leq \|\Psi_-\|$ (Lemma 6.8). □

Proposition 6.29 is reminiscent of the so-called Bogoliubov approximation, which formally consists in replacing specific operators appearing in the Hamiltonian of a given physical system with constants that are determined as solutions to some self-consistency equation or to some associated variational problem.

In light of Proposition 6.29, the set of minimizers of the approximating free energy density $h_{\mathfrak{m},\beta}(\cdot, 0)$ should play an important role. As a consequence, we define the set

$$\mathcal{C}_{\mathfrak{m},\beta} \doteq \left\{ d_- \in \ell^2(\mathbb{N}) : h_{\mathfrak{m},\beta}(d_-, 0) = \inf_{c_- \in \ell^2(\mathbb{N})} h_{\mathfrak{m},\beta}(c_-, 0) \right\} \tag{6.23}$$

for any purely attractive mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$ and every inverse temperature $\beta \in (0, \infty)$. The set $\mathcal{C}_{\mathfrak{m},\beta} \subseteq \ell^2(\mathbb{N})$ is nonempty, norm-bounded, and weakly compact when $\Psi_- \neq 0$. See [1, Lemma 8.4]. The next step is to understand the relation between the above set of minimizers of the approximating free energy density and globally stable equilibrium states.

Recall the definition of globally stable equilibrium states: For any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and $\beta \in (0, \infty)$,

$$\Omega_{\mathfrak{m},\beta} \doteq \Big\{ \omega \in E_1 : \exists (\rho_n)_{n\in\mathbb{N}} \subseteq E_1 \text{ weak}^* \text{ converging to } \omega \text{ so that } \lim_{n\to\infty} \mathfrak{f}_{\mathfrak{m},\beta}(\rho_n) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \Big\} .$$

See Definition 6.15. This set is always convex and weak*-compact, by Lemma 6.22 (i). When the model is purely attractive, the set of globally stable equilibrium states is a *face* of E_1. Recall that a face F of a convex set K is defined to be a subset of K with the property that, if $\rho = \lambda_1 \rho_1 + \cdots + \lambda_n \rho_n \in F$ with $\rho_1, \ldots, \rho_n \in K$, $\lambda_1, \ldots, \lambda_n \in (0, 1)$ and $\lambda_1 + \cdots + \lambda_n = 1$, then $\rho_1, \ldots, \rho_n \in F$. See Definition 7.333.

Lemma 6.30 *For any model* $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$ *and* $\beta \in (0, \infty)$,

$$\Omega_{\mathfrak{m},\beta} = \left\{ \omega \in E_1 : \mathfrak{f}_{\mathfrak{m},\beta}(\omega) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \right\}$$

with extreme points being all ergodic, i.e., $\mathcal{E}(\Omega_{\mathfrak{m},\beta}) = \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1$. *In particular, it is a (nonempty) weak*-closed face of the convex weak*-compact space* E_1 *of invariant states.*

Proof For $\beta \in (0, \infty)$ and any purely attractive mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$, $\mathfrak{f}_{\mathfrak{m},\beta}$ is weak*-lower semicontinuous and affine; see Lemmata 6.7, 6.8, and 6.10 as well as Definition 6.12. The weak*-lower semicontinuity of $\mathfrak{f}_{\mathfrak{m},\beta}$ yields

$$\Omega_{\mathfrak{m},\beta} = \left\{ \omega \in E_1 : \mathfrak{f}_{\mathfrak{m},\beta}(\omega) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \right\} ,$$

while its affineness on the convex set E_1 of invariant states implies that the set $\mathcal{E}\left(\Omega_{\mathfrak{m},\beta}\right)$ of extreme points of $\Omega_{\mathfrak{m},\beta}$ belongs to the set $\mathcal{E}_1$ of ergodic states of E_1, i.e.,

$$\mathcal{E}\left(\Omega_{\mathfrak{m},\beta}\right) = \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1 .$$

□

Lemma 6.30 of course holds true for all interactions $\Phi \equiv (\Phi, 0, 0) \in \mathcal{M}_1$, in particular for all approximating interactions of Definition 6.26, associated with any (not necessarily purely attractive) mean-field model.

Now, we are in a position to establish a precise relation between the solutions to either variational problems given in Proposition 6.29. This is done through globally stable equilibrium states associated with approximating interactions and leads to self-consistency conditions for these equilibrium states:

Proposition 6.31 (Gap Equations) *For any purely attractive mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$, the following properties hold true:*

(i) *For all ergodic globally stable equilibrium states $\hat{\omega} \in \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1$,*

$$d_- \doteq \mathfrak{e}_{\Psi_-}\left(\hat{\omega}\right) \doteq (\mathfrak{e}_{\Psi_{-,n}}\left(\hat{\omega}\right))_{n\in\mathbb{N}} \in \mathcal{C}_{\mathfrak{m},\beta}$$

and $\hat{\omega} \in \Omega_{\Phi_{\mathfrak{m}}(d_-,0),\beta}$.

(ii) *Conversely, for any fixed $d_- \in \mathcal{C}_{\mathfrak{m},\beta}$,*

$$\Omega_{\Phi_{\mathfrak{m}}(d_-,0),\beta} \cap \mathcal{E}_1 \subseteq \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1$$

and every $\omega \in \Omega_{\Phi_{\mathfrak{m}}(d_-,0),\beta}$ satisfies the equality $d_- = \mathfrak{e}_{\Psi_-}(\omega)$.

Proof

(i) Any ergodic equilibrium state $\hat{\omega} \in \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1$ is a solution to the right-hand side of (6.22), and the solution $d_- = d_-\left(\hat{\omega}\right)$ of

$$\inf_{c_- \in \ell^2(\mathbb{N})} \left\{ \|c_-\|_2^2 + \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}\left(\hat{\omega}\right) \right\}$$

satisfies the (Euler-Lagrange) equation $d_-\left(\hat{\omega}\right) = \mathfrak{e}_{\Psi_-}(\omega)$, by Lemma 6.27. The two infima in (6.22) commute with each other and, thus, $d_- = d_-\left(\hat{\omega}\right) \in \mathcal{C}_{\mathfrak{m},\beta}$ and $\hat{\omega} \in \Omega_{\Phi_{\mathfrak{m}}(d_-,0),\beta}$.

(ii) By definition, any sequence $d_- \in \mathcal{C}_{\mathfrak{m},\beta}$ satisfies

$$\|d_-\|_2^2 + \inf_{\rho \in E_1} \mathfrak{f}_{\Phi_{\mathfrak{m}}(d_-,0),\beta}(\rho) = \inf_{c_- \in \ell^2(\mathbb{N})} \left\{ \|c_-\|_2^2 + \inf_{\rho \in E_1} \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}(\rho) \right\} . \tag{6.24}$$

Since the two infima in the right-hand side of this equality commute with each other as before, any equilibrium state $\omega \in \Omega_{\Phi_{\mathfrak{m}}(d_-,0),\beta}$ satisfies $d_- = \mathfrak{e}_{\Psi_-}(\omega)$ because of Lemma 6.27 and

$$\Omega_{\Phi_{\mathfrak{m}}(d_-,0),\beta} \cap \mathcal{E}_1 \subseteq \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1$$

because of Eq. (6.22).

□

Corollary 6.32 *For any purely attractive mean-field model* $\mathfrak{m} \doteq (\Phi, \Psi_-, 0) \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$,

$$\Omega_{\mathfrak{m},\beta} = \overline{\mathrm{co}\left(\bigcup_{d_- \in \mathcal{C}_{\mathfrak{m},\beta}} \Omega_{\Phi_\mathfrak{m}(d_-,0),\beta}\right)} .$$

Proof Combine Proposition 6.31 with Lemma 6.30. □

In the physics literature on superconductors, the self-consistency condition (Euler-Lagrange equation)

$$d_- = \mathfrak{e}_{\Psi_-}(\omega) \ , \qquad d_- \in \mathcal{C}_{\mathfrak{m},\beta}, \ \omega \in \Omega_{\Phi_\mathfrak{m}(d_-,0),\beta} \ ,$$

refers to the so-called gap equation. We keep this terminology here, although in a much broader and abstract sense. Proposition 6.31 and Corollary 6.32 demonstrate that, for all ergodic (globally stable) equilibrium states $\hat{\omega} \in \Omega_{\mathfrak{m},\beta} \cap \mathcal{E}_1$, the pair $(\hat{\omega}, \mathfrak{e}_{\Psi_-}(\hat{\omega}))$ solves the gap equation, since $\hat{\omega} \in \Omega_{\Phi_\mathfrak{m}(d_-,0),\beta}$. This mathematically justifies the theoretical physics approach using the above self-consistency condition to find the infinite-volume properties of mean-field models. Note that we have shown this property only for purely attractive mean-field models, so far, but we will explain it in the sequel for any general mean-field model.

6.6.2 Application to the BCS Theory on Lattices

The gap equation is pivotal to prove the existence of phase transitions for mean-field models. To illustrate this, as a physically relevant application, we describe the (reduced) BCS model of superconductivity:

(i) General Setup Like in Example 6.6, fix $\Omega \doteq \{\uparrow, \downarrow\}$ and $\mathcal{U} \doteq \mathrm{CAR}(\{\uparrow, \downarrow\}, \Gamma)$. We consider fermions in the cubic box

$$\Lambda_\ell \doteq \{(x_1, \ldots, x_d) \in \Gamma : |x_i| \leq \ell\}$$

for some fixed length $\ell \in \mathbb{N}$. As the BCS model is usually written in Fourier space, we additionally define

$$\Lambda_\ell^* \doteq \frac{2\pi}{(2\ell + 1)} \Lambda_\ell \subseteq [-\pi, \pi]^d \ ,$$

the reciprocal lattice of quasi-momenta (referring to periodic boundary conditions). Then, for any spin $\mathrm{s} \in \{\uparrow, \downarrow\}$ and (quasi-) momentum $k \in \Lambda_\ell^*$, let

$$\phi_{k,\mathrm{s}}(\mathrm{t},x) \doteq \frac{1}{|\Lambda_\ell|^{1/2}} \chi_{\Lambda_\ell} \exp(-ik\cdot x)\,\delta_{\mathrm{s,t}}\,, \qquad x\in\Gamma,\ t\in\{\uparrow,\downarrow\}\,,$$

where $\delta_{\mathrm{s,t}}$ is the Kronecker delta, while χ_{Λ_ℓ} is the characteristic function of the cubic box Λ_ℓ.

(ii) The BCS Hamiltonian the Lattice Theoretical foundations of superconductivity go back to the celebrated BCS theory—appeared in the late 1950s (1957)—which explains conventional type I superconductors. The lattice version of this theory is based on the so-called (reduced) BCS Hamiltonian defined, for any $\ell\in\mathbb{N}$, by

$$\mathrm{H}_{\Lambda_\ell}^{BCS} \doteq \underbrace{\sum_{k\in\Lambda_\ell^*,\ \mathrm{s}\in\{\uparrow,\downarrow\}} \hat{\varepsilon}(k)\,\hat{a}_{k,\mathrm{s}}^*\hat{a}_{k,\mathrm{s}}}_{\text{kinetic term}} \underbrace{-\frac{1}{|\Lambda_\ell|}\sum_{k,q\in\Lambda_\ell^*} \eta_{k,q}\hat{a}_{k,\uparrow}^*\hat{a}_{-k,\downarrow}^*\hat{a}_{q,\downarrow}\hat{a}_{-q,\uparrow}}_{\text{attractive interactions}}\,,$$

where $\hat{a}_{k,\mathrm{s}} \doteq a(\phi_{k,\mathrm{s}})$ annihilates a fermion with spin $\mathrm{s}\in\{\uparrow,\downarrow\}$ and (quasi-) momentum $k\in\Lambda_\ell^*$, while $\hat{\varepsilon}$ is the Fourier transform of some real-valued function ε on Γ. In physics, $\{\hat{\varepsilon}(k)\}_{k\in\Lambda_\ell^*}$ is (up to some constant) the spectrum of the discrete Laplacian and

$$\eta_{k,q} = \begin{cases} \eta\geq 0 \text{ for } |k-q|\leq \mathrm{C} \\ 0 \qquad \text{for } |k-q|>\mathrm{C} \end{cases}$$

with constant $\mathrm{C}\in(0,\infty]$. For simplicity, take once and for all $\mathrm{C}=\infty$. In this case, the BCS Hamiltonian can be written in the "x-space" as

$$\mathrm{H}_{\Lambda_\ell}^{BCS} = \sum_{x,y\in\Lambda_\ell,\ \mathrm{s}\in\{\uparrow,\downarrow\}} \varepsilon(x-y)\,a_{x,\mathrm{s}}^*a_{y,\mathrm{s}} - \frac{\eta}{|\Lambda_\ell|}\sum_{x,y\in\Lambda_\ell} a_{x,\uparrow}^*a_{x,\downarrow}^*a_{y,\downarrow}a_{y,\uparrow} \qquad (6.25)$$

for any $\ell\in\mathbb{N}$, where $a_{x,\mathrm{s}} \doteq a(\mathrm{e}_{\mathrm{s},x})$ annihilates a fermion with spin $\mathrm{s}\in\{\uparrow,\downarrow\}$ and lattice position $x\in\Gamma$. Here, $\{\mathrm{e}_{\mathrm{s},x}\}_{(\mathrm{s},x)\in\{\uparrow,\downarrow\}\times\Gamma}$ is the (canonical) Hilbert basis of $\ell^2(\{\uparrow,\downarrow\}\times\Gamma)$.

(iii) BCS Mean-Field Model Like in Example 6.6, the "BCS interaction" Ψ_{BCS} is defined by $\Psi_{\mathrm{BCS}}(\Lambda)\doteq 0$ whenever $|\Lambda|\notin\{1\}$ and $\Psi_{\mathrm{BCS}}(\{x\})\doteq\eta^{1/2}a_{x,\downarrow}a_{x,\uparrow}$ for any $x\in\Gamma$. Then, for the purely attractive mean-field model

$$\mathfrak{n}\doteq(\Phi,(\Psi_{\mathrm{BCS}},0,0,\ldots),0)\in\mathcal{M}_1\,,$$

where $\Phi\in\mathcal{W}_1$ is some invariant interaction, we observe that

$$H_{\Lambda_\ell}^{\mathfrak{n}} = H_{\Lambda_\ell}^{\Phi} - \frac{\eta}{|\Lambda_\ell|}\sum_{x,y\in\Lambda_\ell} a_{x,\uparrow}^*a_{x,\downarrow}^*a_{y,\downarrow}a_{y,\uparrow}\,,$$

$$H_{\Lambda_\ell}^{\Phi_{\mathfrak{n}}(c_-,0)} = H_{\Lambda_\ell}^{\Phi} - \eta^{1/2} \sum_{x\in\Lambda_\ell} \left(c_{-,1} a_{x,\uparrow}^* a_{x,\downarrow}^* + \overline{c_{-,1}} a_{x,\downarrow} a_{x,\uparrow} \right) ,$$

for any $\ell \in \mathbb{N}$ and $c_- \in \ell^2(\mathbb{N})$. Note that the use of general sequences $c_- \in \ell^2(\mathbb{N})$ is not necessary in this example, since the model has only one non-zero attractive mean-field component, Ψ_{BCS}. One can thus consider constants $c_- \equiv c_{-,1} \in \mathbb{C}$, instead of full sequences $c_- \in \ell^2(\mathbb{N})$. By Corollary 6.32, if one is able to determine the set of states

$$\bigcup_{d_-\in\mathcal{C}_{\mathfrak{n},\beta}} \Omega_{\Phi_{\mathfrak{n}}(d_-,0),\beta} \tag{6.26}$$

then we obtain from it all the equilibrium states of the purely attractive mean-field model $\mathfrak{n}$. Under periodic boundary conditions [1, Chapter 3], we would then know all accumulation points of Gibbs states (in particular all correlation functions) associated with local Hamiltonians $H_{\Lambda_\ell}^{\mathfrak{n}}$, $\ell \in \mathbb{N}$.

(iv) Thermodynamic of the BCS Model Recall that ε is some real-valued function on Γ. We define the parameter $\Phi \in \mathcal{W}_1$ of the mean field model $\mathfrak{n}$ by

$$\begin{aligned}\Phi(\Lambda) \doteq \frac{1}{1+\delta_{x,y}} \Big(\varepsilon(x-y) \left(a_{x,\uparrow}^* a_{y,\uparrow} + a_{x,\downarrow}^* a_{y,\downarrow} \right) \\ + \varepsilon(y-x) \left(a_{y,\uparrow}^* a_{x,\uparrow} + a_{y,\downarrow}^* a_{x,\downarrow} \right) \Big)\end{aligned} \tag{6.27}$$

whenever $\Lambda = \{x, y\}$ and $\Phi(\Lambda) = 0$, otherwise. Here, $\delta_{x,y}$ is the Kronecker delta. Observe that, for any $\ell \in \mathbb{N}$,

$$H_{\Lambda_\ell}^{\mathfrak{n}} = \mathrm{H}_{\Lambda_\ell}^{BCS} ,$$

as well as

$$\begin{aligned}H_{\Lambda_\ell}^{\Phi_{\mathfrak{n}}(c_-,0)} = &\sum_{x,y\in\Lambda_\ell,\ \mathrm{s}\in\{\uparrow,\downarrow\}} \varepsilon(x-y)\, a_{x,\mathrm{s}}^* a_{y,\mathrm{s}} \\ &- \eta^{1/2} \sum_{x\in\Lambda_\ell} \left(c_- a_{x,\uparrow}^* a_{x,\downarrow}^* + \overline{c_-} a_{x,\downarrow} a_{x,\uparrow} \right)\end{aligned}$$

for any complex number $c_- \in \mathbb{C}$. This approximating model is quadratic in the annihilation and creation operators. Such Hamiltonians can be exactly diagonalized, which means that the corresponding pressure can be explicitly computed as a function of the parameter c_-. As a consequence, via Theorem 6.13, the approximating free energy density $h_{\mathfrak{n},\beta}(c_-, 0)$ of Definition 6.28, the solutions to the variational problem

$$\inf \left\{ h_{\mathfrak{n},\beta}(c_-, 0) : c_- \in \mathbb{C} \right\} = \inf \left\{ h_{\mathfrak{n},\beta}(c_-, 0) : |c_-| \le R \right\}$$

of Proposition 6.29 and the set

$$\bigcup_{d_- \in \mathcal{C}_{\mathfrak{n},\beta}} \Omega_{\Phi_{\mathfrak{n}}(d_-,0),\beta}$$

can be accurately computed by analytic and/or numerical methods. Thus, the full thermodynamic behavior of the (reduced) BCS Hamiltonian $\mathrm{H}_{\Lambda_\ell}^{BCS}$, as $\ell \to \infty$, can be completely determined. In particular, one can show for large temperatures, i.e., $\beta^{-1} \gg 1$, that

$$\mathcal{C}_{\mathfrak{n},\beta} = \{0\} \quad \text{and} \quad \bigcup_{d_- \in \mathcal{C}_{\mathfrak{n},\beta}} \Omega_{\Phi_{\mathfrak{n}}(d_-,0),\beta} = \Omega_{\Phi_{\mathfrak{n}}(0,0),\beta} = \Omega_{\Phi,\beta} = \left\{\omega_\beta\right\} = \Omega_{\mathfrak{n},\beta} \ ,$$

thanks to Corollary 6.32. Moreover, if $\eta > 0$ is sufficiently large (and fixed for all $\beta > 0$), then there is an inverse temperature β_c such that, for any $\beta > \beta_c$,

$$\mathcal{C}_{\mathfrak{n},\beta} = \left\{\sqrt{\eta r} \exp{(i\varphi)} : \varphi \in [0, 2\pi)\right\}$$

and

$$\bigcup_{d_- \in \mathcal{C}_{\mathfrak{n},\beta}} \Omega_{\Phi_{\mathfrak{n}}(d_-,0),\beta} = \left\{\omega_{\beta,\varphi} : \varphi \in [0, 2\pi)\right\}$$

for some positive number $r > 0$. As a consequence of the self-consistency condition (gap equation), $\omega_{\beta,\varphi_1} \neq \omega_{\beta,\varphi_2}$ for any $\varphi_1, \varphi_2 \in [0, 2\pi)$ with $\varphi_1 \neq \varphi_2$. This refers to the existence of a superconducting (first-order) phase transition at inverse temperature $\beta_c > 0$, with the breakdown of the gauge invariance. Additionally, the order parameter $r \geq 0$ can be shown to be directly related, at all temperatures, to the Cooper pair condensate density

$$r = \lim_{\ell \to \infty} \frac{1}{|\Lambda_\ell|} \omega_{\mathrm{H}_{\Lambda_\ell}^{BCS},\beta} \left(\mathfrak{c}_0^* \mathfrak{c}_0\right) = \lim_{\ell \to \infty} \frac{1}{|\Lambda_\ell|} \frac{\mathrm{Tr}\left(\mathfrak{c}_0^* \mathfrak{c}_0 \exp(-\beta \mathrm{H}_{\Lambda_\ell}^{BCS})\right)}{\mathrm{Tr}\left(\exp(-\beta \mathrm{H}_{\Lambda_\ell}^{BCS})\right)} \ ,$$

where

$$\mathfrak{c}_0 \doteq \frac{1}{\sqrt{|\Lambda_\ell|}} \sum_{x \in \Lambda_\ell} a_{x,\downarrow} a_{x,\uparrow} = \frac{1}{\sqrt{|\Lambda_\ell|}} \sum_{k \in \Lambda_\ell^*} \hat{a}_{k,\downarrow} \hat{a}_{-k,\uparrow}$$

annihilates one Cooper pair within the condensate, i.e., in the zero mode for electron pairs. The adjoint operator $\mathfrak{c}_0^*$ creates such a pair. Here, $\omega_{\mathrm{H}_{\Lambda_\ell}^{BCS},\beta}$ is the Gibbs state of Definition 5.19 associated with the BCS Hamiltonian $\mathrm{H}_{\Lambda_\ell}^{BCS}$. For more details, we recommend [26].

6.7 Thermodynamic Game

In [1], we generalize the results presented in Sect. 6.6 to all mean-field models of the Banach space $\mathcal{M}_1$. In the current subsection, we explain the main lines of this result. As mentioned above, we use the viewpoint of game theory, via the "thermodynamic game," that we now define precisely. First, recall that $h_{\mathfrak{m},\beta} : \ell^2(\mathbb{N}) \times \ell^2(\mathbb{N}) \to \mathbb{R}$ is the approximating free energy density defined by

$$h_{\mathfrak{m},\beta}(c_-, c_+) \doteq -\|c_+\|_2^2 + \|c_-\|_2^2 + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}(E_1)$$

for any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$. See Definitions 6.26 and 6.28. Given $\beta \in (0, \infty)$ and $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$, the thermodynamic game associated with the mean-field model $\mathfrak{m}$ is then the two-person zero-sum game whose payoff function is the approximating free energy density $h_{\mathfrak{m},\beta}$:

(i) The two players are denoted by $(-)$ and $(+)$. In fact, we interpret the mean-field attractions Ψ_- and repulsions Ψ_+ of the model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+)$ as two players that we, respectively, call the attractive and the repulsive player.

(ii) The sets of strategies of the attractive and repulsive player are, respectively, the following subspaces of $\ell^2(\mathbb{N})$:

$$\ell_-^2 \doteq \{c_- \in \ell^2(\mathbb{N}) \ : \ \text{for all } n \in \mathbb{N}, c_{-,n} = 0 \text{ if } \Psi_{-,n} = 0\} ,$$
$$\ell_+^2 \doteq \{c_+ \in \ell^2(\mathbb{N}) \ : \ \text{for all } n \in \mathbb{N}, c_{+,n} = 0 \text{ if } \Psi_{+,n} = 0\} .$$

(iii) The value $h_{\mathfrak{m},\beta}(c_-, c_+) \in \mathbb{R}$ is the loss of the player $(-)$ for the (attractive) strategy $c_- \in \ell_-^2$ and the gain of the second for the (repulsive) strategy $c_+ \in \ell_+^2$:

$(-)$ Without exchange of information, by minimizing

$$h_{\mathfrak{m},\beta}^{\sharp}(c_-) \doteq \sup_{c_+ \in \ell_+^2} h_{\mathfrak{m},\beta}(c_-, c_+) ,$$

the player $(-)$ obtains her/his least maximum loss

$$\mathrm{F}_{\mathfrak{m},\beta}^{\sharp} \doteq \inf_{c_- \in \ell_-^2} h_{\mathfrak{m},\beta}^{\sharp}(c_-) .$$

$(+)$ By maximizing

$$h_{\mathfrak{m},\beta}^{\flat}(c_+) \doteq \inf_{c_- \in \ell_-^2} h_{\mathfrak{m},\beta}(c_-, c_+) ,$$

the player $(+)$ obtains her/his greatest minimum gain

$$\mathrm{F}_{\mathfrak{m},\beta}^{\flat} \doteq \sup_{c_+ \in \ell_+^2} h_{\mathfrak{m},\beta}^{\flat}(c_+) \leq \mathrm{F}_{\mathfrak{m},\beta}^{\sharp} .$$

$\mathrm{F}_{\mathfrak{m},\beta}^{\flat}$ and $\mathrm{F}_{\mathfrak{m},\beta}^{\sharp}$ are called the "conservative values" of the thermodynamic game, while $[\mathrm{F}_{\mathfrak{m},\beta}^{\flat}, \mathrm{F}_{\mathfrak{m},\beta}^{\sharp}]$ is its "duality interval." Observe that, in general, $\mathrm{F}_{\mathfrak{m},\beta}^{\flat} < \mathrm{F}_{\mathfrak{m},\beta}^{\sharp}$. That is, the thermodynamic game may not admit a "cooperative equilibrium," which is, by definition, any saddle point of the payoff function $h_{\mathfrak{m},\beta}$. See [1, p. 42].

(iv) The corresponding sets of "conservative strategies" are

$$\begin{aligned} \mathcal{C}_{\mathfrak{m},\beta}^{\flat} &\doteq \left\{ d_+ \in \ell_+^2 : \mathrm{F}_{\mathfrak{m},\beta}^{\flat} = h_{\mathfrak{m},\beta}^{\flat}(d_+) \right\} , \\ \mathcal{C}_{\mathfrak{m},\beta}^{\sharp} &\doteq \left\{ d_- \in \ell_-^2 : \mathrm{F}_{\mathfrak{m},\beta}^{\sharp} = h_{\mathfrak{m},\beta}^{\sharp}(d_-) \right\} . \end{aligned} \tag{6.28}$$

In the particular case of a purely repulsive mean-field model, i.e., when $\Psi_- = 0$, $\mathcal{C}_{\mathfrak{m},\beta}^{\sharp} = \{0\}$, just because $\ell_-^2 = \{0\}$. Similarly, if $\Psi_+ = 0$, then $\mathcal{C}_{\mathfrak{m},\beta}^{\flat} = \{0\}$. In both cases ($\Psi_- = 0$ or $\Psi_+ = 0$), we have

$$\mathrm{F}_{\mathfrak{m},\beta}^{\flat} = \mathrm{F}_{\mathfrak{m},\beta}^{\sharp} = -\mathfrak{p}_{\beta}(\mathfrak{m}) . \tag{6.29}$$

See Proposition 6.29 for the purely attractive case. For a justification of this equality in the purely repulsive case, see the proof of Theorem 6.34 below. By [1, Lemma 8.4], the sets of conservatives strategies have the following important properties:

Proposition 6.33 *For any mean-field model* $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0, \infty)$, *the sets of conservatives strategies have the following properties:*

($\flat$) $\mathcal{C}_{\mathfrak{m},\beta}^{\flat} \subseteq \ell_+^2 \subseteq \ell^2(\mathbb{N})$ *has exactly one element* d_+.

($\sharp$) $\mathcal{C}_{\mathfrak{m},\beta}^{\sharp} \subseteq \ell_-^2 \subseteq \ell^2(\mathbb{N})$ *is nonempty and norm-bounded.*

The relevance of the thermodynamic game results from the fact that the conservative values $\mathrm{F}_{\mathfrak{m},\beta}^{\flat}$ and $\mathrm{F}_{\mathfrak{m},\beta}^{\sharp}$ of the game can be written as *variational problems over states*, corresponding in particular to the pressure function (Definition 6.14). This refers to a generalization of Proposition 6.29 to all (not necessarily purely attractive) mean-field models. To state the assertions, we recall two free energy functionals associated with mean-field models $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ at a given inverse temperature $\beta \in (0, \infty)$:

- By Definition 6.12, the usual free energy density functional $\mathfrak{f}_{\mathfrak{m},\beta} : E_1 \to \mathbb{R}$ is defined by

$$\mathfrak{f}_{\mathfrak{m},\beta} = \Delta_{\Psi_+} - \Delta_{\Psi_-} + \mathfrak{f}_{\Phi,\beta} .$$

- In the proof of Lemma 6.24, Eq. (6.14), we introduce also a non-conventional free energy density functional $\mathfrak{f}_{\mathfrak{m},\beta}^{\flat} : E_1 \to \mathbb{R}$, defined by

$$\mathfrak{f}_{\mathfrak{m},\beta}^{\flat}(\rho) \doteq \|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2 - \Delta_{\Psi_-}(\rho) + \mathfrak{f}_{\Phi,\beta}(\rho) \ . \tag{6.30}$$

Note that $\mathfrak{f}_{\mathfrak{m},\beta}^{\flat} \leq \mathfrak{f}_{\mathfrak{m},\beta}$, by Theorem 6.11 (v). We are now in a position to give the main statement of this subsection:

Theorem 6.34 *For any mean-field model* $\mathfrak{m} \in \mathcal{M}_1$ *and inverse temperature* $\beta \in (0,\infty)$*, the conservative values equal:*

$$\mathrm{F}_{\mathfrak{m},\beta}^{\flat} \doteq \sup_{c_+ \in \ell_+^2} \inf_{c_- \in \ell_-^2} h_{\mathfrak{m},\beta}(c_-, c_+) = \inf \mathfrak{f}_{\mathfrak{m},\beta}^{\flat}(E_1) \ ,$$

$$\mathrm{F}_{\mathfrak{m},\beta}^{\sharp} \doteq \inf_{c_- \in \ell_-^2} \sup_{c_+ \in \ell_+^2} h_{\mathfrak{m},\beta}(c_-, c_+) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) \ .$$

Idea of Proof The complete proof of this theorem can be found in [1]. See in particular [1, Theorem 2.36]. This is done in a similar way as in Proposition 6.29. The main issue now is that the infimum and supremum defining $\mathrm{F}_{\mathfrak{m},\beta}^{\flat}$ and $\mathrm{F}_{\mathfrak{m},\beta}^{\sharp}$ do not generally commute with each other. In fact, as already remarked above, one has in general that $\mathrm{F}_{\mathfrak{m},\beta}^{\flat}$ is strictly smaller than $\mathrm{F}_{\mathfrak{m},\beta}^{\sharp}$. To circumvent this problem, we proceed as follows: Note from Proposition 6.29 that

$$\begin{aligned}\mathrm{F}_{\mathfrak{m},\beta}^{\flat} &= \sup_{c_+ \in \ell_+^2} \inf_{c_- \in \ell_-^2} \inf_{\rho \in E_1} \left\{ -\|c_+\|_2^2 + \|c_-\|_2^2 + \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}(\rho) \right\} \\ &= \sup_{c_+ \in \ell_+^2} \inf_{\rho \in E_1} \left\{ -\|c_+\|_2^2 + \mathfrak{f}_{(\Phi_{\mathfrak{m}}(0,c_+),\Psi_-,0),\beta}(\rho) \right\} \ .\end{aligned}$$

Now, by the von Neumann min-max theorem [1, Theorem 10.50], the new functional

$$(c_+, \rho) \mapsto -\|c_+\|_2^2 + \mathfrak{f}_{(\Phi_{\mathfrak{m}}(0,c_+),\Psi_-,0),\beta}(\rho)$$

on $\ell_+^2 \times E_1$ has a saddle point and the infimum and supremum in the last equality can be interchanged. Doing this, one computes that

$$\mathrm{F}_{\mathfrak{m},\beta}^{\flat} = \inf \mathfrak{f}_{\mathfrak{m},\beta}^{\flat}(E_1) \ . \tag{6.31}$$

Note that by combining this equality with (6.17), one proves the identity (6.29) for the purely repulsive case. To prove the second part of the theorem, i.e., the equality $\mathrm{F}_{\mathfrak{m},\beta}^{\sharp} = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1)$, the trick with the saddle point is not necessary anymore, because one can directly use (6.31) instead: In fact, observe that (6.31) yields

$$\sup_{c_+\in\ell_+^2} h_{\mathfrak{m},\beta}(c_-,c_+) = \|c_-\|_2^2 + \inf_{\rho\in E_1}\{\|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2 + \mathfrak{f}_{\Phi_{\mathfrak{m}}(c_-,0),\beta}(\rho)\}\ .$$

Thus, by Lemma 6.27 combined with (6.17),

$$\begin{aligned}\inf_{c_-\in\ell_-^2}\sup_{c_+\in\ell_+^2} h_{\mathfrak{m},\beta}(c_-,c_+) &= \inf_{\rho\in E_1}\{\|\mathfrak{e}_{\Psi_+}(\rho)\|_2^2 - \|\mathfrak{e}_{\Psi_-}(\rho)\|_2^2 + \mathfrak{f}_{\Phi,\beta}(\rho)\}\\ &= \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1)\ .\end{aligned}$$

Compare this last argument with the proof of Proposition 6.29. □

By Definition 6.14 and Theorem 6.34, note that the pressure of any mean-field model $\mathfrak{m}\in\mathcal{M}_1$ is equal to

$$\begin{aligned}\mathfrak{p}_\beta(\mathfrak{m}) \doteq -\inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1) &= -\inf_{c_-\in\ell^2(\mathbb{N})}\sup_{c_+\in\ell^2(\mathbb{N})} h_{\mathfrak{m},\beta}(c_-,c_+)\\ &= -\inf_{c_-\in\ell_-^2}\sup_{c_+\in\ell_+^2} h_{\mathfrak{m},\beta}(c_-,c_+)\ .\end{aligned}$$

Recall that the infimum and supremum in this expression do not commute in general. A sufficient condition for them to commute is given through Sion's minimax theorem [27] as follows:

Lemma 6.35 *Let $\beta\in(0,\infty)$ and $\mathfrak{m}\doteq(\Phi,\Psi_-,\Psi_+)\in\mathcal{M}_1$ be any mean-field model such that $\Psi_-\neq 0$ and $\Psi_+\neq 0$. If, for any fixed $c_+\in\ell^2(\mathbb{N})$, the function $h_{\mathfrak{m},\beta}(\cdot,c_+)$ on $\ell^2(\mathbb{N})$ is quasi-convex, i.e., for all $r\in\mathbb{R}$, the level set*

$$\left\{c_-\in\ell^2(\mathbb{N}) : h_{\mathfrak{m},\beta}(c_-,c_+)\leq r\right\}$$

is convex, then $\mathrm{F}^\sharp_{\mathfrak{m},\beta} = \mathrm{F}^\flat_{\mathfrak{m},\beta}$.

Proof [28, Lemma 4.2]. □

To conclude, a result like Theorem 6.34 justifies on the level of thermodynamic functions the replacement of specific operators appearing in the Hamiltonian of a given physical system by constants which are determined as solutions to some self-consistency equation or some associated variational problem. This refers to the Bogoliubov approximation, which was used for (purely attractive mean-field) Fermi systems on lattices, already in 1957, to derive the celebrated Bardeen-Cooper-Schrieffer (BCS) theory for conventional type I superconductors [29–31]. The authors were of course inspired by Bogoliubov and his revolutionary paper [32]. A rigorous justification of this theory was given on the level of ground states by Bogoliubov in 1960 [33]. Then a method for analyzing the Bogoliubov approximation in a systematic way—on the level of the pressure—like in Theorem 6.34 with both mean-field repulsions and attractions was introduced by Bogoliubov Jr. in 1966 [34, 35] and by Brankov, Kurbatov, Tonchev, and Zagrebnov during the 1970s and 1980s [36–38]. This method is known in the literature as the *approximating Hamiltonian method* and leads—on the class of Hamiltonians it applies—to a rigorous proof of the exactness of the Bogoliubov approximation on the level of the pressure, provided it is done in an appropriated manner. Note however that the conditions on model imposed by [36–38] are still much more restrictive than those of Theorem 6.34. See discussions in [1, Section 2.10].

6.8 Self-Consistency of Equilibrium States

In Sect. 6.7, we introduce the thermodynamic game, which provides an efficient method to study phase transitions driven by mean-field interactions. It refers to a two-person zero-sum game whose payoff functions is defined as being the approximating free energy density functional of Definition 6.28. By Theorem 6.34, the conservative values of this game are directly related with variational problems over invariant states, naturally associated with any mean-field model. In fact, as we have seen, the largest of both conservative values is nothing else than the conventional pressure.

It turns out that, like in the special case of purely attractive mean-field models (cf. Corollary 6.32), the thermodynamic game also provides a complete characterization of the set of globally stable equilibrium states (Definition 6.15) of mean-field models, as follows:

Recall that the set of globally stable equilibrium states refers to

$$\Omega_{\mathfrak{m},\beta} \doteq \Big\{\omega \in E_1 : \exists (\rho_n)_{n\in\mathbb{N}} \subseteq E_1 \text{ weak}^* \text{ converging to } \omega \\ \text{so that } \lim_{n\to\infty} \mathfrak{f}_{\mathfrak{m},\beta}(\rho_n) = \inf \mathfrak{f}_{\mathfrak{m},\beta}(E_1)\Big\}$$

for any mean-field model $\mathfrak{m} \in \mathcal{M}_1$ and inverse temperature $\beta \in (0, \infty)$. Having in mind the second variational problem of Theorem 6.34, we also define the set

$$\Omega^\flat_{\mathfrak{m},\beta} \doteq \Big\{\omega \in E_1 : \mathfrak{f}^\flat_{\mathfrak{m},\beta}(\omega) = \inf \mathfrak{f}^\flat_{\mathfrak{m},\beta}(E_1)\Big\}$$

of non-conventional (globally stable) equilibrium states. Note that $\mathfrak{f}^\flat_{\mathfrak{m}}$ is weak*-lower semicontinuous but only convex (and not affine). In particular, $\Omega^\flat_{\mathfrak{m},\beta}$ is a nonempty weak*-compact convex subset of E_1.

In [1, Lemma 8.3 (♯)], it is proven that, for any $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ with $\Psi_+ \neq 0$, and all functions $c_- \in \ell^2(\mathbb{N})$, the set

$$\left\{d_+ \in \ell^2_+ : \max_{c_+\in\ell^2_+} h_{\mathfrak{m},\beta}(c_-, c_+) = h_{\mathfrak{m},\beta}(c_-, d_+)\right\} \tag{6.32}$$

has exactly one element, which is denoted by $\mathrm{r}_+(c_-)$. By [1, Lemma 8.8], if $\Psi_+ \neq 0$, then the mapping

$$\mathrm{r}_+ : c_- \mapsto \mathrm{r}_+ (c_-) \tag{6.33}$$

defines a continuous functional from ℓ^2_- to ℓ^2_+ itself, i.e., from the set of attractive strategies to the set of repulsive strategies of the mean-field model $\mathfrak{m}$. We call this mapping "the thermodynamic decision rule" of the mean-field model $\mathfrak{m} \in \mathcal{M}_1$. Note that in the particular case of purely attractive mean-field models (i.e., when $\Psi_+ = 0$ and $h_{\mathfrak{m},\beta}$ is thus not depending on c_+), one has $\mathrm{r}_+ = 0$.

For any mean-field model $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$, it is convenient to introduce a family of approximating purely attractive mean-field models by

$$\mathfrak{m}\,(c_+) \doteq (\Phi_{\mathfrak{m}}\,(0, c^+)\,, \Psi_-, 0) \in \mathcal{M}_1\ , \qquad c_+ \in \ell^2(\mathbb{N})\ . \tag{6.34}$$

Then, for every pair of strategies $c_- \in \ell^2_-$, $c_+ \in \ell^2_+$, we define the (possibly empty) sets

$$\Omega_{\mathfrak{m},\beta}\,(c_-, c_+) \doteq \left\{\omega \in \Omega_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta} : \mathfrak{e}_{\Psi_-}\,(\omega) = c_- \quad \text{and} \quad \mathfrak{e}_{\Psi_+}\,(\omega) = c_+\right\} \subseteq E_1 \tag{6.35}$$

as well as

$$\Omega_{\mathfrak{m},\beta}\,(c_+) \doteq \left\{\omega \in \Omega_{\mathfrak{m}(c_+),\beta} : \mathfrak{e}_{\Psi_+}\,(\omega) = c_+\right\} \subseteq E_1\ , \tag{6.36}$$

where, for any fixed invariant state $\rho \in E_1$ and $\Psi \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$,

$$\mathfrak{e}_\Psi\,(\rho) \doteq (\mathfrak{e}_{\Psi_n}\,(\rho))_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}) \qquad \text{with} \qquad \mathfrak{e}_{\Psi_n}\,(\rho) \doteq \rho\left(e_{\Psi_n}\right)$$

for all $n \in \mathbb{N}$. By Lemma 6.30, note that $\Omega_{\Phi_{\mathfrak{m}}(c_-,c_+),\beta}$ and $\Omega_{\mathfrak{m}(c_+),\beta}$ are (nonempty) weak*-closed faces of E_1, since $\mathfrak{m}\,(c_+)$ is a purely attractive mean-field model. Then, we obtain a (static) self-consistency condition for (conventional or non-conventional) globally state equilibrium states, which refers, in a sense, to Euler-Lagrange equations for the variational problem defining the thermodynamic game. More precisely, we have the following statements:

Theorem 6.36 *For any mean-field model* $\mathfrak{m} \doteq (\Phi, \Psi_-, \Psi_+) \in \mathcal{M}_1$ *and fixed inverse temperature* $\beta \in (0, \infty)$*, the following properties hold true:*

(i)

$$\Omega_{\mathfrak{m},\beta} = \overline{\mathrm{co}\left(\bigcup_{d_- \in \mathcal{C}^{\sharp}_{\mathfrak{m},\beta}} \Omega_{\mathfrak{m},\beta}\left(d_-, \mathrm{r}_+(d_-)\right) \right)} .$$

(ii) *The set $\mathcal{E}(\Omega_{\mathfrak{m},\beta})$ of extreme points of the weak*-compact convex set $\Omega_{\mathfrak{m},\beta}$ is included in the union of the sets*

$$\mathcal{E}\left(\Omega_{\mathfrak{m},\beta}\left(d_-, \mathrm{r}_+(d_-)\right)\right) , \qquad d_- \in \mathcal{C}^{\sharp}_{\mathfrak{m},\beta} ,$$

of all extreme points of $\Omega_{\mathfrak{m},\beta}(d_-, \mathrm{r}_+(d_-))$, $d_- \in \mathcal{C}^{\sharp}_{\mathfrak{m},\beta}$, which are nonempty, convex, mutually disjoint, weak-closed subsets of E_1.*

(iii) *When $\Psi_+ \neq 0$,*

$$\mathcal{C}^{\flat}_{\mathfrak{m},\beta} = \{d_+\} \qquad \textit{and} \qquad \Omega^{\flat}_{\mathfrak{m},\beta} = \Omega_{\mathfrak{m},\beta}(d_+) .$$

Proof Assertion (i) results from [1, Theorem 2.21 (i)] and [1, Theorem 2.39 (i)], while (ii) corresponds to [1, Theorem 2.39 (ii)]. As already mentioned, the fact that $\mathcal{C}^{\flat}_{\mathfrak{m},\beta} = \{d_+\}$ refers to Proposition 6.33. However, the identity $\Omega^{\flat}_{\mathfrak{m},\beta} = \Omega_{\mathfrak{m},\beta}(d_+)$ was not considered in [1], but its proof is similar to the one of [1, Lemma 9.2]. See [28, Theorem 4.3]. For more details, see also Theorem 7.346 and discussions before and after this theorem, which explain in a general context the strategy of proof used here. □

Theorem 6.36 implies in particular that, for any extreme state $\hat{\omega} \in \mathcal{E}(\Omega_{\mathfrak{m},\beta})$ of $\Omega_{\mathfrak{m},\beta}$, there is a unique $d_- \in \mathcal{C}^{\sharp}_{\mathfrak{m},\beta}$ such that

$$\mathfrak{e}_{\Psi_-}(\omega) = d_- \quad \text{and} \quad \mathfrak{e}_{\Psi_+}(\omega) = \mathrm{r}_+(d_-) . \tag{6.37}$$

In the physics literature on superconductors, recall that the above equality refers to the so-called gap equations. Conversely, for any $d_- \in \mathcal{C}^{\sharp}_{\mathfrak{m},\beta}$, there is some generalized equilibrium state ω satisfying the condition above, but ω is not necessarily an extreme point of $\Omega_{\mathfrak{m},\beta}$.

To conclude, note that Theorem 6.36 yields the equality $\Omega_{\mathfrak{m},\beta} = \Omega^{\flat}_{\mathfrak{m},\beta}$ for any purely repulsive or purely attractive mean-field model $\mathfrak{m} \in \mathcal{M}_1$. However, for mean-field models $\mathfrak{m} \in \mathcal{M}_1$ with both non-trivial attractive and repulsive mean-field interactions, there is no reason for this equality to hold true, in general.

6.9 From Short-Range to Mean-Field Models

Realistic effective interparticle interactions of quantum many-body systems are widely seen as being short-range, not mean-field. However, the rigorous mathematical analysis of phase diagrams of short-range model turns out to be extremely difficult, in general, with many important fundamental questions remaining open still nowadays. By contrast, mean-field models come from different approximations or Ansätze, and are thus less realistic, in a sense, but are technically advantageous, while capturing surprisingly well many real physical phenomena. Indeed, the study of phase diagrams of mean-field models can be performed by self-consistency equations related to the associated thermodynamic game. This is illustrated at the end of Sect. 6.6 for the BCS theory of superconductivity.

Here, we discuss a precise mathematical relation between mean-field and short-range models, by using the long-range limit that is known in the literature as the Kac limit. This is done in [28] in an abstract, model-independent, way. To be more pedagogical, however, we restrict our discussions to a specific example. This gives us, additionally, the opportunity to illustrate results of previous sections, in particular those of Sects. 6.7–6.8, for a specific mean-field model having both positive and attractive mean-field terms.

6.9.1 The Short-Range Model

Like in Example 6.6, fix $\Omega \doteq \{\uparrow, \downarrow\}$ and $\mathcal{U} \doteq \mathrm{CAR}(\{\uparrow, \downarrow\}, \Gamma)$. $\{\mathrm{e}_{\mathrm{s},x}\}_{(\mathrm{s},x)\in\Omega\times\Gamma}$ is, as before, the (canonical) Hilbert basis of $\ell^2(\Omega \times \Gamma)$. We use the shorter notation $a_{x,\mathrm{s}} \doteq a\left(\mathrm{e}_{\mathrm{s},x}\right)$ for the "annihilation operator" of a fermion with spin $\mathrm{s} \in \Omega$ and lattice position $x \in \Gamma$. We consider fermions inside the cubic box

$$\Lambda_\ell \doteq \{(x_1, \ldots, x_d) \in \Gamma : |x_i| \leq \ell\}$$

for any $\ell \in \mathbb{N}$. Fix once and for all, in the present subsection, an invariant interaction $\Phi \in \mathcal{W}_1$. For two parameters $\gamma_-, \gamma_+ \in (0, 1)$ and the fixed invariant interaction $\Phi \in \mathcal{W}_1$, we define the local Hamiltonians

$$H_{\Lambda_\ell}(\gamma_-, \gamma_+) \doteq H_{\Lambda_\ell}^{\Phi} + \sum_{x,y\in\Lambda_\ell,\mathrm{s},\mathrm{t}\in\{\uparrow,\downarrow\}} \gamma_+^d \mathrm{v}_+ \left(\gamma_+ (x-y)\right) a_{y,\mathrm{t}}^* a_{y,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}} - \sum_{x,y\in\Lambda_\ell} \gamma_-^d \mathrm{v}_- \left(\gamma_- (x-y)\right) a_{y,\uparrow}^* a_{y,\downarrow}^* a_{x,\downarrow} a_{x,\uparrow}. \tag{6.38}$$

Here, v_+ is a (non-zero) pair potential characterizing interparticle forces, whose range of action is tuned by the parameter $\gamma_+ \in (0, 1)$. The (non-zero) function v_- encodes the hopping strength of Cooper pairs. The corresponding term of the

Hamiltonian thus implements a BCS-type interaction whose range is tuned by the parameter $\gamma_- \in (0,1)$.

As is usual in theoretical physics, $\mathrm{v}_-, \mathrm{v}_+$ are assumed to be fast decaying, reflection-symmetric,[5] and positive definite, i.e., the Fourier transform $\hat{\mathrm{v}}_-, \hat{\mathrm{v}}_+$ of v are positive functions on $\mathbb{R}^d$. This choice for v_+ is reminiscent of a superstability condition, which is essential in the bosonic case [39, Section 2.2 and Appendix G]. For simplicity, we assume that $\mathrm{v}_-, \mathrm{v}_+ \in C_0^{2d}\left(\mathbb{R}^d, \mathbb{R}\right)$ are both compactly supported. Because of some technical issues, we also assume that

$$\hat{\mathrm{v}}_-(\gamma^{-1}k) \leq \hat{\mathrm{v}}_-(k)\ , \qquad k \in \mathbb{R}^d\ ,\ \gamma \in (0,1)\ .$$

The definition of the Fourier transform of a function v we used here is

$$\hat{\mathrm{v}}(k) \doteq \int_{\mathbb{R}^d} \mathrm{v}(x)\, \mathrm{e}^{-ik\cdot x} \mathrm{d}^d x\ , \qquad k \in \mathbb{R}^d\ . \tag{6.39}$$

Observe that the sequence of local Hamiltonians $H_{\Lambda_\ell}(\gamma_-, \gamma_+)$, $\ell \in \mathbb{N}$, is the one associated with the invariant interaction:

$$\Phi(\gamma_-, \gamma_+) \doteq \Phi + \Psi_{\mathrm{v}_+,\gamma_+} - \Psi_{\mathrm{v}_-,\gamma_-} \in \mathcal{W}_1\ ,$$

where the invariant interactions $\Psi_{\mathrm{v}_-,\gamma_-}, \Psi_{\mathrm{v}_+,\gamma_+} \in \mathcal{W}_1$ are defined by

$$\Psi_{\mathrm{v}_-,\gamma_-}(\Lambda) \doteq 0 \doteq \Psi_{\mathrm{v}_+,\gamma_+}(\Lambda)$$

whenever $|\Lambda| > 2$, while, for any $x, y \in \Gamma$,

$$\Psi_{\mathrm{v}_+,\gamma_+}(\{x,y\}) \doteq \left(2 - \delta_{x,y}\right) \sum_{\mathrm{s,t} \in \{\uparrow,\downarrow\}} \gamma_+^d \mathrm{v}_+(\gamma_+(x-y))\, a_{y,\mathrm{t}}^* a_{y,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}\ ,$$

$$\Psi_{\mathrm{v}_-,\gamma_-}(\{x,y\}) \doteq \left(2 - \delta_{x,y}\right) \gamma_-^d \mathrm{v}_-(\gamma_-(x-y))\, a_{y,\uparrow}^* a_{y,\downarrow}^* a_{x,\downarrow} a_{x,\uparrow}\ ,$$

$\delta_{x,y}$ being the Kronecker delta. Using these definitions, we have

$$H_{\Lambda_\ell}(\gamma_-, \gamma_+) = H_{\Lambda_\ell}^{\Phi(\gamma_-,\gamma_+)}$$

for all natural numbers $\ell \in \mathbb{N}$ and $\gamma_-, \gamma_+ \in (0,1)$. Therefore, we can apply to $\Phi(\gamma_-, \gamma_+)$ all the above results on the thermodynamic behavior of models of $\mathcal{W}_1 \subseteq \mathcal{M}_1$.

For instance, for all parameters $\gamma_-, \gamma_+ \in (0,1)$, the energy density functional

$$\mathfrak{e}_{\Phi(\gamma_-,\gamma_+)} : E_1 \to \mathbb{R}$$

[5] That is, $\mathrm{v}_\pm(x) = \mathrm{v}_\pm(-x)$. Usually, $\mathrm{v}_\pm(x) = v_\pm(|x|)$ for some function $v_\pm : \mathbb{R}_0^+ \to \mathbb{R}$.

associated with the invariant interaction $\Phi(\gamma_-, \gamma_+) \in \mathcal{W}_1$ is defined by

$$\mathfrak{e}_{\Phi(\gamma_-,\gamma_+)}(\rho) \doteq \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|}\rho\left(H_{\Lambda_\ell}(\gamma_-, \gamma_+)\right)$$

for any invariant state $\rho \in E_1$. See Proposition 5.11 and Definition 5.12. It naturally splits into three components:

$$\mathfrak{e}_{\Phi(\gamma_-,\gamma_+)} = \underbrace{\mathfrak{e}_\Phi}_{\text{free term}} + \underbrace{\mathfrak{e}_{\Psi_{\mathrm{v}_+,\gamma_+}}}_{\text{interaction term }+} - \underbrace{\mathfrak{e}_{\Psi_{\mathrm{v}_-,\gamma_-}}}_{\text{interaction term }-} .$$

With this, for any inverse temperature $\beta \in (0, \infty)$ and $\gamma_-, \gamma_+ \in (0, 1)$, the free energy density functional $\mathfrak{f}_{\Phi(\gamma_-,\gamma_+),\beta} : E_1 \to \mathbb{R}$ of Definition 6.12 equals

$$\mathfrak{f}_{\Phi(\gamma_-,\gamma_+),\beta} \doteq \mathfrak{e}_{\Phi(\gamma_-,\gamma_+)} - \beta^{-1}\mathfrak{s} = \mathfrak{e}_{\Psi_{\mathrm{v}_+,\gamma_+}} - \mathfrak{e}_{\Psi_{\mathrm{v}_-,\gamma_-}} + \mathfrak{f}_{\Phi,\beta}, \tag{6.40}$$

where $\mathfrak{s} : E_1 \to \mathbb{R}_0^+$ is the entropy density functional of Theorem 5.20. By Theorem 6.13, the thermodynamic limit of the (grand-canonical) pressure equals

$$P_\beta(\gamma_-, \gamma_+) \doteq \lim_{\ell\to\infty} P_{H_{\Lambda_\ell}(\gamma_-,\gamma_+),\beta} = -\inf \mathfrak{f}_{\Phi(\gamma_-,\gamma_+),\beta}(E_1) < \infty \tag{6.41}$$

for $\beta \in (0, \infty)$ and $\gamma_-, \gamma_+ \in (0, 1)$. See also (6.11). Recall that the globally stable equilibrium states of the short-range model are, by definition, the solutions to this variational problem. They form the set

$$\Omega_{\Phi(\gamma_-,\gamma_+),\beta} \doteq \left\{\omega \in E_1 : \mathfrak{f}_{\Phi(\gamma_-,\gamma_+),\beta}(\omega) = -P_\beta(\gamma_-, \gamma_+)\right\}$$

for any fixed $\beta \in (0, \infty)$ and $\gamma_-, \gamma_+ \in (0, 1)$. By Lemma 6.30, it is a (nonempty) weak*-closed face of the convex weak*-compact space E_1 of invariant states.

6.9.2 The Mean-Field Model

The Kac, or long-range, limits refer here to the limits $\gamma_\pm \to 0^+$ of short-range models that are already in the thermodynamic limit. For small parameters $\gamma_\pm \ll 1$, the short-range model defined in finite volume by (6.38) has an interparticle (+) and BCS (−) interactions with very large range ($\mathcal{O}(\gamma_\pm^{-1})$), but the interaction strength is small as $\gamma_\pm^d$, in such a way that the first Born approximation[6] to the scattering length of the interparticle and BCS potentials remains constant, as is usual. One

[6] That is, $\int_{\mathbb{R}^d} \gamma_\pm^d \mathrm{v}_\pm(\gamma_\pm x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \mathrm{v}_\pm(x)\,\mathrm{d}x \doteq \hat{\mathrm{v}}_\pm(0)$.

therefore expects to have some effective mean-field, or long-range, model in the limits $\gamma_\pm \to 0^+$.

Given $\Phi \in \mathcal{W}_1$, the effective local Hamiltonians in the limits $\gamma_\pm \to 0^+$ of short-range models should be

$$H_{\Lambda_\ell}^\sharp(\eta_-, \eta_+) \doteq H_{\Lambda_\ell}^\Phi + \underbrace{\frac{\eta_+}{|\Lambda_\ell|} \sum_{x,y\in\Lambda_\ell, \mathrm{s},\mathrm{t}\in\{\uparrow,\downarrow\}} a_{y,\mathrm{t}}^* a_{y,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}}_{\text{mean-field repulsion }+} \underbrace{- \frac{\eta_-}{|\Lambda_\ell|} \sum_{x,y\in\Lambda_\ell} a_{y,\uparrow}^* a_{y,\downarrow}^* a_{x,\downarrow} a_{x,\uparrow}}_{\text{mean-field attraction }-} \tag{6.42}$$

for all natural numbers $\ell \in \mathbb{N}$ and some positive parameters $\eta_-, \eta_+ \in \mathbb{R}_0^+$. Compare this Hamiltonian with (6.38). It refers to the mean-field model

$$\mathfrak{m}(\eta_-, \eta_+) \doteq \left(\Phi, \eta_-^{1/2}\Psi_-, \eta_+^{1/2}\Psi_+\right) \in \mathcal{M}_1$$

where

$$\Psi_- \doteq (\Psi_{\mathrm{BCS}}, 0, \ldots), \Psi_+ \doteq (\Psi_{\mathrm{Int}}, 0, \ldots) \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$$

with $\Psi_{\mathrm{BCS}} \in \mathcal{W}_1^{\mathbb{C}}$ being the "BCS interaction" of Example 6.6 for $\eta = 1$, defined by $\Psi_{\mathrm{BCS}}(\Lambda) \doteq 0$ whenever $|\Lambda| \notin \{1\}$ and

$$\Psi_{\mathrm{BCS}}(\{x\}) \doteq a_{x,\downarrow} a_{x,\uparrow}$$

for all lattice sites $x \in \Gamma$, while $\Psi_{\mathrm{Int}} \in \mathcal{W}_1 \subseteq \mathcal{W}_1^{\mathbb{C}}$ is the invariant interaction defined by $\Psi_{\mathrm{Int}}(\Lambda) \doteq 0$ whenever $|\Lambda| \notin \{1\}$ and

$$\Psi_{\mathrm{Int}}(\{x\}) \doteq a_{x,\uparrow}^* a_{x,\uparrow} + a_{x,\downarrow}^* a_{x,\downarrow}$$

for all lattice sites $x \in \Gamma$.

We then apply to the mean-field model $\mathfrak{m}(\eta_-, \eta_+)$ the results obtained above for general elements of $\mathcal{M}_1$. For instance, the space-averaging functionals $\Delta_{\Psi_\pm} : E_1 \to \mathbb{R}$ associated with the above sequences $\Psi_-, \Psi_+ \in \ell^2(\mathbb{N}; \mathcal{W}_1^{\mathbb{C}})$ are equal to

$$\Delta_{\Psi_\pm}(\rho) = \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|^2} \sum_{x,y\in\Lambda_\ell} \rho\left(\tau_y\left(A_\pm^*\right)\tau_x\left(A_\pm\right)\right) \in \left[|\rho(A_\pm)|^2, \|A_\pm\|^2\right],$$

for any invariant state $\rho \in E_1$, where

$$A_- \doteq a_{0,\downarrow} a_{0,\uparrow} = e_{\Psi_{\mathrm{BCS}}} \qquad \text{and} \qquad A_+ \doteq a_{0,\uparrow}^* a_{0,\uparrow} + a_{0,\downarrow}^* a_{0,\downarrow} = e_{\Psi_{\mathrm{Int}}}.$$

See Eqs. (6.6) and (6.8) as well as Definitions 6.9 and 6.12. For any inverse temperature $\beta \in (0, \infty)$ and $\eta_-, \eta_+ \in \mathbb{R}_0^+$, the free energy density functional $\mathfrak{f}_{\mathfrak{m}(\eta_-,\eta_+),\beta} : E_1 \to \mathbb{R}$ of Definition 6.12 equals

$$\mathfrak{f}_{\mathfrak{m}(\eta_-,\eta_+),\beta} \doteq \eta_+ \Delta_{\Psi_+} - \eta_- \Delta_{\Psi_-} + \mathfrak{f}_{\Phi,\beta} \ . \tag{6.43}$$

By Theorem 6.13, the thermodynamic limit of the (grand-canonical) pressure equals

$$P_\beta^\sharp(\eta_-, \eta_+) \doteq \lim_{\ell \to \infty} P_{H_{\Lambda_\ell}^\sharp(\gamma_-,\gamma_+),\beta} = -\inf \mathfrak{f}_{\mathfrak{m}(\eta_-,\eta_+),\beta}(E_1) < \infty \tag{6.44}$$

for any $\beta \in (0, \infty)$ and $\eta_-, \eta_+ \in \mathbb{R}_0^+$. As before, the globally stable equilibrium states of the mean-field model are the limits of minimizing sequences for the functional $\mathfrak{f}_{\mathfrak{m}(\eta_-,\eta_+),\beta}$. They form the set

$$\Omega_{\mathfrak{m}(\eta_-,\eta_+),\beta} \doteq \left\{ \begin{array}{c} \omega \in E_1 : \exists (\rho_n)_{n\in\mathbb{N}} \subseteq E_1 \text{ weak}^* \text{ converging to } \omega \text{ so that} \\ \lim\limits_{n\to\infty} \mathfrak{f}_{\mathfrak{m}(\eta_-,\eta_+),\beta}(\rho_n) = -P_\beta^\sharp(\eta_-, \eta_+) \end{array} \right\}$$

for $\beta \in (0, \infty)$ and $\eta_-, \eta_+ \in \mathbb{R}_0^+$. By Lemma 6.22, it is a (nonempty) convex weak*-compact subspace of the space E_1 of invariant states.

6.9.3 Thermodynamic Game and Bogoliubov Approximation

A mathematically rigorous computation of the pressure and equilibrium states of the short-range model to show possible phase transitions is elusive, beyond perturbative arguments, even after decades of mathematical studies. By contrast, such a question can be solved for the mean-field model. This is done by using the thermodynamic game explained in Sect. 6.7.

In this case, the approximating interactions of the mean-field model $\mathfrak{m}(\eta_-, \eta_+)$ equal

$$\Phi_{\mathfrak{m}(\eta_-,\eta_+)}(c_-, c_+) \doteq \Phi + 2\left(\eta_+^{1/2} \operatorname{Re}\left\{\overline{c_{+,1}}\right\} \Psi_{\mathrm{Int}} - \eta_-^{1/2} \operatorname{Re}\left\{\overline{c_{-,1}} \Psi_{\mathrm{BCS}}\right\}\right) \in \mathcal{W}_1$$

for all sequences $c_-, c_+ \in \ell^2(\mathbb{N})$; see Definition 6.26. Note that the use of full sequences $c_-, c_+ \in \ell^2(\mathbb{N})$ is not necessary here since the model has only one non-zero attractive and repulsive mean-field part. In other words, both sets of attractive and repulsive strategies for the associated thermodynamic game are identified with the set of complex numbers: $c_- \equiv c_{-,1} \in \mathbb{C}$ and $c_+ \equiv c_{+,1} \in \mathbb{C}$. The approximating interaction of the mean-field model leads to the following sequence of local Hamiltonians

$$\tilde{H}_{\Lambda_\ell}(\eta_-, \eta_+, c_-, c_+) \doteq H^{\Phi}_{\Lambda_\ell} + \eta_+^{1/2}(\overline{c_+} + c_+) \sum_{x\in\Lambda_\ell, \mathrm{s}\in\{\uparrow,\downarrow\}} a^*_{x,\mathrm{s}} a_{x,\mathrm{s}}$$
$$+ \eta_-^{1/2} \sum_{x\in\Lambda_\ell} \left(\overline{c_-} a^*_{x,\uparrow} a^*_{x,\downarrow} + c_- a_{x,\downarrow} a_{x,\uparrow}\right) \qquad (6.45)$$

for any two complex numbers $c_-, c_+ \in \mathbb{C}$, natural numbers $\ell \in \mathbb{N}$, and some positive parameters $\eta_-, \eta_+ \in \mathbb{R}_0^+$. Then, by Theorem 6.34 and Eq. (6.44), the conservative values of the thermodynamic game equal

$$\mathrm{F}^\sharp_{\mathfrak{m}(\eta_-,\eta_+),\beta} \doteq \inf_{c_-\in\mathbb{C}} \sup_{c_+\in\mathbb{C}} \left\{- |c_+|^2 + |c_-|^2 - P_\beta(c_-, c_+, \eta_+, \eta_-)\right\}$$
$$= -P^\sharp_\beta(\eta_-, \eta_+) \qquad (6.46)$$

and

$$\mathrm{F}^\flat_{\mathfrak{m}(\eta_-,\eta_+),\beta} \doteq \sup_{c_+\in\mathbb{C}} \inf_{c_-\in\mathbb{C}} \left\{- |c_+|^2 + |c_-|^2 - P_\beta(c_-, c_+, \eta_+, \eta_-)\right\}$$
$$= -P^\flat_\beta(\eta_-, \eta_+) \ . \qquad (6.47)$$

Here, we have the non-conventional pressure defined by

$$P^\flat_\beta(\eta_-, \eta_+) \doteq -\inf \mathfrak{f}^\flat_{\mathfrak{m}(\eta_-,\eta_+),\beta}(E_1) \qquad (6.48)$$

where, for any invariant state $\rho \in E_1$,

$$\mathfrak{f}^\flat_{\mathfrak{m}(\eta_-,\eta_+),\beta}(\rho) = \eta_+ \left|\rho(a^*_{0,\uparrow} a_{0,\uparrow} + a^*_{0,\downarrow} a_{0,\downarrow})\right|^2 - \eta_- \Delta_{\Psi_-}(\rho) + \mathfrak{f}_{\Phi,\beta}(\rho)$$
$$\leq \mathfrak{f}_{\mathfrak{m}(\eta_-,\eta_+),\beta}(\rho) \ ,$$

(see (6.30)), while $P_\beta : \mathbb{C}^2 \times \left(\mathbb{R}_0^+\right)^2 \to \mathbb{R}$ is the function defined by

$$P_\beta(c_-, c_+, \eta_+, \eta_-) = \lim_{\ell\to\infty} P_{\tilde{H}_{\Lambda_\ell}(\eta_-,\eta_+,c_-,c_+),\beta}$$
$$= -\inf \mathfrak{f}_{\Phi_{\mathfrak{m}(\eta_-,\eta_+)}(c_-,c_+)}(E_1) < \infty \ , \qquad (6.49)$$

thanks to Theorem 6.13. Note that a usual choice for the free interaction $\Phi \in \mathcal{W}_1$ is given by (6.27). In this case, the approximating Hamiltonians (6.45) are quadratic in the annihilation and creation operators. It can be exactly diagonalized, and the variational problems (6.46) and (6.47) can be analytically and numerically studied, in this case. The sets $\Omega_{\mathfrak{m}(\eta_-,\eta_+),\beta}$ and

$$\Omega^{\flat}_{\mathfrak{m}(\eta_-,\eta_+),\beta} \doteq \left\{\omega \in E_1 : \mathfrak{f}^{\flat}_{\mathfrak{m}(\eta_-,\eta_+),\beta}(\omega) = \inf \mathfrak{f}^{\flat}_{\mathfrak{m}(\eta_-,\eta_+),\beta}(E_1) = \mathrm{F}^{\flat}_{\mathfrak{m}(\eta_-,\eta_+),\beta}\right\}$$

of equilibrium states can also be explicitly determined, thanks to Theorem 6.36.

6.9.4 The Kac Limit

We now perform the Kac, or long-range, limits $\gamma_\pm \to 0^+$ of short-range models. First, using a cyclic representation of the C^*-algebra $\mathcal{U}$ induced by any invariant state (Theorem 4.113) as well as the spectral theorem, one can prove [28] that the energy densities associated with the invariant interactions $\Psi_{\mathrm{v}_-,\gamma_-}$ and $\Psi_{\mathrm{v}_+,\gamma_+}$ converge pointwise to

$$\lim_{\gamma_\pm\to 0^+} \mathfrak{e}_{\Psi_{\mathrm{v}_\pm,\gamma_\pm}}(\rho) \doteq \hat{\mathrm{v}}_\pm(0)\Delta_\pm(\rho) \tag{6.50}$$

for any invariant state $\rho \in E_1$, where we have from (6.39) that

$$\hat{\mathrm{v}}_\pm(0) \doteq \int_{\mathbb{R}^d} \mathrm{v}_\pm(x)\,\mathrm{d}^d x \geq 0 .$$

Recall that $\mathrm{v}_-, \mathrm{v}_+$ are assumed to be positive definite, i.e., the Fourier transforms $\hat{\mathrm{v}}_-, \hat{\mathrm{v}}_+$ of $\mathrm{v}_-, \mathrm{v}_+$, respectively, are positive functions on $\mathbb{R}^d$. Comparing (6.40)–(6.41) and (6.43)–(6.44) in light of (6.50), this suggests that the parameters $\eta_-, \eta_+ \in \mathbb{R}_0^+$ of the mean-field models to be taken in the limits $\gamma_\pm \to 0^+$ are

$$\eta_\pm = \hat{\mathrm{v}}_\pm(0) \in \mathbb{R}_0^+.$$

This is partially confirmed by [28, Theorem 5.15], which in the example presented here refers to the following theorem:

Theorem 6.37 *Let $\Phi \in \mathcal{W}_1$ and $\mathrm{v}_-, \mathrm{v}_+ \in C_0^{2d}\left(\mathbb{R}^d, \mathbb{R}\right)$ be reflection-symmetric, positive definite functions on $\mathbb{R}^d$ with $\hat{\mathrm{v}}_-(\gamma^{-1}k) \leq \hat{\mathrm{v}}_-(k)$ for $k \in \mathbb{R}^d$. Fix an inverse temperature $\beta \in (0,\infty)$.*

(i) *Convergence of infinite-volume pressures:*

$$\lim_{\gamma_+\to 0^+} \lim_{\gamma_-\to 0^+} P_\beta(\gamma_-,\gamma_+) = P_\beta^\sharp\left(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)\right) .$$

(ii) *Convergence of equilibrium states: For any $\gamma_+ \in (0,1)$, take any weak* accumulation point ω_{γ_+} of any net $(\omega_{\gamma_-,\gamma_+})_{\gamma_-\in(0,1)} \subseteq \Omega_{\Phi(\gamma_-,\gamma_+),\beta}$ as $\gamma_- \to 0^+$. Pick any weak* accumulation point ω of the net $(\omega_{\gamma_+})_{\gamma_+\in(0,1)}$, as $\gamma_+ \to 0^+$. Then,*

$$\omega_{\gamma_-,\gamma_+} \underset{weak^*,\gamma_-\to 0^+}{\to} \omega_{\gamma_+} \underset{weak^*,\gamma_+\to 0^+}{\to} \omega \in \Omega_{\mathfrak{m}(\hat{\mathrm{v}}_-(0),\hat{\mathrm{v}}_+(0)),\beta} \ .$$

This theorem demonstrates that the mean-field model is generally an idealization of short-range models in the long-range limit. In addition, [28] gives some explicit error estimates, and one can deduce approximated phase diagrams on short-range models for sufficiently small parameters $\gamma_\pm \in (0,1)$.

Note, however, that Theorem 6.37 uses a special order for the limit of small $\gamma_\pm \in (0,1)$: First $\gamma_- \to 0^+$ and then $\gamma_+ \to 0^+$. It means that the attractive forces have a much larger range than the one of repulsive forces. One can ask whether this is just a technical artifact. As a matter of fact, it is generally **not** so, and the hierarchy of ranges does have a strong effect on the equilibrium states and pressure of the model:

Proposition 6.38 *Let $\Phi \in \mathcal{W}_1$ and $\mathrm{v}_-, \mathrm{v}_+ \in C_0^{2d}(\mathbb{R}^d, \mathbb{R})$ be reflection-symmetric, positive definite functions on $\mathbb{R}^d$ with $\hat{\mathrm{v}}_-(\gamma^{-1}k) \leq \hat{\mathrm{v}}_-(k)$ for all $k \in \mathbb{R}^d$ and $\gamma \in (0,1)$. Fix $\beta \in (0,\infty)$. If $(\gamma_{-,n})_{n\in\mathbb{N}}$ and $(\gamma_{+,n})_{n\in\mathbb{N}}$ converges to zero, then*

$$P_\beta^\sharp\left(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)\right) \leq \liminf_{n\to\infty} P_\beta\left(\gamma_{+,n}, \gamma_{-,n}\right) \leq \limsup_{n\to\infty} P_\beta\left(\gamma_{+,n}, \gamma_{-,n}\right)$$
$$\leq P_\beta^\flat\left(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)\right) \ .$$

Proof See [28, Proposition 5.14]. □

Recall that the supremum and infimum in (6.46) and (6.47) do not commute, in general. See [1, p. 42]. A sufficient condition for them to commute is given by Lemma 6.35. Thus, we generally have

$$P_\beta^\sharp\left(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)\right) \neq P_\beta^\flat\left(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)\right)$$

and Proposition 6.38 suggests that the limits $\gamma_\pm \to 0^+$ of short-range models can lead to a different system from the one described by the *conventional* mean-field model, which is the thermodynamic limit the finite-volume system associated with the local Hamiltonians (6.42). In fact, applying [28, Theorem 5.17] to the model presented above, one can reach the here called "non-conventional mean-field model":

Theorem 6.39 *Let $\Phi \in \mathcal{W}_1$ and $\mathrm{v}_-, \mathrm{v}_+ \in C_0^{2d}(\mathbb{R}^d, \mathbb{R})$ be reflection-symmetric, positive definite functions on $\mathbb{R}^d$ with $\hat{\mathrm{v}}_-(\gamma^{-1}k) \leq \hat{\mathrm{v}}_-(k)$ for $k \in \mathbb{R}^d$. Fix an inverse temperature $\beta \in (0,\infty)$.*

(i) *Convergence of infinite-volume pressures:*

$$\lim_{\gamma_-\to 0^+} \lim_{\gamma_+\to 0^+} P_\beta(\gamma_-, \gamma_+) = P_\beta^\flat\left(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)\right) \ .$$

(ii) *Convergence of equilibrium states: For any $\gamma_- \in (0,1)$, take any weak* accumulation point ω_{γ_-} of any net $(\omega_{\gamma_-,\gamma_+})_{\gamma_+\in(0,1)} \subseteq \Omega_{\Phi(\gamma_-,\gamma_+),\beta}$ as $\gamma_+ \to 0^+$. Pick any weak* accumulation point ω of the net $(\omega_{\gamma_-})_{\gamma_-\in(0,1)}$, as $\gamma_- \to 0^+$.*

Then,

$$\omega_{\gamma_-,\gamma_+} \underset{weak^*,\gamma_+\to 0^+}{\to} \omega_{\gamma_-} \underset{weak^*,\gamma_-\to 0^+}{\to} \omega \in \Omega^{\flat}_{\mathfrak{m}(\hat{\mathrm{v}}_-(0),\hat{\mathrm{v}}_+(0)),\beta} \, .$$

As there is no reason to have the equality $\Omega^{\sharp}_{\mathfrak{m}} = \Omega^{\flat}_{\mathfrak{m}}$ for a given arbitrary mean-field model $\mathfrak{m} \in \mathcal{M}_1$, Theorems 6.37 and 6.39 generally describe different physical situations. In fact, one can even prove that the limit of Kac pressures can attain **all** the values of the duality interval

$$\mathrm{I} \doteq \left[P^{\sharp}_{\beta}(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)),\, P^{\flat}_{\beta}(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)) \right]$$

of the thermodynamic game associated with the mean-field model $\mathfrak{m}(\hat{\mathrm{v}}_-(0), \hat{\mathrm{v}}_+(0)) \in \mathcal{M}_1$:

Theorem 6.40 *Let* $\Phi \in \mathcal{W}_1$ *and* $\mathrm{v}_-, \mathrm{v}_+ \in C_0^{2d}\left(\mathbb{R}^d, \mathbb{R}\right)$ *be reflection-symmetric, positive definite functions on* $\mathbb{R}^d$ *with* $\hat{\mathrm{v}}_-(\gamma^{-1}k) \le \hat{\mathrm{v}}_-(k)$ *for* $k \in \mathbb{R}^d$. *Fix an inverse temperature* $\beta \in (0, \infty)$. *For any* $\mathrm{p} \in \mathrm{I}$, *there are two sequences* $(\gamma_{+,n})_{n\in\mathbb{N}}$ *and* $(\gamma_{-,n})_{n\in\mathbb{N}}$ *of real numbers in the interval* $(0, 1)$ *converging to zero, such that*

$$\lim_{n\to\infty} P_\beta\left(\gamma_{-,n}, \gamma_{+,n}\right) = \mathrm{p} \, .$$

Proof See [28, Theorem 5.19]. □

This theorem shows that interplay of the long-range limits $\gamma_\pm \to 0^+$ of short-range models can be highly non-trivial. In fact, as expected, any such long-range (Kac) limit leads to mean-field pressures and equilibrium states. However, in the presence of both repulsive and attractive forces, the limit mean-field model is **not necessarily** what one traditionally guesses. In fact, it strongly depends upon the hierarchy of ranges between attractive and repulsive interparticle forces. We have seen that if the range of repulsive forces is much larger than the range of the attractive ones, then in the Kac limit for these forces, one may get a limit mean-field model that is **unconventional**. See Theorems 6.39 and 6.40.

6.9.5 Historical Observations

The study on long-range limits presented here follows a rather old sequence of works on the Kac limit, basically starting from 1959, with Kac's own work on classical one-dimensional spin systems. The first important result [40] on this subject was provided by Penrose and Lebowitz in 1966, who proved the convergence of the free energy of a classical system toward the one of the van der Waals theory. Shortly after, the results of this seminal paper were extended to quantum systems (Boltzmann, Bose, or Fermi statistics) by Lieb [43]. In 1971, Penrose and Lebowitz went considerably further than [40] with [41]. See also [42] for a review of all these results of classical statistical mechanics. These outcomes form the mainstays of the subsequent results on the Kac limit, and we recommend the book [44] for a more recent review on the subject in classical statistical mechanics, including the so-called Lebowitz-Penrose theorem and a more exhaustive list of references.

Studies on the Kac limit are still performed nowadays in classical statistical mechanics; see, e.g., [45–47]. By contrast, to our knowledge, [28, Theorem 5.19] is the unique recent study on the subject for quantum systems, and the sole important results before [28, Theorem 5.19] are those of [43], which refer to quantum particles in the continuum, but may certainly be extended to lattice systems. The main innovation of [28, Theorem 5.19] is the fact that the convergence in the Kac limit is proven not only for pressure-like quantities (for instance, the thermodynamic limit of the logarithm of canonical or grand-canonical partition functions), as in previous works, but also for equilibrium states, i.e., for **all** correlation functions. These results on states were made possible by the variational approach of [1] for equilibrium states of mean-field models, which we present in a simpler setting in the first part of the current chapter. Additionally, also in contrast with previous results on Kac limits, our method allows for coexistence of both attractive and repulsive long-range forces. This important extension is related to the game theoretical characterization of equilibrium states of mean-field models (cf. thermodynamic game) discussed in Sects. 6.7 and 6.8. This approach thus paves the way for the study of phase transitions,[7] or at least important fingerprints of them like strong correlations at long distances, for models having interactions whose ranges are finite, but very large. It also sheds a new light on mean-field models by connecting them with short-range ones, in a mathematically precise manner. Such studies can be important for future theoretical developments in many-body theory, since long-range interactions are expected to imply effective, classical background fields, in the spirit of the Higgs mechanism of quantum field theory. This is shown in [48–50] for mean-field models.

6.10 The Generalized Hartree-Fock Theory as a Mean-Field Theory

In Sect. 5.7, we introduce the generalized Hartree-Fock theory [77, Definition 3.1], which approximates equilibrium states of fermion systems by means of (general) quasi-free states. Here, we illustrate the affinity of this method with mean-field theories. To this end, we consider an explicit, albeit still very general, fermion system that is similar to the model (6.38) studied in Sect. 6.9 in the context of the Kac limit.

[7] Mean-field repulsions have generally a geometrical effect by possibly breaking the face structure of the set of (generalized) equilibrium states (see [1, Lemma 9.8]). When this appears, we have long-range order for correlations. See [1, Section 2.9].

6.10.1 The Short-Range Model

In the current section, we only consider fermion systems, i.e., $\mathcal{U} \doteq \mathrm{CAR}(\Omega, \Gamma)$. As before, Ω denotes an arbitrary finite subset, which is fixed once and for all, and we use the short notation $a_{x,\mathrm{s}} \doteq a\left(\mathrm{e}_{\mathrm{s},x}\right)$ for the "annihilation operator" of a fermion with spin $\mathrm{s} \in \Omega$ at lattice position $x \in \Gamma$. Again, $\{\mathrm{e}_{\mathrm{s},x}\}_{(\mathrm{s},x)\in\Omega\times\Gamma}$ is the (canonical) Hilbert basis of $\ell^2(\Omega \times \Gamma)$, defined by $\mathrm{e}_{\mathrm{s},x}(\tilde{\mathrm{s}}, \tilde{x}) = 1$ if $(\mathrm{s}, x) = (\tilde{\mathrm{s}}, \tilde{x})$ and $\mathrm{e}_{\mathrm{s},x}(\tilde{\mathrm{s}}, \tilde{x}) = 0$, else. Considering fermions inside the cubic box

$$\Lambda_\ell \doteq \{(x_1, \ldots, x_d) \in \Gamma : |x_i| \leq \ell\}$$

for any $\ell \in \mathbb{N}$, the local Hamiltonians of our prototypical example studied here are equal to

$$H_{\Lambda_\ell} \doteq \sum_{x,y\in\Lambda_\ell,\ \mathrm{s}\in\Omega} \mathrm{h}\,(x-y)\, a^*_{x,\mathrm{s}} a_{y,\mathrm{s}} + \sum_{x,y\in\Lambda_\ell,\ \mathrm{s},\mathrm{t}\in\Omega} \mathrm{v}\,(x-y)\, a^*_{y,\mathrm{t}} a_{y,\mathrm{t}} a^*_{x,\mathrm{s}} a_{x,\mathrm{s}}\ ,$$

where $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}$ are two reflection-symmetric[8] functions. The (non-zero) function h encodes the hopping strength of fermions, while v is a (non-zero) pair potential characterizing interparticle forces. In contrast with (6.38), the function v has not necessarily positive values.

Such a family $(H_{\Lambda_\ell})_{\ell\in\mathbb{N}}$ of Hamiltonians is encoded by the (translation) invariant interaction $\Phi_{\mathrm{h},\mathrm{v}} = \Phi_{\mathrm{h}} + \Psi_{\mathrm{v}} \in \mathcal{V} \subseteq \mathcal{V}^{\mathbb{C}}$ (Definition 5.5), where the invariant self-conjugate interactions $\Phi_{\mathrm{h}}, \Psi_{\mathrm{v}} \in \mathcal{V}$ are defined by

$$\Psi_{\mathrm{v}}(\Lambda) \doteq 0 \doteq \Phi_{\mathrm{h}}(\Lambda)$$

whenever $|\Lambda| > 2$, while, for any $x, y \in \Gamma$,

$$\Phi_{\mathrm{h}}\left(\{x,y\}\right) \doteq \left(1 - \frac{1}{2}\delta_{x,y}\right) \sum_{\mathrm{s}\in\Omega} \mathrm{h}\,(x-y) \left(a^*_{x,\mathrm{s}} a_{y,\mathrm{s}} + a^*_{y,\mathrm{s}} a_{x,\mathrm{s}}\right)\ ,$$

$$\Psi_{\mathrm{v}}\left(\{x,y\}\right) \doteq \left(2 - \delta_{x,y}\right) \sum_{\mathrm{s},\mathrm{t}\in\Omega} \mathrm{v}\,(x-y)\, a^*_{y,\mathrm{t}} a_{y,\mathrm{t}} a^*_{x,\mathrm{s}} a_{x,\mathrm{s}}\ ,$$

$\delta_{x,y}$ being the Kronecker delta. In fact, using these definitions, we have

$$H_{\Lambda_\ell} = H_{\Lambda_\ell}^{\Phi_{\mathrm{h},\mathrm{v}}}$$

for all natural numbers $\ell \in \mathbb{N}$ and functions $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}$.

---

[8] That is, $\mathrm{h}\,(x) = \mathrm{h}\,(-x)$ and $\mathrm{v}\,(x) = \mathrm{v}\,(-x)$ for every $x \in \Gamma$.

We additionally impose the two reflection-symmetric functions h and v to be summable, i.e.,

$$\|\mathrm{h}\|_\Gamma \doteq \sum_{x\in\Gamma} |\mathrm{h}(x)| < \infty \qquad \text{and} \qquad \|\mathrm{v}\|_\Gamma \doteq \sum_{x\in\Gamma} |\mathrm{v}(x)| < \infty .$$

This implies that $\Phi_{\mathrm{h,v}} \in \mathcal{W}_1 \subseteq \mathcal{W}_1^{\mathbb{C}}$. See Definitions 5.6 and 6.2. In fact, note that the absolute summability of h and v is a necessary and sufficient condition to have $\Phi_{\mathrm{h,v}} \in \mathcal{W}_1$. It is a very weak condition in view of applications in condensed matter physics. For instance, taking $\mathrm{h}(x) = 0$ when $|x| > 1$ and $\mathrm{v}(x) = 0$ for $x \neq 0$, one obtains the celebrated Hubbard model. Note, moreover, that the summability of h and v is important to ensure the existence of the infinite-volume dynamics, via the celebrated Lieb-Robinson bounds (see, e.g., [94, Sections 4.1–4.2]). In fact, as explained in Paragraph 6.10.3, the existence of an infinite-volume dynamics is used in our arguments in order to link the generalized Hartree-Fock theory to mean-field models, via the KMS theory.

Observe from Proposition 5.11 and Definition 5.12 that the energy density functional

$$\begin{aligned} \mathfrak{e}_{\Phi_{\mathrm{h,v}}} : E_1 &\to \mathbb{R} \\ \rho &\mapsto \mathfrak{e}_{\Phi_{\mathrm{h,v}}}(\rho) \doteq \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|} \rho\left(H_{\Lambda_\ell}\right) \end{aligned}$$

associated with the invariant interaction $\Phi_{\mathrm{h,v}} \in \mathcal{W}_1$ naturally splits into two components:

$$\mathfrak{e}_{\Phi_{\mathrm{h,v}}} = \underbrace{\mathfrak{e}_{\Phi_{\mathrm{h}}}}_{\text{free term}} + \underbrace{\mathfrak{e}_{\Psi_{\mathrm{v}}}}_{\text{interaction term}} ,$$

where $\mathfrak{e}_{\Phi_{\mathrm{h}}} : E_1 \to \mathbb{R}$ and $\mathfrak{e}_{\Psi_{\mathrm{v}}} : E_1 \to \mathbb{R}$ are, respectively, equal to

$$\begin{aligned} \mathfrak{e}_{\Phi_{\mathrm{h}}}(\rho) &\doteq \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|} \sum_{x,y\in\Lambda_\ell,\ \mathrm{s}\in\Omega} \mathrm{h}(x-y)\, \rho\left(a_{x,\mathrm{s}}^* a_{y,\mathrm{s}}\right) \\ &= \frac{1}{2} \sum_{x\in\Gamma,\ \mathrm{s}\in\Omega} \mathrm{h}(x)\, \rho\left(a_{x,\mathrm{s}}^* a_{0,\mathrm{s}} + a_{0,\mathrm{s}}^* a_{x,\mathrm{s}}\right) \end{aligned} \tag{6.51}$$

and

$$\begin{aligned} \mathfrak{e}_{\Psi_{\mathrm{v}}}(\rho) &\doteq \lim_{\ell\to\infty} \frac{1}{|\Lambda_\ell|} \sum_{x,y\in\Lambda_\ell,\ \mathrm{s,t}\in\Omega} \mathrm{v}(x-y)\, \rho\left(a_{y,\mathrm{t}}^* a_{y,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}\right) \\ &= \sum_{x\in\Gamma,\ \mathrm{s,t}\in\Omega} \mathrm{v}(x)\, \rho\left(a_{0,\mathrm{t}}^* a_{0,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}\right) \end{aligned} \tag{6.52}$$

for any invariant state $\rho \in E_1$. With this, for any inverse temperature $\beta \in (0, \infty)$, the free energy density functional $\mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta} : E_1 \to \mathbb{R}$ of Definition 6.12 can be written as

$$\mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta} \doteq \mathfrak{e}_{\Phi_{\mathrm{h,v}}} - \beta^{-1}\mathfrak{s} = \mathfrak{e}_{\Phi_{\mathrm{h}}} + \mathfrak{e}_{\Psi_{\mathrm{v}}} - \beta^{-1}\mathfrak{s} \, ,$$

where $\mathfrak{s} : E_1 \to \mathbb{R}_0^+$ is the entropy density functional of Theorem 5.20. By Theorem 6.13, the thermodynamic limit of the (grand-canonical) pressure equals

$$P_\beta \doteq \lim_{\ell \to \infty} P_{H_{\Lambda_\ell},\beta} = -\inf \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}(E_1) < \infty \tag{6.53}$$

for any $\beta \in (0, \infty)$ and absolutely summable, reflection-symmetric, functions $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}$. See also (6.11). Recall that the globally stable equilibrium states of the short-range model are, by definition, the solutions to this variational problem. They form the set

$$\Omega_{\Phi_{\mathrm{h,v}},\beta} \doteq \left\{ \omega \in E_1 : \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}(\omega) = -P_\beta \right\} \tag{6.54}$$

for any fixed $\beta \in (0, \infty)$ and summable reflection-symmetric functions $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}$. By Lemma 6.30, it is a (nonempty) weak*-closed face of the convex weak*-compact space E_1 of invariant states.

6.10.2 Restriction to Quasi-Free States

In Definition 4.217, we introduce the notion of quasi-free states on self-dual CAR C^*-algebras. As explained after Proposition 4.219 with $G = H = \ell^2(\Omega \times \Gamma)$, recall from Corollary 4.207 that

$$\mathcal{U} \doteq \mathrm{CAR}(\Omega, \Gamma) = \mathrm{CAR}(\ell^2(\Omega \times \Gamma))$$

can be naturally identified with $\mathrm{sCAR}(\ell^2(\Omega \times \Gamma)_{\mathrm{sd}})$, where

$$\ell^2(\Omega \times \Gamma)_{\mathrm{sd}} \doteq \ell^2(\Omega \times \Gamma) \oplus_2 \ell^2(\Omega \times \Gamma)^{\mathrm{td}} \, ;$$

see Definition 4.195. We can thus use this identification of C^*-algebras to define from Definition 4.217 quasi-free states on $\mathcal{U}$. They are uniquely defined via Pfaffians and the two-point correlation functions $\rho(a_{x,\mathrm{s}}a_{y,\mathrm{t}})$ and $\rho(a_{x,\mathrm{s}}a_{y,\mathrm{t}}^*)$ for $x, y \in \Gamma$ and $\mathrm{s}, \mathrm{t} \in \Omega$. Such a quasi-free state ρ on $\mathcal{U}$ is called here simple, whenever $\rho(a_{x,\mathrm{s}}a_{y,\mathrm{t}}) = 0$ for all $x, y \in \Gamma$ and $\mathrm{s}, \mathrm{t} \in \Omega$.

Let

$$\mathcal{Q}_1 \doteq \{\rho \in E_1 : \ \rho \text{ is a quasi-free state}\}$$

be the (nonempty[9]) set of quasi-free states on $\mathcal{U}$. Note that a convex combination of quasi-free states is a state that is not necessarily quasi-free. In particular, $\mathcal{Q}_1$ is *not* a convex subset of the convex weak*-compact set E_1, but it is weak*-closed and therefore weak*-compact. This can be straightforwardly deduced from the definition of quasi-free states.

As explained in Sect. 5.7, we follow Bach, Lieb, and Solovej's approach [77] to the Hartree-Fock theory applied to the Hubbard model and, thus, minimize the free energy density functional in the set of all (not necessarily simple) quasi-free states. In other words, instead of (6.53), we study the variational problem

$$\inf \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}\left(\mathcal{Q}_1\right)$$

for any inverse temperature $\beta \in (0,\infty)$ and summable reflection-symmetric functions $\mathrm{h}:\Gamma\to\mathbb{R}$ and $\mathrm{v}:\Gamma\to\mathbb{R}$. Its solutions form a set denoted by

$$\mathrm{Q}_{\Phi_{\mathrm{h,v}},\beta} \doteq \left\{\omega\in\mathcal{Q}_1 : \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}(\omega) = \inf \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}\left(\mathcal{Q}_1\right)\right\} ,$$

which is in general rather different from the set $\Omega_{\Phi_{\mathrm{h,v}},\beta}$ of globally stable equilibrium states defined by (6.54).

To show that $\mathrm{Q}_{\Phi_{\mathrm{h,v}},\beta}$ is not empty, we observe that any invariant state $\rho\in E_1$ uniquely defines a quasi-free state via the corresponding symbol, and we rewrite the variational problem over quasi-free states as follows:

Lemma 6.41 *For any invariant state $\rho\in E_1$, there is a (unique) quasi-free state $q_\rho\in\mathcal{Q}_1$ satisfying*

$$q_\rho\left(a_{x,\mathrm{s}}a_{y,\mathrm{t}}\right)=\rho\left(a_{x,\mathrm{s}}a_{y,\mathrm{t}}\right) \qquad \text{and} \qquad q_\rho\left(a_{x,\mathrm{s}}a_{y,\mathrm{t}}^{*}\right)=\rho(a_{x,\mathrm{s}}a_{y,\mathrm{t}}^{*})$$

for all $x,y\in\Gamma$ and $\mathrm{s},\mathrm{t}\in\Omega$. The mapping $q:\rho\mapsto q_\rho$ from E_1 to $\mathcal{Q}_1$ is weak-continuous and satisfies $q_\rho=\rho$ for any quasi-free state $\rho\in\mathcal{Q}_1$.*

Proof The existence and uniqueness of $q_\rho\in\mathcal{Q}_1$ for any given $\rho\in E_1$ are a consequence of Exercise 4.216 and Proposition 4.219 together with Corollary 4.207, keeping in mind the definition of quasi-free states on CAR algebras just explained above. The weak* continuity of q can be verified by direct computations. See Definitions 4.80 and 4.217. We omit the details. □

Corollary 6.42 *For each inverse temperature $\beta\in(0,\infty)$ and any invariant interaction $\Phi\in\mathcal{W}_1$,*

$$\inf \mathfrak{f}_{\Phi,\beta}\left(\mathcal{Q}_1\right)=\inf \mathfrak{f}_{\Phi,\beta}\circ q\left(E_1\right)$$

and

[9] See, e.g., Proposition 4.219.

$$\left\{\omega \in \mathcal{Q}_1 : \mathfrak{f}_{\Phi,\beta}(\omega) = \inf \mathfrak{f}_{\Phi,\beta}(\mathcal{Q}_1)\right\} = q\left(\left\{\omega \in E_1 : \mathfrak{f}_{\Phi,\beta}(\omega) = \inf \mathfrak{f}_{\Phi,\beta} \circ q\,(E_1)\right\}\right)$$

is a nonempty weak-compact subset of* $\mathcal{Q}_1 \subseteq E_1$.

Proof The assertions are consequences of Lemma 6.41 combined with Lemmata 6.7 and 6.8. Note in particular from these statements that $\mathfrak{f}_{\Phi,\beta} \circ q$ is a weak*-lower semicontinuous functional on E_1, which is a weak*-compact set. As a consequence, the set of minimizers of $\mathfrak{f}_{\Phi,\beta} \circ q$ in E_1 is a nonempty weak*-closed, and thus compact, subset of $\mathcal{Q}_1 \subseteq E_1$. □

Applying this last corollary to the (short-range) model $\Phi_{\mathrm{h,v}} \in \mathcal{W}_1 \subseteq \mathcal{M}_1$ for any summable functions $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}$, we conclude in particular that $\mathrm{Q}_{\Phi_{\mathrm{h,v}},\beta}$ is a nonempty weak*-compact subset of $\mathcal{Q}_1 \subseteq E_1$ for any inverse temperature $\beta \in (0, \infty)$. In addition, the energy density functionals, respectively, associated with the kinetic and interparticle interactions have the following properties:

Corollary 6.43 *For any summable reflection-symmetric functions* $\mathrm{h} : \Gamma \to \mathbb{R}$ *and* $\mathrm{v} : \Gamma \to \mathbb{R}$, *we have* $\mathfrak{e}_{\Phi_{\mathrm{h}}} \circ q = \mathfrak{e}_{\Phi_{\mathrm{h}}}$ *and*

$$\mathfrak{e}_{\Psi_{\mathrm{v}}} \circ q\,(\rho) = \mathrm{v}(0) \sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^* a_{0,\mathrm{s}}\right) + \left(\sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^* a_{0,\mathrm{s}}\right)\right)^2 \sum_{x\in\Gamma} \mathrm{v}(x)$$
$$+ \sum_{x\in\Gamma} \mathrm{v}(x) \sum_{\mathrm{s,t}\in\Omega} \left(\left|\rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right)\right|^2 - \left|\rho\left(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}\right)\right|^2\right) .$$

Proof To obtain the equality $\mathfrak{e}_{\Phi_{\mathrm{h}}} \circ q = \mathfrak{e}_{\Phi_{\mathrm{h}}}$ it suffices to combine Lemma 6.41 with the explicit expression of the energy density $\mathfrak{e}_{\Phi_{\mathrm{h}}}$ given in Eq. (6.51). Now, recall from (6.52) that

$$\mathfrak{e}_{\Psi_{\mathrm{v}}}(\rho) = \sum_{x\in\Gamma,\ \mathrm{s,t}\in\Omega} \mathrm{v}(x)\, \rho\left(a_{0,\mathrm{t}}^* a_{0,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}\right)$$

for any invariant state $\rho \in E_1$. If $\rho \in \mathcal{Q}_1$ is a quasi-free state, then the 4-point correlation function $\rho\left(a_{0,\mathrm{t}}^* a_{0,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}\right)$ for $x \in \Gamma$ and $\mathrm{s,t} \in \Omega$ can be written in terms of 2-point correlation functions via the corresponding Pfaffians. More explicitly, one gets from Definitions 4.195 and 4.217 that, for any $\rho \in \mathcal{Q}_1$, $x \in \Gamma$ and $\mathrm{s,t} \in \Omega$,

$$\rho\left(a_{0,\mathrm{t}}^* a_{0,\mathrm{t}} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}\right) = \rho\left(a\left(\mathrm{e}_{\mathrm{t},0}^*\right) a\left(\mathrm{e}_{\mathrm{t},0}\right) a\left(\mathrm{e}_{\mathrm{s},x}^*\right) a\left(\mathrm{e}_{\mathrm{s},x}\right)\right)$$
$$= \mathrm{Pf}\begin{pmatrix} 0 & \rho(a_{0,\mathrm{t}}^* a_{0,\mathrm{t}}) & \rho(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}^*) & \rho(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}) \\ -\rho(a_{0,\mathrm{t}}^* a_{0,\mathrm{t}}) & 0 & \rho(a_{0,\mathrm{t}} a_{x,\mathrm{s}}^*) & \rho(a_{0,\mathrm{t}} a_{x,\mathrm{s}}) \\ -\rho(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}^*) & -\rho(a_{0,\mathrm{t}} a_{x,\mathrm{s}}^*) & 0 & \rho(a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}) \\ -\rho(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}) & -\rho(a_{0,\mathrm{t}} a_{x,\mathrm{s}}) & -\rho(a_{x,\mathrm{s}}^* a_{x,\mathrm{s}}) & 0 \end{pmatrix},$$

where $\mathrm{e}_{\mathrm{s},x}^{*} = (0, \langle \mathrm{e}_{\mathrm{s},x}, \cdot \rangle) \in \ell^{2}(\Omega \times \Gamma)_{\mathrm{sd}}$, the right-hand side of the first equality being written within the self-dual approach. It remains to compute this Pfaffian from its definition. See, e.g., the equation before Definition 4.217. By the CAR (Definition 4.163), note that, for any $x, y \in \Gamma$ and $\mathrm{s}, \mathrm{t} \in \Omega$,

$$a_{x,\mathrm{s}} a_{y,\mathrm{t}} + a_{y,\mathrm{t}} a_{x,\mathrm{s}} = 0 , \qquad a_{x,\mathrm{s}} a_{y,\mathrm{t}}^{*} + a_{y,\mathrm{t}}^{*} a_{x,\mathrm{s}} = \delta_{x,y} \delta_{\mathrm{s},\mathrm{t}} \mathbb{1} ,$$

keeping in mind that $a_{x,\mathrm{s}} = a\left(\mathrm{e}_{\mathrm{s},x}\right) \in \mathcal{U}$ with $\{\mathrm{e}_{\mathrm{s},x}\}_{(\mathrm{s},x) \in \Omega \times \Gamma}$ being, as is usual, the (canonical) Hilbert basis of $\ell^{2}(\Omega \times \Gamma)$. By plugging the CAR relations just stated, as well as the self-conjugate (or Hermitian) property of states, into the computation of the above Pfaffian, we arrive at

$$\begin{aligned}
\rho\left(a_{0,\mathrm{t}}^{*} a_{0,\mathrm{t}} a_{x,\mathrm{s}}^{*} a_{x,\mathrm{s}}\right) &= \rho\left(a_{0,\mathrm{t}}^{*} a_{0,\mathrm{t}}\right) \rho\left(a_{x,\mathrm{s}}^{*} a_{x,\mathrm{s}}\right) - \rho\left(a_{0,\mathrm{t}}^{*} a_{x,\mathrm{s}}^{*}\right) \rho\left(a_{0,\mathrm{t}} a_{x,\mathrm{s}}\right) \\
&\quad + \rho\left(a_{0,\mathrm{t}} a_{x,\mathrm{s}}^{*}\right) \rho\left(a_{0,\mathrm{t}}^{*} a_{x,\mathrm{s}}\right) \\
&= \rho\left(a_{0,\mathrm{t}}^{*} a_{0,\mathrm{t}}\right) \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right) + \rho\left(\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right)^{*}\right) \rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right) \\
&\quad - \rho\left(a_{x,\mathrm{s}}^{*} a_{0,\mathrm{t}}\right) \rho\left(a_{0,\mathrm{t}}^{*} a_{x,\mathrm{s}}\right) + \delta_{x,0} \delta_{\mathrm{s},\mathrm{t}} \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right) \\
&= \rho\left(a_{0,\mathrm{t}}^{*} a_{0,\mathrm{t}}\right) \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right) + \rho\left(\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right)^{*}\right) \rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right) \\
&\quad - \rho\left(\left(a_{0,\mathrm{t}}^{*} a_{x,\mathrm{s}}\right)^{*}\right) \rho\left(a_{0,\mathrm{t}}^{*} a_{x,\mathrm{s}}\right) + \delta_{x,0} \delta_{\mathrm{s},\mathrm{t}} \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right) \\
&= \rho\left(a_{0,\mathrm{t}}^{*} a_{0,\mathrm{t}}\right) \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right) + \left|\rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right)\right|^{2} - \left|\rho\left(a_{0,\mathrm{t}}^{*} a_{x,\mathrm{s}}\right)\right|^{2} \\
&\quad + \delta_{x,0} \delta_{\mathrm{s},\mathrm{t}} \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right)
\end{aligned}$$

for any invariant quasi-free state $\rho \in \mathcal{Q}_{1}$, lattice position $x \in \Gamma$, and spin $\mathrm{s}, \mathrm{t} \in \Omega$. Using this result together with Lemma 6.41, we deduce the expression in the corollary for the energy density functional $\mathfrak{e}_{\Psi_{\mathrm{v}}} \circ q$. □

Since the free energy density functional of the model studied here equals

$$\mathfrak{f}_{\Phi_{\mathrm{h},\mathrm{v}},\beta} = \mathfrak{e}_{\Phi_{\mathrm{h}}} + \mathfrak{e}_{\Psi_{\mathrm{v}}} - \beta^{-1} \mathfrak{s} ,$$

for any inverse temperature $\beta \in (0, \infty)$ and summable reflection-symmetric functions $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}$, we conclude from Corollaries 6.42 and 6.43 that the variational problem

$$\inf \mathfrak{f}_{\Phi_{\mathrm{h},\mathrm{v}},\beta} \left(\mathcal{Q}_{1}\right)$$

on the weak*-compact set $\mathcal{Q}_{1}$ of invariant quasi-free states on $\mathcal{U}$ can be studied by minimizing on the set E_{1} of all invariant states the weak*-lower semicontinuous[10]

[10] To prove the lower semicontinuity of $\tilde{g}_{\mathrm{h},f,\beta}$, combine Lemmata 6.7, 6.8, and 6.41, as is already done in the proof of Corollary 6.42.

functional $\mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta} \circ q$, which is *quadratic* in the energy densities and thus similar to the nonlinear free energy density functional $\mathfrak{g}_{\mathfrak{m},\beta} : E_1 \to \mathbb{R}$ of Definition 6.17, for a mean-field model $\mathfrak{m} \in \mathcal{M}_1$. In fact, by Definition 6.17 and Corollary 6.43, there is a mean-field model $\mathfrak{m}_{\mathrm{h,v}} \in \mathcal{M}_1$ such that, for any invariant state $\rho \in E_1$ and every inverse temperature $\beta \in (0, \infty)$,

$$\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(\rho) = \tilde{f}_{\mathrm{h,v},\beta}(\rho) + \left(\sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^* a_{0,\mathrm{s}}\right) \right)^2 \sum_{x\in\Gamma} \mathrm{v}(x) + \sum_{x\in\Gamma} \mathrm{v}(x) \sum_{\mathrm{s,t}\in\Omega} \left(\left|\rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right)\right|^2 - \left|\rho\left(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}\right)\right|^2 \right)$$

where $\tilde{f}_{\mathrm{h,v},\beta} : E_1 \to \mathbb{R}$ is the weak*-lower semicontinuous and affine functional defined by

$$\tilde{f}_{\mathrm{h,v},\beta}(\rho) \doteq \mathfrak{e}_{\Phi_{\mathrm{h}}}(\rho) + \mathrm{v}(0) \sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^* a_{0,\mathrm{s}}\right) - \beta^{-1} \mathfrak{s}(\rho)$$

for $\rho \in E_1$ and $\beta \in (0, \infty)$. Remark that $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta} = \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}$ on the set $\mathcal{Q}_1$ of invariant quasi-free states but this equality does *not* a priori hold true on the whole set $E_1 \supseteq \mathcal{Q}_1$.

The mean-field model $\mathfrak{m}_{\mathrm{h,v}}$ can be explicitly written by using some bijection from $\mathbb{N}$ to $\Omega \times \Gamma$, but we omit its explicit form to simplify our discussions and focus on the main arguments. Observe only that the mean-field model $\mathfrak{m}_{\mathrm{h,v}} \in \mathcal{M}_1$ refers to a fermion system, whose local Hamiltonians (Definition 6.4) are

$$H_\Lambda^{\mathfrak{m}_{\mathrm{h,v}}} = \sum_{x,y\in\Lambda,\ \mathrm{s}\in\Omega} \mathrm{h}(x-y)\, a_{x,\mathrm{s}}^* a_{y,\mathrm{s}} + \mathrm{v}(0) \sum_{x\in\Lambda,\ \mathrm{s}\in\Omega} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}} + \frac{1}{|\Lambda|} \sum_{z\in\Gamma} \mathrm{v}(z) \left| \sum_{x\in\Lambda,\ \mathrm{s}\in\Omega} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}} \right|^2 + \frac{1}{|\Lambda|} \sum_{z\in\Gamma} \mathrm{v}(z) \sum_{\mathrm{s,t}\in\Omega} \left(\left| \sum_{x,x+z\in\Lambda} a_{x+z,\mathrm{s}} a_{x,\mathrm{t}} \right|^2 - \left| \sum_{x,x+z\in\Lambda} a_{x,\mathrm{t}}^* a_{x+z,\mathrm{s}} \right|^2 \right)$$

for any finite subset $\Lambda \in \mathcal{P}_f$. This mean-field model is highly non-trivial and even includes BCS-type interactions; see, e.g., Example 6.6. This may give the impression that the original short-range model can imply a superconducting phase transition at low temperatures (for non-positive v), but one shall refrain from making such rapid conclusions, since the generalized Hartree-Fock theory could significantly overestimate the true free energy density.

6.10.3 Thermodynamic Game and Bogoliubov Approximation

The set $\hat{M}_{\mathfrak{m}_{h,v},\beta}$ of minimizers of the variational problem

$$\inf \mathfrak{g}_{\mathfrak{m}_{h,v},\beta}\left(E_1\right)$$

(see Definition 6.21) can be completely described via Bogoliubov approximations for the associated mean-field models, as explained in Sect. 6.5. This brings us to the thermodynamic game introduced in Sect. 6.7. In fact, similar to Lemma 6.27, the following assertions hold true:

Lemma 6.44 (Bogoliubov Approximation) *Let $\rho \in E_1$ be any invariant state.*

(i) *Given $\gamma \geq 0$, the positive number*

$$r\left(\rho\right) = \gamma \sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right)$$

is the unique maximizer of the variational problem

$$\sup_{r\in\mathbb{R}_0^+}\left\{-r^2 + 2r\gamma \sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right)\right\} = \left(\gamma \sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right)\right)^2 .$$

(ii) *For any function $\xi \in \ell^2(\Omega^2 \times \Gamma, \mathcal{U})$,*

$$\sup_{c\in\ell^2(\Omega^2\times\Gamma)}\left\{-\left\|c\right\|_2^2 + 2\,\mathrm{Re}\left\langle c, \rho\left(\xi\right)\right\rangle\right\} = \left\|\rho\left(\xi\right)\right\|_2^2$$

with unique a maximizer

$$d\left(\rho\right) = \rho\left(\xi\right) \doteq \left(\rho\left(\xi\left(\mathrm{s},\mathrm{t},x\right)\right)\right)_{(\mathrm{s},\mathrm{t},x)\in\Omega^2\times\Gamma} \in \ell^2(\Omega^2\times\Gamma) .$$

Proof The proof is the same as the one of Lemma 6.27 and it is therefore omitted. We only remark that the variational problem in (i) can be restricted to positive numbers because $\rho\left(a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}}\right) \in \mathbb{R}_0$, a state being by definition a positive functional and $a_{0,\mathrm{s}}^{*} a_{0,\mathrm{s}} \geq 0$ in $\mathcal{U}$. □

To apply Lemma 6.44, we need to keep track of the sign of the function v at each lattice site. With this aim, as is usual, one splits v into its positive and negative components, $\mathrm{v} = \mathrm{v}_+ - \mathrm{v}_-$, where $\mathrm{v}_+(x) \doteq \sup\{\mathrm{v}(x), 0\}$ and $\mathrm{v}_-(x) \doteq \sup\{-\mathrm{v}(x), 0\}$ for all $x \in \Gamma$. This is similar to what is done for mean-field models for which we must distinguish between the effects of mean-field attractions and repulsions. For the sake of simplicity, we assume from now that $\mathrm{v} = \mathrm{v}_+ \geq 0$. Notice, however, that this special case already yields a non-trivial thermodynamic game, that is, the corresponding mean-field model is neither purely attractive nor repulsive. In fact, the generalization to functions v not having a definite sign is straightforward; it is only a matter of "bookkeeping."

Theorem 6.45 (Hartree-Fock Thermodynamic Game—Repulsive Case) *For any inverse temperature $\beta \in (0, \infty)$ and summable reflection-symmetric functions* $\mathrm{h} : \Gamma \to \mathbb{R}$ *and* $\mathrm{v} : \Gamma \to \mathbb{R}_0^+$,

$$\inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(E_1) = \inf_{c_- \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v}))} \sup_{r \in \mathbb{R}_0^+} \sup_{c_+ \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v}))} \left\{ \|c_-\|_2^2 - \|c_+\|_2^2 - r^2 + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta}(E_1) \right\}$$

where

$$\mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta} \doteq \tilde{f}_{\mathrm{h,v},\beta}(\rho) + 2r \|\mathrm{v}\|_\Gamma^{1/2} \sum_{\mathrm{s}\in\Omega} \rho\left(a_{0,\mathrm{s}}^* a_{0,\mathrm{s}}\right) + 2 \sum_{x \in \mathrm{supp}(\mathrm{v})} \sqrt{\mathrm{v}(x)} \times \sum_{\mathrm{s,t}\in\Omega} \mathrm{Re}\left\{ \overline{c_+(\mathrm{s,t},x)} \rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right) - \overline{c_-(\mathrm{s,t},x)} \rho\left(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}\right) \right\} .$$

Idea of the Proof The proof is a slightly simplified version of the one of Theorem 6.34, which uses among other things Lemma 6.44 together with the von Neumann min-max theorem [1, Theorem 10.50] to be able to exchange the two suprema of the assertion with the infimum over invariant states $\rho \in E_1$. □

Recall that

$$\|\mathrm{v}\|_\Gamma \doteq \sum_{x\in\Gamma} |\mathrm{v}(x)| < \infty$$

and $\mathrm{supp}(\mathrm{v}) \subseteq \Gamma$ stands for the support of the function v. Here,

$$\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-, c_+, r) \equiv \Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-, (c_+, r)) \in \mathcal{W}_1$$

refers to the approximating interactions associated with the mean-field model $\mathfrak{m}_{\mathrm{h,v}}$. See Definition 6.26.

It is now clear from Theorem 6.45 that the variational problem $\inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(E_1)$ for non-zero functions v can be seen as the conservative value of a ("Hartree-Fock thermodynamic") game, with players (−) and (+), whose sets of strategies are $\ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v}))$ for (−) and $\ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v})) \times \mathbb{R}_0^+$ for (+). The minimizers of $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ can also be derived from this game:

Note that $\mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta}$ is an affine and lower weak* semicontinuous functional on E_1, which is a weak*-compact and convex set. Thus, define its weak*-compact and convex set of minimizers by

$$\Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta} \doteq \left\{\omega \in E_1 : \mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta}(\omega) = \inf \mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta}(E_1)\right\} .$$

By [1, Lemma 8.3 (♯)], for any $\beta \in (0,\infty)$ and summable reflection-symmetric functions $\mathrm{h} : \Gamma \to \mathbb{R}$ and $\mathrm{v} : \Gamma \to \mathbb{R}_0^+$, $\mathrm{v} \neq 0$, and all $c_- \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v}))$, there is exactly one element, which is denoted by

$$\mathrm{r}_+(c_-) \doteq (d_+(c_-), \mathbf{r}(c_-)) \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v})) \times \mathbb{R}_0^+ ,$$

such that

$$\begin{aligned} h^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}(c_-) &\doteq \|c_-\|_2^2 + \sup_{r\in\mathbb{R}_0^+} \sup_{c_+\in\ell^2(\Omega^2\times\mathrm{supp}(\mathrm{v}))} \Big\{-\|c_+\|_2^2 - r^2 \\ &\qquad + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta}(E_1)\Big\} \\ &= \|c_-\|_2^2 - \|d_+(c_-)\|_2^2 - \mathbf{r}(c_-)^2 + \inf \mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,d_+(c_-),\mathbf{r}(c_-)),\beta}(E_1) . \end{aligned}$$

Using the set

$$\mathcal{C}^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta} \doteq \left\{d_- \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v})) : \inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(E_1) = h^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}(d_-)\right\} ,$$

one can characterize minimizers of the nonlinear free energy functional $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ as follows: For all strategies $c_- \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v}))$ and $(c_+, r) \in \ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v})) \times \mathbb{R}_0^+$, we define the (possibly empty) set

$$\begin{aligned} \Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}},\beta}(c_-, c_+, r) &\doteq \Big\{\omega \in \Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta} \ : \ \|\mathrm{v}\|_\Gamma^{1/2} \sum_{\mathrm{s}\in\Omega} \omega\left(a_{0,\mathrm{s}}^* a_{0,\mathrm{s}}\right) = r , \\ \sqrt{\mathrm{v}(x)}\omega\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right) &= c_+(\mathrm{s},\mathrm{t},x) , \ \sqrt{\mathrm{v}(x)}\omega\left(a_{0,\mathrm{t}}^* a_{x,\mathrm{s}}\right) = c_-(\mathrm{s},\mathrm{t},x) , \ (\mathrm{s},\mathrm{t},x) \in \Omega^2 \\ &\qquad \times \mathrm{supp}(\mathrm{v})\Big\} \subseteq E_1 . \end{aligned}$$

Note that the above set $\Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}},\beta}(c_-, c_+, r)$ of self-consistent equilibrium states is an instance of (6.35). With these definitions, we have the following assertion:

Theorem 6.46 (Self-Consistency—Repulsive Case) *For any inverse temperature* $\beta \in (0,\infty)$ *and summable reflection-symmetric functions* $\mathrm{h} : \Gamma \to \mathbb{R}$ *and* $\mathrm{v} : \Gamma \to \mathbb{R}_0^+$,

$$\hat{M}_{\mathfrak{m}_{\mathrm{h,v}},\beta} = \bigcup_{d_-\in\mathcal{C}^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}} \Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}},\beta}(d_-, d_+(d_-), \mathbf{r}(d_-))$$

Proof This theorem is proven like Theorem 6.36 (i). For a complete proof, see [1, Theorem 9.4]. □

Similar to Theorem 6.36, this last theorem characterizes minimizers of the nonlinear free energy functional $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ by (static) *self-consistency conditions*, which refer, in a sense, to Euler-Lagrange equations for the variational problem defining the thermodynamic game.

It turns out that under mild conditions on the functions h and v, the approximating interactions $\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-, c_+, r)$ have exactly one equilibrium state. By Theorem 6.36 (iii), note additionally that $\Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(d_-,d_+(d_-),\mathbf{r}(d_-)),\beta}$ is never empty, for any $d_- \in \mathcal{C}^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}$. Thus, under these conditions, the corresponding sets $\Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}},\beta}(c_-, c_+, r)$ of self-consistent equilibrium states have at most one element, and, from the last theorem, we arrive at

$$\hat{M}_{\mathfrak{m}_{\mathrm{h,v}},\beta} = \bigcup_{d_- \in \mathcal{C}^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}} \Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(d_-,d_+(d_-),\mathbf{r}(d_-)),\beta} \ .$$

In other words, in this case, the set of minimizers of $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ is nothing else than the collection of the unique equilibrium states of the corresponding approximating interactions $\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(d_-, d_+(d_-), \mathbf{r}(d_-))$ for $d_- \in \mathcal{C}^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}$. What is more, beyond this nice property of $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$, it turns out that the same condition guaranteeing the uniqueness of the equilibrium state of $\Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}},\beta}(c_-, c_+, r)$ also implies that $\hat{M}_{\mathfrak{m}_{\mathrm{h,v}},\beta} \subseteq \mathcal{Q}_1$, i.e., all the minimizers of $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ are quasi-free states. In other words, the Hartree-Fock equilibrium states for model considered in this subsection are exactly the minimizers of $\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ in the set of all (i.e., not necessarily quasi-free) invariant states, that is, the nonlinear equilibrium states of an explicit mean-field model $\mathfrak{m}_{\mathrm{h,v}}$.

In order to prove this claim, we now discuss in more detail the relation between the variational problem

$$\inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(E_1) \ ,$$

along with its set $\hat{M}_{\mathfrak{m}_{\mathrm{h,v}},\beta}$ of minimizers, and the variational problem

$$\inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(\mathcal{Q}_1) = \inf \mathfrak{f}_{\Phi_{\mathrm{h,v}},\beta}(\mathcal{Q}_1) \ ,$$

along with its set $\mathrm{Q}_{\Phi_{\mathrm{h,v}},\beta}$ of minimizers, given by the generalized Hartree-Fock theory, which is our main concern in Sect. 6.10:

Since $\mathcal{Q}_1 \subseteq E_1$, one has trivially the inequality

$$\inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(E_1) \leq \inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(\mathcal{Q}_1) \ .$$

Further, one observes that the approximating (invariant) interaction $\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-, c_+, r)$ associated with the mean-field model $\mathfrak{m}_{\mathrm{h,v}} \in \mathcal{M}_1$ for $c_-, c_+ \in \ell^2(\Omega^2 \times \mathrm{supp(v)})$ and $r \in \mathbb{R}_0^+$ corresponds to even self-adjoint elements of $\mathcal{U}$ that are *quadratic* in the creation and annihilation elements $a_{x,\mathrm{s}}, a_{x,\mathrm{s}}^*$, for $x \in \Gamma$ and $\mathrm{s} \in \Omega$. For instance, the associated local Hamiltonians (Definition 6.4) (for positive $\mathrm{v} \geq 0$) are equal to

$$\begin{aligned} H_\Lambda^{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r)} = & \sum_{x,y\in\Lambda,\ \mathrm{s}\in\Omega} \mathrm{h}\,(x-y)\, a_{x,\mathrm{s}}^* a_{y,\mathrm{s}} + \left(\mathrm{v}\,(0) + 2r\,\|\mathrm{v}\|_\Gamma^{1/2}\right) \\ & \times \sum_{x\in\Lambda,\ \mathrm{s}\in\Omega} a_{x,\mathrm{s}}^* a_{x,\mathrm{s}} \\ & +2\sum_{z\in\Gamma}\sqrt{\mathrm{v}\,(z)} \sum_{\mathrm{s,t}\in\Omega}\sum_{x\in\Lambda} \mathrm{Re}\left\{\overline{c_+\,(\mathrm{s,t},x)} a_{x+z,\mathrm{s}} a_{x,\mathrm{t}}\right. \\ & \left. -\overline{c_-\,(\mathrm{s,t},x)} a_{x,\mathrm{t}}^* a_{x+z,\mathrm{s}}\right\} \end{aligned}$$

for any finite subset $\Lambda \in \mathcal{P}_f$ and every $c_-, c_+ \in \ell^2(\Omega^2 \times \Gamma)$ and $r \in \mathbb{R}_0^+$. Such kind of quadratic, or bilinear, Hamiltonians can be explicitly diagonalized by a so-called Bogoliubov transformation, as already shown in Berezin's book [95], published in 1966. In other words, the thermodynamic game associated with Theorem 6.45 can be studied from finite-volume systems for which explicit computations can be made.

More generally, if the reflection-symmetric functions h and v are not only summable but also decaying sufficiently fast,[11] as $|x| \rightarrow \infty$, such bilinear Hamiltonians are well-known to generate an infinite-volume dynamics which is a strongly continuous group of Bogoliubov $*$-automorphisms, as given by Definition 4.181 and Corollary 4.183. See, for instance, [96, Lemma 2.8]. Araki proves in [69, Theorem 3] the existence of a unique KMS (Kubo-Martin-Schwinger) state associated with such a group of Bogoliubov $*$-automorphisms, which turns out to be a *quasi-free* state. See, for instance, Proposition 3.33 for the KMS condition in finite dimensions. For more details, see, e.g., [55, Sections 5.3–5.4]. If the bilinear model defining the dynamics is invariant, its KMS state is also invariant. If it is gauge-invariant, then the KMS state is also (globally) gauge-invariant and therefore a *simple* quasi-free state.

The relation between KMS states and minimizers of a variational problem derived from the same sufficiently short-range interaction has been studied for lattice fermions by Araki and Moriya [15]: It turns out that all minimizers of a variational problem like

$$\inf \mathfrak{f}_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta}\,(E_1)$$

for sufficiently decaying reflection-symmetric functions h and v are KMS states. See, e.g., [97, Theorem 3.1]. Since, in this case, the KMS state is unique, invariant, and quasi-free, we conclude that

$$\Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,c_+,r),\beta} = \{\omega\} \subseteq \mathcal{Q}_1 .$$

[11] For instance, they show a sufficiently fast polynomial decay.

Therefore, it follows, in this situation, that

$$\hat{M}_{\mathfrak{m}_{\mathrm{h,v}},\beta} = \bigcup_{d_- \in \mathcal{C}^\sharp_{\mathfrak{m}_{\mathrm{h,v}},\beta}} \Omega_{\Phi_{\mathfrak{m}_{\mathrm{h,v}}}(c_-,d_+(c_-),\mathbf{r}(c_-)),\beta} = \mathrm{Q}_{\Phi_{\mathrm{h,v}},\beta} \subseteq \mathcal{Q}_1 \ ,$$

meaning in particular that

$$\inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(E_1) = \inf \mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(\mathcal{Q}_1) = \inf \mathfrak{f}_{\Phi_{\mathrm{h},f},\beta}(\mathcal{Q}_1) \ .$$

This explicitly shows the mean-field character of the general Hartree-Fock theory applied on our prototypical (though very general, quartic) short-range model (Sect. 6.10.1), which includes the celebrated Hubbard model, widely used in Physics, as one simple example.

To conclude, notice that if one would only consider simple (i.e., gauge-invariant) quasi-free states in the Hartree-Fock theory for our prototypical model, everything that is said above still applies, mutatis mutandis, by considering the (simpler) nonlinear energy density functional

$$\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(\rho) = \tilde{f}_{\mathrm{h,v},\beta}(\rho) + \left(\sum_{\mathrm{s}\in\Omega} \rho\left(a^*_{0,\mathrm{s}} a_{0,\mathrm{s}}\right)\right)^2 \sum_{x\in\Gamma} \mathrm{v}(x) - \sum_{x\in\Gamma} \mathrm{v}(x) \sum_{\mathrm{s,t}\in\Omega} \left|\rho\left(a^*_{0,\mathrm{t}} a_{x,\mathrm{s}}\right)\right|^2 \ ,$$

instead of

$$\mathfrak{g}_{\mathfrak{m}_{\mathrm{h,v}},\beta}(\rho) = \tilde{f}_{\mathrm{h,v},\beta}(\rho) + \left(\sum_{\mathrm{s}\in\Omega} \rho\left(a^*_{0,\mathrm{s}} a_{0,\mathrm{s}}\right)\right)^2 \sum_{x\in\Gamma} \mathrm{v}(x) + \sum_{x\in\Gamma} \mathrm{v}(x) \sum_{\mathrm{s,t}\in\Omega} \left(\left|\rho\left(a_{x,\mathrm{s}} a_{0,\mathrm{t}}\right)\right|^2 - \left|\rho\left(a^*_{0,\mathrm{t}} a_{x,\mathrm{s}}\right)\right|^2\right) \ .$$

In particular, even in this simpler case, for positive $\mathrm{v} \geq 0$, we still have a non-trivial thermodynamic game with two players, $(-)$ and $(+)$, whose sets of strategies are now $\ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v}))$ (as before) for $(-)$ and $\mathbb{R}_0^+$ (instead of $\ell^2(\Omega^2 \times \mathrm{supp}(\mathrm{v})) \times \mathbb{R}_0^+$) for $(+)$.

Chapter 7
Appendix

7.1 Vector Spaces and Algebras

We discuss in the present section elementary definitions and properties related to vector spaces and algebras, as well as basic constructions with these spaces, like quotients, tensor products, free algebras, and the like. We start by recalling the formal definition of a vector space:

Definition 7.1 (Vector Space) The triple $(V, +, \cdot)$, where V is a nonempty set with two operations, $+ : V \times V \to V$ (sum) and $\cdot : \mathbb{K} \times V \to V$ (scalar multiplication), is a "vector space" over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ if the operations have the following properties:

(i) $(V, +)$ is an abelian group:

- (i.a) *Commutativity*. For all $v_1, v_2 \in V$, $v_1 + v_2 = v_2 + v_1$.
- (i.b) *Associativity.* For all $v_1, v_2, v_3 \in V$, $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3) \doteq v_1 + v_2 + v_3$.
- (i.c) *Neutral element*. There is an element $0 \in V$ (which is necessarily unique) such that, for all $v \in V$, $0 + v = v$.
- (i.d) *Inverse element.* For any $v \in V$, there is $-v \in V$ (which is necessarily unique), for which $v + (-v) = 0$.

(ii) The scalar multiplication $(\alpha, v) \mapsto \alpha \cdot v \in V$, $\alpha \in \mathbb{K}$, $v \in V$, has the following properties:

- (ii.a) *Associativity.* $\beta \cdot (\alpha \cdot v) = (\beta\alpha) \cdot v$ for all $\alpha, \beta \in \mathbb{K}$, $v \in V$.
- (ii.b) $1 \cdot v = v$ for all $v \in V$, where $1 \in \mathbb{K}$ is the unit of $\mathbb{K}$.
- (ii.c) *Distributivity with respect to* $(V, +)$. For all $v_1, v_2 \in V$ and $\alpha \in \mathbb{K}$, $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$.
- (ii.d) *Distributivity with respect to* $(\mathbb{K}, +)$. For all $v \in V$ and $\alpha, \beta \in \mathbb{K}$, $(\alpha + \beta) \cdot v = \alpha \cdot v + \beta \cdot v$.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1_7

Additionally, $V' \subseteq V$, $V' \neq \emptyset$, is a (vector) “subspace” of the vector space $(V, +, \cdot)$ if, for all $v'_1, v'_2 \in V'$ and $\alpha \in \mathbb{K}$,

$$v'_1 + v'_2 \in V' \qquad \text{and} \qquad \alpha \cdot v'_1 \in V' \ .$$

Note that the field $\mathbb{K} = \mathbb{R}, \mathbb{C}$ itself is (canonically) a vector space over $\mathbb{K}$. By a slight abuse of notation, for simplicity, a generic vector space $(V, +, \cdot)$, which is formally a triple, is frequently denoted here by the simple name of the set on which its operations are defined, i.e., V.

Definition 7.2 (Direct Sum and Cartesian Product of Vector Spaces)

(i) Let V be any vector space and $V_1, \ldots, V_N \subseteq V$, $N \in \mathbb{N}$, any finite sequence of vector subspaces of V. We say that V is the “direct sum” of this sequence of subspaces if, for every $v \in V$, there is a *unique* sequence of vectors $v_k \in V_k$, $k \in \{1, \ldots, N\}$, such that $v = v_1 + \cdots + v_N$.

(ii) Let $V_1, \ldots, V_N$ be an arbitrary sequence of vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then the Cartesian product $V_1 \times \cdots \times V_N$ is (canonically) a $\mathbb{K}$-vector space with the following vector space operations:

$$(v_1, \ldots, v_N) + (v'_1, \ldots, v'_N) \doteq (v_1 + v'_1, \ldots, v_N + v'_N) \ ,$$
$$\alpha \cdot (v_1, \ldots, v_N) \doteq (\alpha \cdot v_1, \ldots, \alpha \cdot v_N) \ ,$$

for any $v_k, v'_k \in V_k$, $k \in \{1, \ldots, N\}$ and $\alpha \in \mathbb{K}$. This vector space is denoted by $V_1 \times \cdots \times V_N$, by a slight abuse of notation. By identifying the vector $v_k \in V_k$, $k \in \{1, \ldots, N\}$, with $(0, \ldots, v_k, \ldots, 0) \in V_1 \times \cdots \times V_N$, V_k is (canonically) a vector subspace of the Cartesian product $V_1 \times \cdots \times V_N$, which is, in turn, the direct sum of the subspaces $V_1, \ldots, V_N$.

In vector spaces, there is a natural notion of vector space dimension based on linearly independent vectors, defined as follows:

Definition 7.3 (Linear (In)dependence and Bases) Let V be any vector space and $\Omega \subseteq V$ any nonempty subset.

(i) Elements of Ω are “linearly independent” if, for any finite sequence $v_1, \ldots, v_N \subseteq \Omega$, $N \in \mathbb{N}$, the equality $\alpha_1 v_1 + \cdots + \alpha_N v_N = 0$ holds true only if $\alpha_1 = \cdots = \alpha_N = 0$. Otherwise, they are said to be “linearly dependent.”

(ii) $\text{span}(\Omega) \subseteq V$ denotes the smallest[1] vector subspace of V that comprises Ω and is called the subspace (linearly) “generated” by Ω. If $\text{span}(\Omega) = V$, we say that Ω is “generating” for V.

[1] Note that such a subspace always exists, for arbitrary intersections of vector subspaces of a given vector spaces are new subspaces.

(iii) Ω is a (Hamel) "basis" of V when its elements are linearly independent and it generates the whole space V.

Recall that every vector space has a (Hamel) basis, as a consequence of Zorn's lemma (or, equivalently, the axiom of choice).

Definition 7.4 (Linear Transformation) Let V_1 and V_2 be two vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. The mapping $\Theta : V_1 \to V_2$ is "($\mathbb{K}$-)linear" if, for all $v_1, v_1' \in V_1$ and $\alpha \in \mathbb{K}$,

$$\Theta(v_1 + v_1') = \Theta(v_1) + \Theta(v_1') \qquad \text{and} \qquad \Theta(\alpha v_1) = \alpha\Theta(v_1) \; .$$

The set of all linear transformations $V_1 \to V_2$ is denoted by $\mathcal{L}(V_1; V_2)$. If $V_1 = V_2$, then we use the shorter notation $\mathcal{L}(V_1)$. If the linear transformation $\Theta : V_1 \to V_2$ is a one-to-one correspondence (i.e., it is a bijection), then it is called an "isomorphism" of the vector spaces V_1 and V_2. In this case, V_1 and V_2 are said to be "isomorphic" vector spaces.

There is a special terminology for linear transformations, whose codomain is the field $\mathbb{K} = \mathbb{R}, \mathbb{C}$:

Definition 7.5 (Dual Spaces and Linear Functionals) Let V be any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We define its dual vector space by $V' \doteq \mathcal{L}(V; \mathbb{K})$. The elements of $V'$are called "linear functionals" on V. The vector space V is canonically identified with a vector subspace of its "bidual" V'' (i.e., the dual vector space of V') as follows:

$$v(\varphi) \doteq \varphi(v) \; , \qquad v \in V, \; \varphi \in V'.$$

An important vector subspace associated with an arbitrary linear transformation is its kernel, defined as follows:

Definition 7.6 (Kernel of a Linear Transformation) Let V_1 and V_2 be any two vector spaces and $\Theta \in \mathcal{L}(V_1; V_2)$. The vector subspace

$$\ker(\Theta) \doteq \{v_1 \;:\; \Theta(v_1) = 0\} \subseteq V_1$$

is called the "kernel" of the linear transformation Θ.

Let Ω be an arbitrary nonempty set and V a vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. The set $\mathcal{F}(\Omega; V)$ of all functions $\Omega \to V$ is a vector space over $\mathbb{K}$ with the following operations:

(i) For any $f_1, f_2 \in \mathcal{F}(\Omega; V)$, we define $f_1 + f_2 \in \mathcal{F}(\Omega; V)$ by

$$[f_1 + f_2](p) \doteq f_1(p) + f_2(p) \; , \qquad p \in \Omega \; .$$

(ii) For any $f \in \mathcal{F}(\Omega; V)$ and $\alpha \in \mathbb{K}$, $\alpha \cdot f \in \mathcal{F}(\Omega; V)$ is defined by

$$[\alpha \cdot f](p) \doteq \alpha \cdot f(p)\,, \qquad p \in \Omega\,.$$

Note, moreover, that if V_1 and V_2 are two vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ then $\mathcal{L}(V_1; V_2)$ is a vector subspace of $\mathcal{F}(V_1; V_2)$.

Let $\mathcal{F}_0(\Omega; V)$ denote the set of all functions $f : \Omega \to V$ whose support

$$\mathrm{supp}(f) \doteq \{\omega \in \Omega \ : \ f(\omega) \neq 0\} \subseteq \Omega$$

is finite. Remark that $\mathcal{F}_0(\Omega; V)$ is a vector subspace of $\mathcal{F}(\Omega; X)$. For notational simplicity, we identify any element $\omega \in \Omega$ with the function that takes the value $1 \in \mathbb{K}$ in ω and $0 \in \mathbb{K}$ else, i.e., the characteristic function $\chi_{\{\omega\}} \in \mathcal{F}_0(\Omega; \mathbb{K})$ of the set $\{\omega\}$.

We now define the quotient of a vector space by one of its subspace:

Definition 7.7 (Quotient of a Vector Space by a Subspace) Let V be any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $V' \subseteq V$ an arbitrary vector subspace.

(i) Two vectors $v_1, v_2 \in V$ are "V'-equivalent" whenever $v_1 - v_2 \in V'$. It is easy to see that the V'-equivalence is a congruence relation for the vector space structure $(V, +, \cdot)$, i.e., an equivalence relation in V that is compatible with the vector space operations[2].

(ii) V/V' denotes the set of all V'-equivalence classes in V, and $[v] \in V/V'$ the V'-equivalence class of the vector $v \in V$. As the V'-equivalence is a congruence relation, there are two unique operations $+ : V/V' \times V/V' \to V/V'$ (sum) and $\cdot : \mathbb{K} \times V/V' \to V/V'$ (scalar multiplication) for which $[v_1] + [v_2] = [v_1 + v_2]$ and $\alpha[v_1] = [\alpha v_1]$ for any $v_1, v_2 \in V$ and $\alpha \in \mathbb{K}$. In particular, $(V/V', +, \cdot)$ is a $\mathbb{K}$-vector space, called the "quotient of V by V'." $\mathfrak{q}$ denotes the linear transformation $v \mapsto [v]$ from V to V/V'.

Exercise 7.8 Let V_1 and V_2 be arbitrary vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\Theta : V_1 \to V_2$ a linear transformation. Let $V_1' \subseteq \ker(\Theta) \subseteq V_1$ be any subspace of the kernel of Θ. Show that there is a unique linear transformation $\bar{\Theta} : V_1/V_1' \to V_2$ such that $\Theta = \bar{\Theta} \circ \mathfrak{q}$, i.e., $\Theta(v_1) = \bar{\Theta}([v_1])$ for any $v_1 \in V_1$. Here, $\mathfrak{q}$ is the linear transformation $v \mapsto [v]$ from V to V/V' as in the last definition. Note that if $V_1' = \ker(\Theta)$, then $\bar{\Theta}$ is injective, by construction.

We now introduce (algebraic) tensor products of vector spaces and discuss their basic properties. In fact, these spaces are very important in quantum theory. For more details on this subject, as well as complete proofs, we recommend, for instance, [2].

[2] For all $v_1, v_2, v_3, v_4 \in V$ and $\alpha \in \mathbb{K}$, $v_1 + v_2$ is V'-equivalent to $v_3 + v_4$ and αv_1 to αv_3, whenever v_1 and v_2 are, respectively, V'-equivalent to v_3 and v_4, respectively.

Definition 7.9 (Tensor Products of Vector Spaces) Let $V_1, \ldots, V_n$ be $n \in \mathbb{N}$ vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then, $M(V_1, \ldots, V_n)$ denotes the ($\mathbb{K}$-vector) space of multilinear forms on $V_1 \times \cdots \times V_n$, i.e., mappings $V_1 \times \cdots \times V_n \to \mathbb{K}$ for which, for all $k \in \{1, \ldots, n\}$ and $v_1 \in V_1, \ldots, v_n \in V_n$, the mapping[3]

$$M(v_1, \ldots, v_{k-1}, (\cdot), v_{k+1}, \ldots, v_n) : V_K \to \mathbb{K}$$

is linear. For all $v_1 \in V_1, \ldots, v_n \in V_n$, let $v_1 \otimes \cdots \otimes v_n \in M(V_1', \ldots, V_n')$ be defined by

$$v_1 \otimes \cdots \otimes v_n(\varphi_1, \ldots, \varphi_n) \doteq \varphi_1(v_1) \cdots \varphi_n(v_n) , \quad v_1 \in V_1, \ldots, v_n \in V_n, \\ \varphi_1 \in V_1', \ldots, \varphi_n \in V_n' ,$$

where we recall that $V_1', \ldots, V_n'$ are the dual vector spaces of $V_1, \ldots, V_n$. The "tensor product" of the vector spaces $V_1, \ldots, V_n$ is the following subspace of $M(V_1', \ldots, V_n')$:

$$V_1 \otimes \cdots \otimes V_n \doteq \mathrm{span}\{v_1 \otimes \cdots \otimes v_n \ : \ v_1 \in V_1, \ldots, v_n \in V_n\} \subseteq M(V_1', \ldots, V_n') .$$

The following universal property of tensor products is well-known:

Proposition 7.10 *Let $V, V_1, \ldots, V_n$, $n \in \mathbb{N}$, be vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and let $f : V_1 \times \cdots \times V_n \to V$ be any multilinear transformation, i.e., for all $k \in \{1, \ldots, n\}$ and $v_1 \in V_1, \ldots, v_n \in V_n$, the mapping*[4]

$$f(v_1, \ldots, v_{k-1}, (\cdot), v_{k+1}, \ldots, v_n) : V_K \to V$$

is linear. There is a unique linear transformation $\Xi : V_1 \otimes \cdots \otimes V_n \to V$ such that

$$f(v_1, \ldots, v_n) = \Xi(v_1 \otimes \cdots \otimes v_n) , \qquad v_1 \in V_1, \ldots, v_n \in V_n .$$

Note here that the mapping $(v_1, \ldots, v_n) \mapsto v_1 \otimes \cdots \otimes v_n$ from $V_1 \times \cdots \times V_n$ to $V_1 \otimes \cdots \otimes V_n$ is multilinear.

As a simple application of the above proposition, note that, for all elements $m \in M(V_1'', \ldots, V_n'')$ in the space of multilinear forms on $V_1'' \times \cdots \times V_n''$, there is a unique linear mapping

$$V_1 \otimes \cdots \otimes V_n \to \mathbb{K} ,$$

i.e., a linear functional on $V_1 \otimes \cdots \otimes V_n$ also denoted by m, such that

$$m(v_1 \otimes \cdots \otimes v_n) = m(v_1, \ldots, v_n) , \qquad v_1 \in V_1 \subseteq V_1'', \ldots, v_n \in V_n \subseteq V_n'' .$$

[3] Of course, for $k = 1$, it means that $M((\cdot), v_2, \ldots, v_n)$ is linear. Mutatis mutandis for $k = n$.

[4] As already said, for $k = 1$, it means that $f((\cdot), v_2, \ldots, v_n)$ is linear. Mutatis mutandis for $k = n$.

In particular, one has the canonical inclusion

$$V_1' \otimes \cdots \otimes V_n' \subseteq (V_1 \otimes \cdots \otimes V_n)' \; .$$

With these observations, for all $\varphi_1 \in V_1', \ldots, \varphi_n \in V_n'$, $\varphi_1 \otimes \cdots \otimes \varphi_1$ is naturally seen as an element of the dual vector space $(V_1 \otimes \cdots \otimes V_n)'$.

From the last proposition, one can also give a meaning of the tensor product of linear transformations:

Corollary 7.11 *Let $V_1, \ldots, V_n$ and $W_1, \ldots, W_n$ be vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, where $n \in \mathbb{N}$. For any n linear transformations $\Theta_1 \in \mathcal{L}(V_1; W_1), \ldots, \Theta_n \in \mathcal{L}(V_n; W_n)$, there is a unique* linear *transformation*

$$\Theta_1 \otimes \cdots \otimes \Theta_n : V_1 \otimes \cdots \otimes V_n \to W_1 \otimes \cdots \otimes W_n$$

satisfying

$$\Theta_1 \otimes \cdots \otimes \Theta_n (v_1 \otimes \cdots \otimes v_n) = \Theta_1(v_1) \otimes \cdots \otimes \Theta_n(v_n) \; , \qquad v_1 \in V_1, \ldots, v_n \in V_n \; .$$

The next proposition refers to (Hamel) bases of tensor products and is also well-known:

Proposition 7.12 (Hamel Bases for Tensor Products) *Let $V_1, \ldots, V_n$ be $n \in \mathbb{N}$ vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and let $B_1 \subseteq V_1, \ldots, B_n \subseteq V_n$ be (Hamel) bases of these spaces. Then*

$$\{v_1 \otimes \cdots \otimes v_n : \; v_1 \in B_1, \ldots, v_n \in B_n\} \subseteq V_1 \otimes \cdots \otimes V_n$$

is a (Hamel) basis of the tensor product $V_1 \otimes \cdots \otimes V_n$.

The following corollary says that the tensor product of vector spaces is an associative operation, which is an important, albeit simple, consequence of the last proposition:

Corollary 7.13 (Associativity of the Tensor Product) *Let V_1, V_2, V_3 be three vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then, the unique linear mappings*

$$\begin{array}{rcl} V_1 \otimes V_2 \otimes V_3 & \to & (V_1 \otimes V_2) \otimes V_3 \\ v_1 \otimes v_2 \otimes v_3 & \mapsto & (v_1 \otimes v_2) \otimes v_3 \end{array} \quad \textit{and} \quad \begin{array}{rcl} V_1 \otimes V_2 \otimes V_3 & \to & V_1 \otimes (V_2 \otimes V_3) \\ v_1 \otimes v_2 \otimes v_3 & \mapsto & v_1 \otimes (v_2 \otimes v_3) \end{array}$$

for any $v_1 \in V_1$, $v_2 \in V_2$, $v_3 \in V_3$, are bijective. Thus, the three spaces as well as their elements $(v_1 \otimes v_2) \otimes v_3$, $v_1 \otimes (v_2 \otimes v_3)$, and $v_1 \otimes v_2 \otimes v_3$ are canonically identified with each other.

From Proposition 7.12, we also deduce that the tensor product of spaces of linear operators can be identified with a subspace of linear operators acting on some tensor products of vector spaces:

Corollary 7.14 *Let* $V_1, \ldots, V_n$ *and* $W_1, \ldots, W_n$ *be vector spaces over* $\mathbb{K} = \mathbb{R}, \mathbb{C}$*, where* $n \in \mathbb{N}$*. The unique linear mapping*

$$\mathcal{L}(V_1; W_1) \otimes \cdots \otimes \mathcal{L}(V_n; W_n) \to \mathcal{L}(V_1 \otimes \cdots \otimes V_n; W_1 \otimes \cdots \otimes W_n)$$

satisfying

$$\Theta_1 \otimes \cdots \otimes \Theta_n \mapsto \Theta_1 \otimes \cdots \otimes \Theta_n \ , \qquad \Theta_1 \in \mathcal{L}(V_1; W_1), \ldots, \Theta_n \in \mathcal{L}(V_n; W_n) \ ,$$

is injective. If the vector spaces $V_1, \ldots, V_n, W_1, \ldots, W_n$ *are finite dimensional, then this mapping is bijective.*

Proof The existence and uniqueness of the mapping are a direct consequence of the universal property of tensor products (see Proposition 7.10), by observing that the mapping

$$(\Theta_1, \ldots, \Theta_n) \mapsto \Theta_1 \otimes \cdots \otimes \Theta_n$$

from $\mathcal{L}(V_1; W_1) \times \cdots \times \mathcal{L}(V_n; W_n)$ to $\mathcal{L}(V_1 \otimes \cdots \otimes V_n; W_1 \otimes \cdots \otimes W_n)$ is multilinear. By noting that for all $\Theta_1 \in \mathcal{L}(V_1; W_1), \ldots, \Theta_n \in \mathcal{L}(V_n; W_n)$, all $v_1 \in V_1, \ldots, v_n \in V_n$, and all $\varphi_1 \in W_1', \ldots, \varphi_n \in W_n'$, one has

$$\varphi_1 \otimes \cdots \otimes \varphi_n(\Theta_1 \otimes \cdots \otimes \Theta_n(v_1 \otimes \cdots \otimes v_n)) = \varphi_1(\Theta_1(v_1)) \cdots \varphi_n(\Theta_n(v_n)) \ ,$$

the injectivity of the mapping can be proven by a simple adaptation of the proof of Lemma 4.157. Recall that $\varphi_1 \otimes \cdots \otimes \varphi_n \in W_1' \otimes \cdots \otimes W_n'$ is canonically seen as a linear form on $W_1 \otimes \cdots \otimes W_n$. If the vector spaces $V_1, \ldots, V_n, W_1, \ldots, W_n$ are finite dimensional, then, by Proposition 7.12, both vector spaces $\mathcal{L}(V_1; W_1) \otimes \cdots \otimes \mathcal{L}(V_n; W_n)$ and $\mathcal{L}(V_1 \otimes \cdots \otimes V_n; W_1 \otimes \cdots \otimes W_n)$ have the same dimension, and, thus, the mapping is bijective. □

We now define the usual notion of algebras, i.e., vector spaces endowed with a product:

Definition 7.15 (Algebra)

(i) The structure $(\mathcal{A}, +, \cdot, \circ)$ is, by definition, an "algebra" over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ if $(\mathcal{A}, +, \cdot)$ is a vector space over $\mathbb{K}$ and the binary operation[5] $\circ : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$ (product in $\mathcal{A}$) is bilinear, i.e., for all $A_1, A_2, A_3 \in \mathcal{A}$,

$$A_1 \circ (A_2 + A_3) = A_1 \circ A_2 + A_1 \circ A_3 \ , \ (A_1 + A_2) \circ A_3 = A_1 \circ A_3 + A_2 \circ A_3 \ ,$$

while for all $\alpha \in \mathbb{K}$, $\alpha \cdot (A_1 \circ A_2) = (\alpha \cdot A_1) \circ A_2 = A_1 \circ (\alpha \cdot A_2)$.

(ii) The algebra $\mathcal{A}$ is called "commutative" if $A_1 \circ A_2 = A_2 \circ A_1$ for all $A_1, A_2 \in \mathcal{A}$.

[5] As is usual, for simplicity, for $A_1, A_2 \in \mathcal{A}$, the element $\circ(A_1, A_2) \in \mathcal{A}$ is denoted by $A_1 \circ A_2$.

(iii) The algebra $\mathcal{A}$ is said to be "associative" if, for all $A_1, A_2, A_3 \in \mathcal{A}$,

$$(A_1 \circ A_2) \circ A_3 = A_1 \circ (A_2 \circ A_3) \doteq A_1 \circ A_2 \circ A_3 \ .$$

(iv) The element $\mathbb{1} \in \mathcal{A}$ (when it exists) is called "unit" of this algebra whenever $\mathbb{1} \circ A = A \circ \mathbb{1} = A$ for all $A \in \mathcal{A}$. In this case, $\mathcal{A}$ is called a "unital algebra." If $\mathcal{A}$ is a unital algebra, then its unit is necessarily unique.

(v) The vector subspace $\mathcal{B} \subseteq \mathcal{A}$ is called a "subalgebra" of the algebra $\mathcal{A}$ when $B_1 \circ B_2 \in \mathcal{B}$ for all $B_1, B_2 \in \mathcal{B}$.

In these notes, the symbol "$\mathbb{1}$ " will always stand for the unit of any unital algebra. The algebras that we consider in this book are by default unital. In fact, any non-unital algebra can be canonically extended to a unital one:

Definition 7.16 (Unitization of an Algebra) Let $(\mathcal{A}, +, \cdot, \circ)$ be any algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and define the vector space $\tilde{\mathcal{A}} \doteq \mathbb{K} \times \mathcal{A}$ which has $\mathcal{A}$ as subspace, as in Definition 7.2. Then we define an operation $\circ : \tilde{\mathcal{A}} \times \tilde{\mathcal{A}} \to \tilde{\mathcal{A}}$ (product) as follows:

$$(\alpha_1, A_1) \circ (\alpha_2, A_2) \doteq (\alpha_1\alpha_2, \alpha_1 A_2 + \alpha_2 A_1 + A_1 A_2) \ .$$

It is easy to check that $(\tilde{\mathcal{A}}, +, \cdot, \circ)$ is a unital algebra with $\mathbb{1} = (1, 0) \in \mathbb{K} \times \mathcal{A}$.

Similar to the case of general vector spaces, again by a slight abuse of notation and for simplicity, a generic algebra $(\mathcal{A}, +, \cdot, \circ)$ is frequently denoted here by the name of the set on which its operations are defined, i.e., $\mathcal{A}$.

Let $\Omega \neq \emptyset$ be a nonempty set. The vector space $\mathcal{F}(\Omega; \mathbb{K})$ is an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ with the following product: For all $f_1, f_2 \in \mathcal{F}(\Omega; \mathbb{K})$, $f_1 \circ f_2 \in \mathcal{F}(\Omega; \mathbb{K})$ is (pointwise) defined by

$$[f_1 \circ f_2]\,(p) \doteq f_1(p) f_2(p) \ , \qquad p \in \Omega \ .$$

With this product, $\mathcal{F}(\Omega; \mathbb{K})$ is an associative and commutative unital algebra. More generally, for an arbitrary algebra $(\mathcal{A}, +, \cdot, \circ)$, we define a product $\circ$ in $\mathcal{F}(\Omega; \mathcal{A})$ by

$$[f_1 \circ f_2]\,(p) \doteq f_1(p) \circ f_2(p) \ , \qquad p \in \Omega \ .$$

Again, $\mathcal{F}(\Omega; \mathcal{A})$ is an algebra with respect to this product. However, it is generally non-commutative and non-associative. In fact, it is commutative iff $\mathcal{A}$ is itself commutative. Mutatis mutandis for the associativity or the existence of a unit. Note also that the set $\mathcal{F}_0(\Omega; \mathcal{A})$ of all functions $f : \Omega \to V$, whose support is finite, is a subalgebra of $\mathcal{F}(\Omega; \mathcal{A})$.

We give now a second important example of associative algebras, the so-called free algebras. For an arbitrary nonempty set Ω, $\tilde{\mathcal{S}}(\Omega)$ denotes the set of all finite sequences of elements of Ω. The elements of $\tilde{\mathcal{S}}(\Omega)$ are called "words on the alphabet Ω." The "empty, or zero-length, word" is denoted by $\mathbb{1}$. The elements of Ω are identified with the corresponding length-one words. We define a product

$\star : \tilde{\mathcal{S}}(\Omega) \times \tilde{\mathcal{S}}(\Omega) \to \tilde{\mathcal{S}}(\Omega)$ by simple concatenation of words. Remark that $(\tilde{\mathcal{S}}(\Omega), \star)$ is a monoid[6] whose unit is the empty word $\imath \in \tilde{\mathcal{S}}(\Omega)$. The set of nonempty words is denoted by $\mathcal{S}(\Omega) \doteq \tilde{\mathcal{S}}(\Omega) \backslash \{\imath\}$. Note that the product $\star$ preserves $\mathcal{S}(\Omega)$, i.e.,

$$\mathcal{S}(\Omega) \star \mathcal{S}(\Omega) \subseteq \mathcal{S}(\Omega) ,$$

and thus defines a binary operation (also called a product and denoted by $\star$) on $\mathcal{S}(\Omega)$, by restriction. In contrast to $\tilde{\mathcal{S}}(\Omega)$, $\mathcal{S}(\Omega)$ is not a monoid anymore.

These objects are used to define free algebras as follows:

Definition 7.17 (Unital Free Algebra) Let $\Omega \neq \emptyset$ be a nonempty set and $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Let $\tilde{F}(\Omega; \mathbb{K})$ be the vector space $\mathcal{F}_0(\tilde{\mathcal{S}}(\Omega); \mathbb{K})$ of $\mathbb{K}$-valued functions on words. There is a unique operation

$$\star : \tilde{F}(\Omega; \mathbb{K}) \times \tilde{F}(\Omega; \mathbb{K}) \to \tilde{F}(\Omega; \mathbb{K})$$

extending the product of the monoid $(\tilde{\mathcal{S}}(\Omega), \star)$ in such a way that $(\tilde{F}(\Omega; \mathbb{K}), +, \cdot, \star)$ is an algebra, where $+, \cdot$ are the usual vector space operations of $(\mathcal{F}_0(\tilde{\mathcal{S}}(\Omega); \mathbb{K}) +, \cdot)$. This algebra is associative and unital, $\imath \in \Omega \subseteq \tilde{\mathcal{S}}(\Omega) \subseteq \tilde{F}(\Omega; \mathbb{K})$ being its unit, but it is generally non-commutative. It is known as the "unital free algebra over $\mathbb{K}$ generated by Ω" and $\star$ is the "free product" for $\tilde{F}(\Omega; \mathbb{K})$.

The non-unital case is defined in a similar way by replacing the set $\tilde{\mathcal{S}}(\Omega)$ of all words with the set $\mathcal{S}(\Omega)$ of nonempty words:

Definition 7.18 (Non-Unital Free Algebra) Let $\Omega \neq \emptyset$ be a nonempty set and $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Let $\mathfrak{F}(\Omega; \mathbb{K})$ be the vector space $\mathcal{F}_0(\mathcal{S}(\Omega); \mathbb{K})$ of $\mathbb{K}$-valued functions on nonempty words. As before, there is a unique operation

$$\star : \mathfrak{F}(\Omega; \mathbb{K}) \times \mathfrak{F}(\Omega; \mathbb{K}) \to \mathfrak{F}(\Omega; \mathbb{K})$$

extending the product of $(\mathcal{S}(\Omega), \star)$ in such a way that $(\mathfrak{F}(\Omega; \mathbb{K}), +, \cdot, \star)$ is an algebra, where $+, \cdot$ are the usual vector space operations of $(\mathcal{F}_0(\mathcal{S}(\Omega); \mathbb{K}) +, \cdot)$. This algebra is associative, but not unital and generally non-commutative. It is known as the "free algebra over $\mathbb{K}$ generated by Ω" and $\star$ is the "free product" for $\mathfrak{F}(\Omega; \mathbb{K})$.

In fact, the unital free algebra $\tilde{F}(\Omega; \mathbb{K})$ can be canonically identified with the unitization of the free algebra $\mathfrak{F}(\Omega; \mathbb{K})$. See Definitions 7.16 and 7.20 and Exercise 7.22.

A third important example of an associative unital algebra is the vector space $\mathcal{L}(V)$ of linear transformations $V \to V$, where V is any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For all $\Theta_1, \Theta_2 \in \mathcal{L}(V)$, the product $\Theta_1 \circ \Theta_2 \in \mathcal{L}(V)$ is defined by composition of mappings:

[6] A monoid $\mathcal{X}$ is a set endowed with an associative bilinear operation $\star : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ that has a unit.

$$[\Theta_1 \circ \Theta_2](v) \doteq \Theta_1(\Theta_2(v)) , \qquad v \in V .$$

Note that $\mathcal{L}(V)$ is a non-commutative algebra, whenever the dimension of V is bigger than one.

Tensor products of vector spaces that are algebras have a natural algebra structure:

Proposition 7.19 (Tensor Products of Algebras) *Let* $(\mathcal{A}_1, +, \cdot, \circ), \ldots, (\mathcal{A}_n, +, \cdot, \circ)$ *be* $n \in \mathbb{N}$ *algebras over* $\mathbb{K} = \mathbb{R}, \mathbb{C}$. *There is a unique product in* $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$*, i.e., a bilinear mapping*

$$\circ : (\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n) \times (\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n) \to \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n ,$$

such that

$$(A_1 \otimes \cdots \otimes A_n) \circ (A'_1 \otimes \cdots \otimes A'_n) = (A_1 \circ A'_1) \otimes \cdots \otimes (A_n \circ A'_n) ,$$
$$A_1, A'_1 \in \mathcal{A}_1, \ldots, A_n, A'_n \in \mathcal{A}_n .$$

Additionally, $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ *is a commutative (associative) algebra if* $\mathcal{A}_1, \ldots, \mathcal{A}_n$ *are commutative (associative) algebras. If the algebras* $\mathcal{A}_1, \ldots, \mathcal{A}_n$ *are unital, then* $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ *is unital, and its unit is* $\mathbb{1} \otimes \cdots \otimes \mathbb{1}$.

Proof For any fixed $A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_n$, by universality of tensor products, there is a unique linear mapping $L_{A_1,\ldots,A_n} \in \mathcal{L}(\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n)$, such that

$$L_{A_1,\ldots,A_n}(A'_1 \otimes \cdots \otimes A'_n) = (A_1 \circ A'_1) \otimes \cdots \otimes (A_n \circ A'_n) , \qquad A'_1 \in \mathcal{A}_1, \ldots, A'_n \in \mathcal{A}_n .$$

Again by universality of tensor products (see Proposition 7.10), there is a unique ("left multiplication") linear mapping

$$L : \mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n \to \mathcal{L}(\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n) ,$$

such that

$$L(A_1 \otimes \cdots \otimes A_n)(A'_1 \otimes \cdots \otimes A'_n) = (A_1 \circ A'_1) \otimes \cdots \otimes (A_n \circ A'_n) .$$

The bilinear mapping $L(\cdot)(\cdot)$ on $(\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n) \times (\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n)$ has the required properties for the product. The second part of the proposition is clear. □

We now define algebra homomorphisms and representations:

Definition 7.20 (Homomorphism and Algebra Representation) Let $(\mathcal{A}_1, +, \cdot, \circ)$ and $(\mathcal{A}_2, +, \cdot, \circ)$ be two algebras over $\mathbb{K} = \mathbb{R}, \mathcal{C}$.

(i) A linear transformation $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ is a "(algebra) homomorphism" from $\mathcal{A}_1$ to $\mathcal{A}_2$ whenever

$$\Theta(A \circ A') = \Theta(A) \circ \Theta(A') , \qquad A, A' \in \mathcal{A}_1 .$$

It is, by definition, an "(algebra) isomorphism" when it is a one-to-one correspondence (i.e., a bijection). In this case, $\mathcal{A}_1$ and $\mathcal{A}_2$ are said to be "isomorphic."

(ii) An algebra homomorphism $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ is a "representation" of the algebra $\mathcal{A}_1$ on the vector space V if $\mathcal{A}_2 = \mathcal{L}(V)$. An algebra representation is said to be "faithful" if it is one-to-one (i.e., injective).

(iii) If $\mathcal{A}_1$ and $\mathcal{A}_2$ are both unital and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ is an algebra homomorphism mapping the unit of $\mathcal{A}_1$ to the unit of $\mathcal{A}_2$, then Θ is said to be "unital."

Note that if $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ is an isomorphism of two unital algebras, $\mathcal{A}_1$ and $\mathcal{A}_2$, then it is automatically a unital algebra homomorphism.

Exercise 7.21 (Unitization of an Algebra Homomorphism) Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two algebras over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, $\mathcal{A}_2$ unital, and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ an algebra homomorphism. Show that there is a unique unital algebra homomorphism $\tilde{\Theta} : \tilde{A}_1 \to \mathcal{A}_2$ extending Θ to the unitization $\tilde{A}_1$ of $\mathcal{A}_1$.

Exercise 7.22 Let $\Omega \neq \emptyset$ be a nonempty set and $\mathcal{A}$ an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. If $\mathcal{A}$ has no unit, prove that, for any mapping $i : \Omega \to \mathcal{A}$, there is a unique algebra homomorphism $\mathfrak{F}(\Omega; \mathbb{K}) \to \mathcal{A}$ extending i. Otherwise, prove the same property with $\tilde{F}(\Omega; \mathbb{K})$ instead of $\mathfrak{F}(\Omega; \mathbb{K})$, the algebra homomorphism $\tilde{F}(\Omega; \mathbb{K}) \to \mathcal{A}$ being now unital. Recall that one has the canonical inclusions $\Omega \subseteq \mathcal{S}(\Omega) \subseteq \mathfrak{F}(\Omega; \mathbb{K})$ and $\Omega \subseteq \tilde{\mathcal{S}}(\Omega) \subseteq \tilde{F}(\Omega; \mathbb{K})$. This algebra homomorphism is in any case denoted again by i.

The above exercise demonstrates that the properties of $\mathfrak{F}(\Omega; \mathbb{K})$ or $\tilde{F}(\Omega; \mathbb{K})$ uniquely determine the (possibly unital) free algebra, up to an algebra isomorphism.

We define now the so-called center and ideals of algebras as follows:

Definition 7.23 (Center of an Algebra) Let $\mathcal{A}$ be an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We define its "center" by

$$\mathcal{Z}(\mathcal{A}) \doteq \{A \in \mathcal{A} \ : \ [A, A'] = 0 \text{ for all } A' \in \mathcal{A}\} ,$$

where $[A, A']$ denotes the so-called commutator

$$[A, A'] \doteq AA' - A'A \in \mathcal{A} .$$

Observe that the center of a subalgebra is one of its subalgebras.

Definition 7.24 (Ideal of an Algebra) Let $(\mathcal{A}, +, \cdot, \circ)$ be an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

(i) The vector subspace $\mathcal{I} \subseteq \mathcal{A}$ is a "left (right) ideal" if $\mathcal{A} \circ \mathcal{I} \subseteq \mathcal{I}$ ($\mathcal{I} \circ \mathcal{A} \subseteq \mathcal{I}$), where

$$\mathcal{A} \circ \mathcal{I} \doteq \{A \circ I : I \in \mathcal{I},\ A \in \mathcal{A}\}$$

(mutatis mutandis for $\mathcal{I} \circ \mathcal{A}$).

(ii) $\mathcal{I}$ is called an "ideal" of $\mathcal{A}$ if it is simultaneously a left and a right ideal of $\mathcal{A}$.
(iii) Note that $\{0\}$ and $\mathcal{A}$ itself are trivially ideals of $\mathcal{A}$. If they are the only ideals of $\mathcal{A}$, we say that the algebra is "simple."

Exercise 7.25 Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Prove that, for any dimension $n \in \mathbb{N}$, the algebra $\mathcal{L}(\mathcal{K}^n)$ of all linear operators $\mathbb{K}^n \to \mathbb{K}^n$ is simple.

Note that any ideal of an algebra is a subalgebra, but a subalgebra is not necessarily an ideal. Observe also that, if $\mathcal{A}$ is a non-unital algebra with its unitization denoted by $\tilde{\mathcal{A}}$, then $\mathcal{A}$ is not only a subalgebra of $\tilde{\mathcal{A}}$ but also an ideal. See Definition 7.16. Another example of non-trivial algebra ideals is given by the kernel of algebra homomorphisms (Definition 7.6):

Exercise 7.26 Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two algebras over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ an algebra homomorphism. Show that $\ker(\Theta)$ is an ideal of $\mathcal{A}_1$.

The next exercise shows that congruence relations for algebra structures can be defined via ideals:

Exercise 7.27 Let $\mathcal{A}$ be an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\mathcal{I} \subseteq \mathcal{A}$ an ideal. Recall that the $\mathcal{I}$-equivalence of elements of $\mathcal{A}$ is a congruence relation for the vector space structure $(\mathcal{A}, +, \cdot)$, $\mathcal{I}$ being a vector subspace, by definition of an ideal. See Definition 7.7. Show that it is also a congruence relation for the algebra structure $(\mathcal{A}, +, \cdot, \circ)$: for all $A_1, A_2, A_3, A_4 \in \mathcal{A}$, $A_1 \circ A_2$ is $\mathcal{I}$-equivalent to $A_3 \circ A_4$, whenever A_1 and A_2 are, respectively, $\mathcal{I}$-equivalent to A_3 and A_4.

This exercise allows us to define a new algebra via the quotient of an algebra by one of its ideals: Given an algebra $\mathcal{A}$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and an ideal $\mathcal{I} \subseteq \mathcal{A}$, recall that $\mathcal{A}/\mathcal{I}$ is a $\mathbb{K}$-vector space, $\mathcal{I}$ being a subspace of the vector space $\mathcal{A}$. See Definition 7.7. Then, because the $\mathcal{I}$-equivalence is a congruence relation for the algebra structure $(\mathcal{A}, +, \cdot, \circ)$, we obtain the following definition for quotients on algebras:

Definition 7.28 (Quotient of an Algebra by an Ideal) Let $\mathcal{A}$ be an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\mathcal{I} \subseteq \mathcal{A}$ an ideal. By Exercise 7.27, there is a unique operation, $\circ : \mathcal{A}/\mathcal{I} \times \mathcal{A}/\mathcal{I} \to \mathcal{A}/\mathcal{I}$ (product) for which $[A_1] \circ [A_2] = [A_1 \circ A_2]$ for any $A_1, A_2 \in \mathcal{A}$. Then, $(\mathcal{A}/\mathcal{I}, +, \cdot, \circ)$ is an algebra over $\mathbb{K}$, named the "quotient of the algebra $\mathcal{A}$ by its ideal $\mathcal{I}$. " Remark that $\mathcal{A}/\mathcal{I}$ is unital whenever $\mathcal{A}$ is unital. Mutatis mutandis for an associative or commutative algebra $\mathcal{A}$.

Exercise 7.29 Let $\mathcal{A}$ be an algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\mathcal{I} \subseteq \mathcal{A}$ an ideal. Show that the mapping $\mathfrak{q} : A \mapsto [A]$ from the algebra $\mathcal{A}$ to the quotient algebra $\mathcal{A}/\mathcal{I}$ is an algebra homomorphism. Show, additionally, that if $\mathcal{A}_1$ and $\mathcal{A}_2$ are two algebras over $\mathbb{K}$ and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ an algebra homomorphism with $\mathcal{I} \subseteq \ker(\Theta)$, then the mapping $\bar{\Theta} : \mathcal{A}_1/\mathcal{I} \to \mathcal{A}_2$, as defined in Exercise 7.8, is an algebra homomorphism, which is faithful when $\mathcal{I} = \ker(\Theta)$. Recall that $\ker(\Theta) \subseteq \mathcal{A}_1$ is an ideal, Θ being an algebra homomorphism.

Recall that if $\mathcal{A}$ is a non-unital algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ with its unitization denoted by $\tilde{\mathcal{A}}$, then $\mathcal{A}$ is an ideal of $\tilde{\mathcal{A}}$. Then, the quotient $\tilde{\mathcal{A}}/\mathcal{A}$ can be canonically identified with $\mathbb{K}$, seen as an algebra over $\mathbb{K}$.

Exercise 7.30 Let $\mathcal{A}$ be a non-unital algebra over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Prove that there exists at least one non-zero, multiplicative, linear functional $\rho : \mathcal{A} \tilde{\to} \mathbb{K}$, i.e., a non-zero linear functional for which $\rho(A_1 A_2) = \rho(A_1)\rho(A_2)$ for any $A_1, A_2 \in \mathcal{A}$.

We now continue our discussions on vector spaces and algebras by introducing an additional structure, namely, a (semi)norm on these spaces:

Definition 7.31 (Normed Space) Let $(X, +, \cdot)$ be a vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

(i) A "seminorm" in X is a mapping $\|\cdot\| : X \to \mathbb{R}_0^+ \doteq [0, \infty)$ such that

$$\|\alpha \cdot x\| = |\alpha| \, \|x\| \text{ (homogeneity of degree one)} \quad \text{and}$$
$$\|x + x'\| \leq \|x\| + \|x'\| \text{ (subadditivity)}$$

for all $x, x' \in X$ and $\alpha \in \mathbb{K}$.

(ii) A "norm" in X is a seminorm $\|\cdot\|$ (in X) that is "non-degenerated," i.e., for all $x \in X$, $\|x\| = 0 \in \mathbb{R}_0^+$ iff $x = 0 \in X$. In this case, $(X, +, \cdot, \|\cdot\|)$ is, by definition, a "normed space." Any vector subspace of a normed space is canonically a normed space, by restriction of the norm.

(iii) If $(X_1, +, \cdot, \|\cdot\|_{X_1})$ and $(X_2, +, \cdot, \|\cdot\|_{X_2})$ are two normed spaces for which there is some norm-preserving[7] linear one-to-one correspondence[8] $\Theta \in \mathcal{L}(X_1; X_2)$, then X_1 and X_2 are said to be "isomorphic" normed spaces. Such a linear transformation Θ is called an "isomorphism of normed spaces."

(iv) If $(X, +, \cdot, \|\cdot\|)$ is a normed space, an arbitrary subspace $Y \subseteq X$ is said to be "closed" whenever, for all $x \in X$,

$$\inf\{\|x - y\| \ : \ y \in Y\} = 0 \qquad \text{iff} \qquad x \in Y \, .$$

Once again, by a slight abuse of notation and for simplicity, a generic normed (vector) space $(X, +, \cdot, \|\cdot\|)$ is often denoted here by the pair $(X, \|\cdot\|)$ or, even simpler, by X. By default, we see any vector subspace Y of a normed space X as a normed space, by restricting the norm of X to $Y \subseteq X$.

One trivial example of a normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ is the field $\mathbb{K}$ itself with the corresponding absolute value as a norm:

$$\|x\| \doteq |x|, \qquad x \in \mathbb{K} \, .$$

Another elementary example is provided by (finite-dimensional) Euclidean spaces:

[7] That is, $\|\Theta(x_1)\|_{X_2} = \|x_1\|_{X_1}$ for all $x_1 \in X_1$.

[8] That is, the linear transformation Θ is additionally bijective.

Definition 7.32 (Euclidean Space) For any fixed (finite) dimension $D \in \mathbb{N}$, the "Euclidean norm" in $\mathbb{R}^D$ (as a real vector space) is defined by

$$\|\mathbf{x}\|_{\mathrm{e}} \doteq \sqrt{|x_1|^2 + \cdots + |x_D|^2} , \qquad \mathbf{x} = (x_1, \ldots, x_D) \in \mathbb{R}^D .$$

The Euclidean norm for the complex vector space $\mathbb{C}^D$ is defined in the same way.

Let $\Omega \neq \emptyset$ be a nonempty set and $(X, \|\cdot\|)$ a normed space. For all $f \in \mathcal{F}(\Omega; X)$, we define the "supremum norm" of f by

$$\|f\|_{\infty} \doteq \sup_{p\in\Omega} \|f(p)\| \in [0, \infty] .$$

Let

$$\mathcal{F}_{\mathrm{b}}(\Omega; X) \doteq \left\{ f \in \mathcal{F}(\Omega; X) \ : \ \|f\|_{\infty} < \infty \right\} . \tag{7.1}$$

The elements of $\mathcal{F}_{\mathrm{b}}(\Omega; X)$ are called "bounded functions" from Ω to X . Observe that $\mathcal{F}_{\mathrm{b}}(\Omega; X)$ is a vector subspace of $\mathcal{F}(\Omega; X)$. Similarly, $\mathcal{F}_{\mathrm{b}}(\Omega; \mathbb{K})$ is a subalgebra of the algebra $\mathcal{F}(\Omega; \mathbb{K})$, the (pointwise) product of which is defined above. Note further that $(\mathcal{F}_{\mathrm{b}}(\Omega; X), \|\cdot\|_{\infty})$ is a normed space and that $\mathcal{F}_0(\Omega; X)$ is a subspace of $\mathcal{F}_{\mathrm{b}}(\Omega; X)$.

Using Definition 7.2, we define now the Cartesian product of normed spaces as follows:

Definition 7.33 (Cartesian Products of Normed Spaces) If $(X_1, \|\cdot\|_1), \ldots,$ $(X_N, \|\cdot\|_N)$, $N \in \mathbb{N}$, are normed spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ then the Cartesian product $X_1 \times \cdots \times X_N$ is (canonically) a normed space with the norm

$$\|(x_1, \ldots, x_N)\|_{X_1\times\cdots\times X_N} \doteq \max\{\|x_1\|_1 , \ldots, \|x_N\|_N\} ,$$
$$(x_1, \ldots, x_N) \in X_1 \times \cdots \times X_N .$$

By construction, for any $k \in \{1, \ldots, N\}$, the restriction of $\|\cdot\|_{X_1\times\cdots\times X_N}$ to the subspace $X_k \subseteq X_1 \times \cdots \times X_N$ is the original norm $\|\cdot\|_k$. Note however that this property does not completely determine the norm $\|\cdot\|_{X_1\times\cdots\times X_N}$. In fact, in some situations, other extensions of the norms $\|\cdot\|_k$, $k = 1, \ldots, N$, to the Cartesian product $X_1 \times \cdots \times X_N$ are more convenient than the norm defined above, which is only the one that is used by default.

Using the quotient of a vector space by one of its subspace (Definition 7.7), one shows that the quotient of a seminormed space by a subspace is naturally a seminormed space:

Exercise 7.34 Let X be any vector space, $\|\cdot\|$ a seminorm in X, and $Y \subseteq X$ a vector subspace. Define the mapping $\|\cdot\|_{X/Y} : X/Y \to \mathbb{R}_0^+$ by

$$\|[x]\|_{X/Y} \doteq \inf\left\{ \left\|x'\right\|_X \ : \ x' \in [x] \right\} , \qquad x \in X .$$

(i) Prove that $\|\cdot\|_{X/Y}$ is a seminorm in the vector space X/Y.
(ii) Show that if $\|\cdot\|$ is a norm and Y is closed (in the normed space $(X, \|\cdot\|)$), then $\|\cdot\|_{X/Y}$ is also a norm.
(iii) Prove that $\|\cdot\|$ satisfies the parallelogram identity (see Sect. 7.3.1) only if $\|\cdot\|_{X/Y}$ also satisfies this identity.

The following construction is standard to obtain a normed space from a seminormed one:

Exercise 7.35 Let X be any vector space and $\|\cdot\|$ a seminorm in X. Show that

$$X_0 \doteq \{x \in X \ : \ \|x\| = 0\}$$

is a vector subspace of X and that $\|\cdot\|_{X/X_0}$ is a norm in the vector space X/X_0.

We now define bounded linear operators between (semi)normed spaces as follows:

Definition 7.36 (Bounded Linear Operators) Let $(X_1, \|\cdot\|_{X_1})$ be a seminormed space and $(X_2, \|\cdot\|_{X_2})$ another normed (seminormed) space, both over $\mathbb{K} = \mathcal{R}, \mathbb{C}$. For any linear operator $\Theta \in \mathcal{L}(X_1; X_2)$, its "operator norm (seminorm)" is denoted by

$$\|\Theta\|_{\mathrm{op}} \doteq \sup_{x_1 \in X_1,\ \|x_1\|_{X_1}=1} \|\Theta(x_1)\|_{X_2} \in [0, \infty] ,$$

which is used to define the vector subspace

$$\mathcal{B}(X_1; X_2) \doteq \{\Theta \in \mathcal{L}(X_1; X_2) \ : \ \|\Theta\|_{\mathrm{op}} < \infty\} \subseteq \mathcal{L}(X_1; X_2)$$

of "bounded" linear operators from $(X_1, \|\cdot\|_{X_1})$ to $(X_2, \|\cdot\|_{X_2})$. We say that $\Theta \in \mathcal{B}(X_1; X_2)$ is a "contraction" if $\|\Theta\|_{\mathrm{op}} \leq 1$. Observe that $\|\cdot\|_{\mathrm{op}}$ is a norm (seminorm) in $\mathcal{B}(X_1; X_2)$ and that $\mathcal{B}(X_1; X_2)$ is a vector subspace of $\mathcal{L}(X_1; X_2)$. If $X_1 = X_2$, we use the short notation $\mathcal{B}(X_1)$ for $\mathcal{B}(X_1; X_2)$.

Note that the use of the term "bounded" for elements of $\mathcal{B}(X_1; X_2)$ does not have the same meaning as for elements of $\mathcal{F}_{\mathrm{b}}(X_1; X_2)$: In fact, the *bounded operator* $\Theta \in \mathcal{B}(X_1; X_2)$ is an element of $\mathcal{F}_{\mathrm{b}}(X_1; X_2)$, i.e., it is a *bounded function* $X_1 \to X_2$, iff $\Theta = 0$.

Exercise 7.37 Let $(X, \|\cdot\|)$ be any seminormed space and $Y \subseteq X$ a subspace. Show that the linear transformation $\mathfrak{q} : x \mapsto [x]$ from X to the seminormed space $(X/Y, \|\cdot\|_{X/Y})$ is a contraction.

Exercise 7.38 Let X and Y be two seminormed spaces and $X_0 \subseteq X$ a subspace. Let $\Theta \in \mathcal{B}(X; Y)$ be a bounded linear transformation with $X_0 \subseteq \ker(\Theta)$. Show that $\bar{\Theta} \in \mathcal{B}(X/X_0; Y)$ with $\left\|\bar{\Theta}\right\|_{\mathrm{op}} = \|\Theta\|_{\mathrm{op}}$, where the linear transformation $\bar{\Theta} : X/X_0 \to Y$ is defined as in Exercise 7.8.

We define now the dual space of a normed space, as itself a normed space:

Definition 7.39 (Dual of a Normed Space) Let $(X, \|\cdot\|)$ be any normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. The vector space

$$X^{\mathrm{td}} \doteq \mathcal{B}(X; \mathbb{K})$$

is called the "(topological) dual" of X . It is canonically a normed space, the norm of which is $\|\cdot\|_{\mathrm{op}}$.

The usual notation for the topological dual of a normed space X is X^*, and not X^{td}. We prefer, however, to use a different notation here, because the upper-right symbol "*" is very frequently used with different meanings and could therefore be misleading.

The following important fact about bounded linear functionals is well-known:

Theorem 7.40 (Hahn-Banach Theorem for Normed Spaces) *Let $(X, \|\cdot\|)$ be any normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For any vector subspace $Y \subseteq X$, every bounded linear functional $\varphi \in Y^{\mathrm{td}}$ is the restriction of some bounded linear function on X, having the same operator norm as φ.*

See, e.g., [18, 3.4 Theorem and its corollary]. One important consequence of this theorem is the following assertion:

Corollary 7.41 (Bounded Linear Functionals Separate Points) *Let $(X, \|\cdot\|)$ be any normed space. For any $x \in X$, one has $\varphi(x) = 0$ for every $\varphi \in X^{\mathrm{td}}$ iff $x = 0$.*

See, for instance, [18, 3.4 Theorem and its corollary].

We discuss next a few important, albeit basic, facts about tensor products of normed spaces:

Definition 7.42 (Tensor Products of Normed Spaces) Let $X_1, , \ldots, X_n$ be $n \in \mathbb{N}$ normed spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then the tensor product $X_1 \otimes \cdots \otimes X_n$ of these spaces is defined as in Definition 7.9, but $x_1 \otimes \cdots \otimes x_n$, $x_1 \in X_1, \ldots, x_n \in X_n$, are now seen (by restriction) as multilinear forms on $X_1^{\mathrm{td}} \times \cdots \times X_n^{\mathrm{td}} \subseteq X_1' \times \cdots \times X_n'$.

Observe that the tensor product $X_1 \otimes \cdots \otimes X_n$ in the sense of Definition 7.9 and the one of the last definition are equivalent vector spaces, because of the following fact:

Exercise 7.43 Show that every (multilinear form) $m \in X_1 \otimes \cdots \otimes X_n$ (see Definition 7.9) vanishes on the whole $X_1' \times \cdots \times X_n'$,whenever it vanishes on the subspace

$$X_1^{\mathrm{td}} \times \cdots \times X_n^{\mathrm{td}} \subseteq X_1' \times \cdots \times X_n' .$$

Note additionally that every element $m \in X_1 \otimes \cdots \otimes X_n$ is a "bounded multilinear" form on $X_1^{\mathrm{td}} \times \cdots \times X_n^{\mathrm{td}}$, i.e., for some $C \in \mathbb{R}_0^+$,

$$|m(\varphi_1, \ldots, \varphi_n)| \leq C \|\varphi_1\|_{\mathrm{op}} \cdots \|\varphi_n\|_{\mathrm{op}} , \qquad \varphi_1 \in X_1^{\mathrm{td}}, \ldots, \varphi_n \in X_n^{\mathrm{td}} .$$

We introduce now the notion of crossnormed tensor products:

Definition 7.44 (Crossnormed Tensor Products) Let $(X_1, \|\cdot\|^{(1)}), \ldots, (X_n, \|\cdot\|^{(n)})$ be $n \in \mathbb{N}$ normed spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We say that a norm $\|\cdot\|$ in the tensor product $X_1 \otimes \cdots \otimes X_n$ is a "crossnorm" if, for all

$$\|x_1 \otimes \cdots \otimes x_n\| = \|x_1\|^{(1)} \cdots \|x_n\|^{(n)} \ , \qquad x_1 \in X_1, \ldots, x_n \in X_n \ .$$

We call the a tensor product $X_1 \otimes \cdots \otimes X_n$ endowed with such a norm a "crossnormed tensor product." If

$$\|\varphi_1 \otimes \cdots \otimes \varphi_n\|_{\mathrm{op}} \leq \|\varphi_1\|_{\mathrm{op}} \cdots \|\varphi_n\|_{\mathrm{op}} \ , \qquad \varphi_1 \in X_1^{\mathrm{td}}, \ldots, \varphi_n \in X_n^{\mathrm{td}} \ ,$$

then we say that the tensor norm is "reasonable." Recall that $\varphi_1 \otimes \cdots \otimes \varphi_n$ is canonically seen as a linear form on $X_1 \otimes \cdots \otimes X_n$. In particular, $\varphi_1 \otimes \cdots \otimes \varphi_n \in (X_1 \otimes \cdots \otimes X_n)^{\mathrm{td}}$, in this case.

The following example of a reasonable crossnorm is very important:

Definition 7.45 (Injective Norms for Tensor Products) Let $X_1, , \ldots, X_n$ be $n \in \mathbb{N}$ normed spaces over $\mathcal{K} = \mathbb{R}, \mathbb{C}$. For all $m \in V_1 \otimes \cdots \otimes V_n$, define

$$\|m\|_\varepsilon \doteq \sup\{|m(\varphi_1, \ldots, \varphi_n)| \ : \ \varphi_k \in V_k^{\mathrm{td}} \ , \ \|\varphi_k\|_{\mathrm{op}} = 1 \ , \ k = 1, \ldots, n\} \ .$$

The mapping $\|\cdot\|_\varepsilon : X_1 \otimes \cdots \otimes X_n \to \mathbb{R}_0^+$ is a norm on $X_1 \otimes \cdots \otimes X_n$, which is called "injective norm" associated with the normed spaces $X_1, \ldots, X_n$. A completion of $X_1 \otimes \cdots \otimes X_n$ with respect to the norm $\|\cdot\|_\varepsilon$ is called "injective tensor product" of the normed spaces $X_1, \ldots, X_n$.

Exercise 7.46 Show that $\|\cdot\|_\varepsilon$ is a reasonable crossnorm on $X_1 \otimes \cdots \otimes X_n$. Prove additionally that this norm is the smallest reasonable crossnorm on $X_1 \otimes \cdots \otimes X_n$.

In fact, for any finite sequence of normed spaces, there is also a largest reasonable crossnorm on the corresponding tensor product. This norm is called "projective norm" associated with the normed spaces. For more details, see, for instance, [2, Section 6.1].

In the next exercise, we claim that the tensor product of bounded linear operators in normed spaces defines a linear bounded operator in the corresponding injective tensor product:

Exercise 7.47 Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be arbitrary normed spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, where $n \in \mathbb{N}$. For any bounded linear operators $A_1 \in \mathcal{B}(X_1, Y_1), \ldots, A_n \in \mathcal{B}(X_n, Y_n)$, the linear mapping

$$\begin{aligned} A_1 \otimes \cdots \otimes A_n : X_1 \otimes \cdots \otimes X_n &\to Y_1 \otimes \cdots \otimes Y_n \\ x_1 \otimes \cdots \otimes x_n &\mapsto A_1 x_1 \otimes \cdots \otimes A_n x_n \end{aligned}$$

(see Corollary 7.11) is bounded with respect to the injective norms of $X_1 \otimes \cdots \otimes X_n$ and $Y_1 \otimes \cdots \otimes Y_n$ with

$$\|A_1 \otimes \cdots \otimes A_n\|_{\mathrm{op}} = \|A_1\|_{\mathrm{op}} \cdots \|A_n\|_{\mathrm{op}} \ .$$

We now continue by studying (semi)normed spaces that are algebras (Definition 7.15):

Definition 7.48 (Normed Algebra) Let $(\mathcal{A}, +, \cdot, \circ)$ be an algebra and $\|\cdot\|$ a seminorm in the vector space $(\mathcal{A}, +, \cdot)$.

(i) The structure $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ is called a "seminormed algebra" if the seminorm $\|\cdot\|$ is "submultiplicative," i.e., for all $A_1, A_2 \in \mathcal{A}$,

$$\|A_1 \circ A_2\| \leq \|A_1\| \, \|A_2\| \ .$$

(ii) A seminormed algebra $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ is called a "normed algebra" if $\|\cdot\|$ is a norm.
(iii) A normed algebra $\mathcal{A}$ is "simple" if $\{0\}$ and $\mathcal{A}$ are the only *closed* ideals of $\mathcal{A}$.

Compare Definition 7.48 (iii) with Definition 7.24 (iii). Note, in particular, that even if the normed algebra $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ is simple, the corresponding algebra $(\mathcal{A}, +, \cdot, \circ)$ may be non-simple as it may have non-trivial ideals that are not closed.

Again, we often use the simpler notations $(\mathcal{A}, \|\cdot\|)$ or $\mathcal{A}$ for a generic normed algebra $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$. For any normed algebra $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$, $\mathcal{F}_{\mathrm{b}}(\Omega; \mathcal{A})$ with its supremum norm is again a normed algebra. Given any normed space X, the space $\mathcal{B}(X)$ endowed with its operator norm is another important example of a normed algebra.

Exercise 7.49 Let $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ be any normed algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed ideal. Prove that $(\mathcal{A}/\mathcal{I}, +, \cdot, \circ, \|\cdot\|_{\mathcal{A}/\mathcal{I}})$ is also a normed algebra.

Exercise 7.50 Let $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ be any seminormed algebra. Show that

$$\mathcal{A}_0 \doteq \{A \in \mathcal{A} \ : \ \|A\| = 0\}$$

is an ideal of $\mathcal{A}$ and that $(\mathcal{A}/\mathcal{A}_0, \|\cdot\|_{\mathcal{A}/\mathcal{A}_0})$ is a normed algebra.

The following structure associated with complex vector spaces is very important in our considerations. It refers to the existence of an antilinear involution in this spaces and generalizes the notion of complex conjugate of a complex number:

Definition 7.51 (Complex Conjugation) Let $(V, +, \cdot)$ be a *complex* vector space. A "complex conjugation" in V is a antilinear involution $(\cdot)^* : V \to V$, meaning that it has the following properties:

(i) *Involution property*. For all $v \in V$, $v^{**} \doteq (v^*)^* = v$, i.e., the mapping $(\cdot)^*$ is its own inverse mapping.

(ii) *Additivity*. For all $v, v' \in V$, $(v + v')^* = v^* + v'^*$.
(iii) For all $v \in V$ and all $\alpha \in \mathbb{C}$, $(\alpha v)^* = \bar{\alpha} v^*$.

This definition leads to the concept of $*$-vector spaces defined as follows:

Definition 7.52 ($*$-Vector Space)

(i) A "$*$-vector space" is a structure $(V, +, \cdot, ^*)$, where $(V, +, \cdot)$ and $(\cdot)^*$ are, respectively, a complex vector space and a complex conjugation in V.
(ii) A vector $v \in V$ is said to be "self-conjugate" in $(V, +, \cdot, ^*)$ whenever it is preserved by the complex conjugation, i.e., $v = v^*$. Similarly, a subset $\Omega \subseteq V$ is self-conjugate if

$$(\Omega)^* \doteq \left\{v^* : v \in \Omega\right\} \subseteq \Omega \, .$$

The set of all self-conjugate vectors of V is denoted by $\mathrm{Re}\{V\} \subseteq V$. In particular, $\mathrm{Re}\{V\}$ is a self-conjugate subset.

Again, we often use the simpler notation V for a generic $*$-vector space $(V, +, \cdot, ^*)$.

In some contexts, for instance, in the theory of $*$-algebras, self-conjugate elements are called "self-adjoint." Any self-conjugate vector subspace of a $*$-vector space is canonically a $*$-vector space, whose complex conjugation is the restriction to the subspace of the original conjugation. Note also that an arbitrary element of a self-conjugate subset of a $*$-vector space is not necessarily self-conjugate itself. One simple example of a $*$-vector space is the complex field $\mathbb{C}$, which is canonically a complex vector space, with the usual conjugation of complex numbers. By default, $\mathbb{C}$ is seen as a $*$-vector space with this (canonical) complex conjugation.

Definition 7.53 (Real and Imaginary Parts) Let $(V, +, \cdot, ^*)$ be a $*$-vector space. We define the mappings $\mathrm{Re}\{\cdot\}, \mathrm{Im}\{\cdot\} : V \to \mathrm{Re}\{V\}$ as follows:

$$\mathrm{Re}\{v\} \doteq \frac{1}{2}(v + v^*) \, , \qquad \mathrm{Im}\{v\} \doteq \frac{1}{2i}(v - v^*) \, , \qquad v \in V \, .$$

$\mathrm{Re}\{v\}$ and $\mathrm{Im}\{v\}$ are naturally called "real part" and "imaginary part" of $v \in V$, respectively.

Given a $*$-vector space $(V, +, \cdot, ^*)$, note that $\mathrm{Re}\{V\}$ is a real vector subspace[9] of V and the mappings $\mathrm{Re}\{\cdot\}, \mathrm{Im}\{\cdot\} : V \to \mathrm{Re}\{V\}$ are real linear. Moreover, the real and imaginary parts have the following important, albeit simple, properties (that are well-known for complex numbers):

(a) For all $v \in V$, the following three conditions are equivalent:

$$\text{(i) } v = v^* \, , \qquad \text{(ii) } \mathrm{Re}\{v\} = v \, , \qquad \text{(ii) } \mathrm{Im}\{v\} = 0 \, .$$

[9] Recall that any complex vector space is canonically a real vector space.

(b) For all $v \in V$,

$$v = \mathrm{Re}\{v\} + i\mathrm{Im}\{v\} .$$

In particular, $\mathrm{Re}\{V\}$ is generating for V. See Definition 7.3 (ii).

We now introduce the concept of $*$-morphisms, which are special cases of linear transformations between $*$-vector spaces:

Definition 7.54 ($*$-Morphism) Let $(V_1, +, \cdot, ^*)$ and $(V_2, +, \cdot, ^*)$ be two $*$-vector spaces. We say that the linear transformation $\Theta : V_1 \to V_2$ is a "$*$-morphism" if, for all $v_1 \in V_1$, $\Theta(v_1)^* = \Theta(v_1^*)$.

Note, in particular, that $\mathrm{Re}\{\Theta(\cdot)\} = \Theta(\mathrm{Re}\{\cdot\})$ and $\mathrm{Im}\{\Theta(\cdot)\} = \Theta(\mathrm{Im}\{\cdot\})$, whenever $\Theta : V_1 \to V_2$ is a $*$-morphism.

If V_1 and V_2 are two $*$-vector spaces, then the space $\mathcal{L}(V_1; V_2)$ of linear transformations is canonically a $*$-vector space:

Definition 7.55 (Complex Conjugation for lin. trafos. Between $*$-Vector Spaces) Let $(V_1, +, \cdot, ^*)$ and $(V_2, +, \cdot, ^*)$ be two $*$-vector spaces. We introduce a complex conjugation in $\mathcal{L}(V_1; V_2)$ as follows: For every $\Theta \in \mathcal{L}(V_1; V_2)$, we define $\Theta^* \in \mathcal{L}(V_1; V_2)$ by

$$\Theta^* (v_1) \doteq \Theta \left(v_1^*\right)^* , \qquad v_1 \in V_1 .$$

$\mathcal{L}(V_1; V_2)$ is canonically endowed with this complex conjugation and seen as a $*$-vector space.

For an arbitrary $*$-vector space V, note that self-conjugate elements of V, i.e., linear functionals on V, are frequently called "Hermitian" (linear) functionals, for instance, in the theory of $*$-algebras. Observe also that, for any two $*$-vector spaces V_1 and V_2, $\mathrm{Re}\{\mathcal{L}(V_1; V_2)\}$ is nothing else than the set of all $*$-morphisms from V_1 to V_2.

Exercise 7.56 Let V_1 and V_2 be two $*$-vector spaces.

(i) Show that $\Theta \in \mathrm{Re}\{\mathcal{L}(V_1; V_2)\}$ iff, for all $v_1 \in \mathrm{Re}\{V_1\}$, one has that $\Theta(v_1) \in \mathrm{Re}\{V_2\}$.
(ii) Show that, for any $\Theta \in \mathrm{Re}\{\mathcal{L}(V_1; V_2)\}$, $\ker(\Theta)$ is a self-conjugate vector subspace of V_1.
(iii) Show that any $\Theta \in \mathrm{Re}\{\mathcal{L}(V_1; V_2)\}$ is uniquely determined by its restriction to $\mathrm{Re}\{V_1\}$.
(iv) Show that the restriction to $\mathrm{Re}\{V_1\}$ of mappings $V_1 \to V_2$ defines an one-to-one correspondence from $\mathrm{Re}\{\mathcal{L}(V_1; V_2)\}$ to $\mathcal{L}(\mathrm{Re}\{V_1\}; \mathrm{Re}\{V_2\})$.
(v) Show that if V_1 and V_2 are $*$-normed spaces, then $\mathcal{B}(V_1; V_2) \subseteq \mathcal{L}(V_1; V_2)$ is a self-conjugate vector subspace of $\mathcal{L}(V_1; V_2)$.

Tensor products of $*$-vector spaces are naturally $*$-vector spaces:

Proposition 7.57 *Let $V_1, \ldots, V_n$ be $n \in \mathbb{N}$ $*$-vector spaces. Then, there is a unique complex conjugation in the tensor product $V_1 \otimes \cdots \otimes V_n$, such that*

$$(v_1 \otimes \cdots \otimes v_n)^* = v_1^* \otimes \cdots \otimes v_n^* \,, \qquad v_1 \in V_1, \ldots, v_n \in V_n \,.$$

Proof This follows directly from Proposition 7.12. □

Similar to the quotient of vector spaces by subspaces of Definition 7.7, we define quotients of $*$-vector spaces by self-conjugate subspaces as follows:

Definition 7.58 (Quotient of a $*$-Vector Space by a Subspace) Let V be a $*$-vector space and $V' \subseteq V$ a self-conjugate subspace. Observe that the V'-equivalence is a congruence relation, i.e., it is an equivalence relation in V that is compatible with the vector space operations (Definition 7.7), as well as with its involution in the sense that, for all $v_1, v_2 \in V$, v_1^* is V'-equivalent to v_2^* whenever v_1 and v_2 are V'-equivalent. Thus, besides the vector space operations for V/V' as defined in Definition 7.7, there is additionally a unique involution $(\cdot)^* : V/V' \to V/V'$ such that $[v]^* = [v^*]$ for any $v \in V$ Then, $(V/V', +, \cdot, ^*)$ is a $*$-vector space, named the "quotient of V by the self-conjugate subspace V'."

Exercise 7.59

(i) Let V be a $*$-vector space and $V' \subseteq V$ a self-conjugate vector subspace. Prove that the mapping $\mathsf{q} : v \mapsto [v]$ from V to the quotient $*$-vector space V/V' is a $*$-morphism.

(ii) Let V_1 and V_2 be arbitrary $*$-vector spaces and $\Theta : V_1 \to V_2$ a linear transformation. Let $V_1' \subseteq \ker(\Theta) \subseteq V_1$ be any subspace of the kernel of Θ. Show that there is a unique $*$-morphism $\bar{\Theta} : V_1/V_1' \to V_2$ such that $\Theta = \bar{\Theta} \circ \mathsf{q}$, i.e., $\Theta(v_1) = \bar{\Theta}([v_1])$ for any $v_1 \in V_1$.

The following definitions are natural for complex algebras and normed spaces having a complex conjugation:

Definition 7.60 ($*$-Algebra) Let $(\mathcal{A}, +, \cdot, \circ)$ be an *complex* algebra and $(\cdot)^* : \mathcal{A} \to \mathcal{A}$ a complex conjugation in the (complex) vector space $(\mathcal{A}, +, \cdot)$.

(i) The structure $(\mathcal{A}, +, \cdot, \circ, ^*)$ is a "$*$-algebra," whenever, for all $A, A' \in \mathcal{A}$, one has that $(AA')^* = A'^* A^*$.

(ii) Elements $A \in \mathcal{A}$ satisfying $A^*A = AA^*$ are said to be "normal."

(iii) A subalgebra of $\mathcal{A}$ is said to be a $*$-subalgebra if it is a self-conjugate vector subspace of $\mathcal{A}$.

(iv) A "$*$-homomorphism" ("$*$-isomorphism") is, by definition, a self-conjugate algebra homomorphism between $*$-algebras, that is, it is simultaneously an algebra homomorphism (isomorphism) and a $*$-morphism of $*$-vector spaces. See Definitions 7.20 and 7.54.

As before for other algebraic structures, we frequently use the simpler notation $\mathcal{A}$ for generic $*$-algebras $(\mathcal{A}, +, \cdot, \circ, ^*)$.

Note that any $*$-subalgebra of a $*$-algebra is (canonically) a $*$-algebra, the complex conjugation of which is the canonical one for self-conjugate subspaces. Observe also that, if $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ is a $*$-isomorphism between two $*$-algebras $\mathcal{A}_1$ and $\mathcal{A}_2$, then $\Theta^{-1} : \mathcal{A}_2 \to \mathcal{A}_1$ is also a $*$-isomorphism.

Example 7.61 Let $\Omega \neq \emptyset$ be a nonempty set. We naturally define a complex conjugation $(\cdot)^* : \mathcal{F}(\Omega; \mathbb{C}) \to \mathcal{F}(\Omega; \mathbb{C})$ by

$$f^*(p) \doteq \overline{f(p)}\ , \qquad p \in \Omega\ .$$

It is easy to check that $(\mathcal{F}(\Omega; \mathbb{C}), +, \cdot, \circ, ^*)$ is a $*$-algebra. For this complex conjugation, observe that

$$\mathrm{Re}\{f\}(p) = \mathrm{Re}\{f(p)\}\ , \qquad f \in \mathcal{F}(\Omega; \mathbb{C}),\ p \in \Omega\ .$$

More generally, for any $*$-vector space $(V, +, \cdot, ^*)$, one can define a complex conjugation $(\cdot)^* : \mathcal{F}(\Omega; V) \to \mathcal{F}(\Omega; V)$ by

$$f^*(p) \doteq f(p)^*\ , \qquad p \in \Omega\ .$$

$(\mathcal{F}(\Omega; V), +, \cdot, ^*)$ is again a $*$-vector space. If $(\mathcal{A}, +, \cdot, \circ, ^*)$ is a $*$-algebra, then so does $(\mathcal{F}(\Omega; \mathcal{A}), +, \cdot, \circ, ^*)$.

For any vector space V, recall that the set $\mathcal{L}(V)$ of linear transformations $V \to V$ (Definition 7.4) endowed with the composition of mappings is (canonically) an algebra. If V is a $*$-vector space, then $\mathcal{L}(V)$ is additionally a $*$-vector space. However, in this case, $\mathcal{L}(V)$ is generally *not* a $*$-algebra, because, for all $A, A' \in \mathcal{L}(V)$, one has that $(AA')^* = A^*A'^*$, instead of $(AA')^* = A'^*A^*$.

We study now tensor products of $*$-algebras:

Proposition 7.62 (Tensor Products of $*$-Algebras) *Let $\mathcal{A}_1, \ldots, \mathcal{A}_n$ be $n \in \mathbb{N}$ $*$-algebras. Then, there is a unique $*$-algebra structure $(\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n, +, \cdot, \circ, ^*)$ such that*

$$(A_1 \otimes \cdots \otimes A_n) \circ (A'_1 \otimes \cdots \otimes A'_n) = (A_1 \circ A'_1) \otimes \cdots \otimes (A_n \circ A'_n)$$

as well as

$$(A_1 \otimes \cdots \otimes A_n)^* = A_1^* \otimes \cdots \otimes A_n^*$$

for all $A_1, A'_1 \in \mathcal{A}_1, \ldots, A_n, A'_n \in \mathcal{A}_n$.

Proof The proposition directly follows from Propositions 7.19 and 7.57. □

As in the case of usual algebras (Definition 7.16), any non-unital $*$-algebra can be canonically extended to a unital one:

Definition 7.63 (Unitization of an $*$-Algebra) Let $(\mathcal{A}, +, \cdot, \circ, ^*)$ be a $*$-algebra and define the unital algebra $\tilde{\mathcal{A}} \doteq \mathbb{C} \times \mathcal{A}$, which has $\mathcal{A}$ as a subalgebra, as in Definition 7.16. Then we define an involution $(\cdot)^* : \tilde{\mathcal{A}} \to \tilde{\mathcal{A}}$ (complex conjugation) as follows:

$$(\alpha, A)^* \doteq (\bar{\alpha}, A^*) , \qquad (\alpha, A) \in \tilde{\mathcal{A}} .$$

It is easy to check that $(\tilde{\mathcal{A}}, +, \cdot, \circ, ^*)$ is a unital $*$-algebra.

Exercise 7.64 (Unitization of a $*$-Homomorphism) Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two $*$-algebras, $\mathcal{A}_2$ being unital and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ a $*$-homomorphism. Show that there is a unique unital $*$-homomorphism $\tilde{\Theta} : \tilde{\mathcal{A}}_1 \to \mathcal{A}_2$ extending Θ to the unitization $\tilde{\mathcal{A}}_1$ of $\mathcal{A}_1$.

The construction done to define the unital free algebras of Definition 7.17 has a natural version for $*$-algebras: For an arbitrary nonempty set Ω, $\tilde{\mathcal{S}}^*(\Omega)$ denotes the set of all finite sequences of elements of $\Omega^* \doteq \Omega \times \{-, +\}$, i.e., $\tilde{\mathcal{S}}^*(\Omega) \doteq \tilde{\mathcal{S}}(\Omega^*)$. For simplicity of notation, $(\omega, -), (\omega, +) \in \Omega^*$ are, respectively, denoted by ω and ω^*, and we identify Ω with the subset

$$\left\{(\omega, -) \in \Omega^* \ : \ \omega \in \Omega\right\} \subseteq \Omega^* .$$

The elements of $\tilde{\mathcal{S}}^*(\Omega)$ are called "$*$-words on the alphabet Ω." The mapping $(\cdot)^* : \tilde{\mathcal{S}}^*(\Omega) \to \tilde{\mathcal{S}}^*(\Omega)$ denotes the unique involution of the monoid $(\tilde{\mathcal{S}}^*(\Omega), \star)$ for which $\omega \mapsto \omega^*$ for any $\omega \in \Omega$, $\imath^* = \imath$ and $(s_1 \star s_2)^* = s_2^* \star s_1^*$ for any $*$-words $s_1, s_2 \in \tilde{\mathcal{S}}^*(\Omega)$. The set of nonempty $*$-words is denoted by $\mathcal{S}^*(\Omega) \doteq \tilde{\mathcal{S}}^*(\Omega)\backslash\{\imath\}$. Note that the involution preserves $\mathcal{S}^*(\Omega)$, i.e.,

$$\mathcal{S}^*(\Omega)^* \subseteq \mathcal{S}^*(\Omega) ,$$

and thus defines an involution (also denoted by $(\cdot)^*$) on $\mathcal{S}^*(\Omega)$, by restriction. Then, similar to Definition 7.17, we define unital free $*$-algebras as follows:

Definition 7.65 (Unital Free $*$-Algebra) Let $\Omega \neq \emptyset$ be any nonempty set and $\tilde{F}^*(\Omega)$ the vector space $\mathcal{F}_0(\tilde{\mathcal{S}}^*(\Omega); \mathbb{C})$ of $\mathbb{C}$-valued functions on $*$-words. Recall that there is a unique operation

$$\star : \tilde{F}^*(\Omega) \times \tilde{F}^*(\Omega) \to \tilde{F}^*(\Omega)$$

extending the product of the monoid $(\tilde{\mathcal{S}}^*(\Omega), \star)$ in such a way that $(\mathfrak{F}^*(\Omega), +, \cdot, \star)$ is an algebra over $\mathbb{C}$. Similarly, there is a unique antiunitary involution

$$(\cdot)^* : \tilde{F}^*(\Omega) \to \tilde{F}^*(\Omega)$$

that extends the involution of $\tilde{\mathcal{S}}^*(\Omega) \subseteq \tilde{F}^*(\Omega)$ in such a way that $(\tilde{F}^*(\Omega), +, \cdot, \star, ^*)$ is a $*$-algebra, which is named the "unital free $*$-algebra generated by Ω."

The non-unital case is defined in a similar way by replacing the set $\tilde{\mathcal{S}}^*(\Omega)$ of all $*$-words with the set $\mathcal{S}^*(\Omega)$ of nonempty $*$-words:

Definition 7.66 (Non-Unital Free $*$-Algebra) Let $\Omega \neq \emptyset$ be any nonempty set. Let $\mathfrak{F}^*(\Omega)$ be the vector space $\mathcal{F}_0(\mathcal{S}^*(\Omega); \mathbb{C})$ of $\mathbb{C}$-valued functions on nonempty words. As before, there is a unique operation

$$\star : \mathfrak{F}^*(\Omega) \times \mathfrak{F}^*(\Omega) \to \mathfrak{F}^*(\Omega)$$

extending the product of $(\mathcal{S}^*(\Omega), \star)$ in such a way that $(\mathfrak{F}^*(\Omega), +, \cdot, \star)$ is an algebra, as well as a unique antiunitary involution

$$(\cdot)^* : \mathfrak{F}^*(\Omega) \to \mathfrak{F}^*(\Omega)$$

that extends the involution of $\mathcal{S}^*(\Omega) \subseteq \mathfrak{F}^*(\Omega)$ in such a way that $(\mathfrak{F}^*(\Omega), +, \cdot, \star, ^*)$ is a $*$-algebra, which is named the "free $*$-algebra generated by Ω."

In fact, the unital free $*$-algebra $\tilde{F}^*(\Omega)$ can be canonically identified with the unitization of the free $*$-algebra $\mathfrak{F}^*(\Omega)$. Additionally, similar to Exercise 7.22, we have the following properties:

Exercise 7.67 Let $\Omega \neq \emptyset$ be any nonempty set and $\mathcal{A}$ a $*$-algebra. If $\mathcal{A}$ has no unit, prove that, for any mapping $i : \Omega \to \mathcal{A}$, there is a unique $*$-homomorphism $\mathfrak{F}^*(\Omega) \to \mathcal{A}$ extending i. Otherwise, prove the same property with $\tilde{F}^*(\Omega)$ instead of $\mathfrak{F}^*(\Omega)$, the $*$-homomorphism $\tilde{F}^*(\Omega) \to \mathcal{A}$ being now unital. Recall that one has the canonical inclusions $\Omega \subseteq \mathcal{S}^*(\Omega) \subseteq \mathfrak{F}^*(\Omega)$ and $\Omega \subseteq \tilde{\mathcal{S}}^*(\Omega) \subseteq \tilde{F}^*(\Omega)$. This $*$-homomorphism is in any case denoted again by i.

The above exercise shows that the properties of $\mathfrak{F}^*(\Omega)$ or $\tilde{F}^*(\Omega)$ uniquely determine the (possibly unital) free $*$-algebra, up to a $*$-isomorphism.

Exercise 7.68 Let $(\mathcal{A}, +, \cdot, \circ, ^*)$ be an arbitrary $*$-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a left or right ideal of the algebra $(\mathcal{A}, +, \cdot, \circ)$ that is a self-conjugate subspace of the $*$-vector space $(\mathcal{A}, +, \cdot, ^*)$. Show that $\mathcal{I}$ is then an ideal.

The above exercise motivates the following definition:

Definition 7.69 ($*$-Ideal of a $*$-Algebra) Let $(\mathcal{A}, +, \cdot, \circ, ^*)$ be a $*$-algebra. We say that the vector subspace $\mathcal{I} \subseteq \mathcal{A}$ is a "$*$-ideal" if it is a self-conjugate left or right ideal of the algebra $(\mathcal{A}, +, \cdot, \circ)$. In particular, $\mathcal{I}$ is in this case an ideal of the algebra (Definition 7.24).

Exercise 7.70 Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be two $*$-algebras and $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ a $*$-homomorphism. Show that $\ker(\Theta)$ is a $*$-ideal of $\mathcal{A}_1$. See Definition 7.6.

Another example of non-trivial $*$-ideals is given by unitizations of non-unital $*$-algebras: Note that if $\tilde{\mathcal{A}}$ is the unitization of a non-unital $*$-algebra $\mathcal{A}$, as in Definition 7.63, then $\mathcal{A}$ is not only a $*$-subalgebra of $\tilde{\mathcal{A}}$ but also a $*$-ideal.

Exercise 7.71 Let $\mathcal{A}$ be a $*$-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a $*$-ideal. Recall that the $\mathcal{I}$-equivalence of elements of $\mathcal{A}$ is a congruence relation for the algebra structure $(\mathcal{A}, +, \cdot, \circ)$, $\mathcal{I}$ being an ideal, by definition of a $*$-ideal. Show that it is also a congruence relation for the $*$-algebra structure $(\mathcal{A}, +, \cdot, \circ, ^*)$, i.e., that for all $A_1, A_2 \in \mathcal{A}$, A_1^* is $\mathcal{I}$-equivalent to A_2^*, whenever A_1 and A_2 are $\mathcal{I}$-equivalent.

This exercise allows us to define the quotient of a $*$-algebra by a $*$-ideal: Given a $*$-algebra $\mathcal{A}$ and a $*$-ideal $\mathcal{I} \subseteq \mathcal{A}$, recall that $\mathcal{A}/\mathcal{I}$ is an algebra over $\mathbb{C}$. See Definitions 7.28 and 7.7. Then, because the $\mathcal{I}$-equivalence is a congruence relation for the $*$-algebra structure $(\mathcal{A}, +, \cdot, \circ, ^*)$, we define quotients as $*$-algebras as follows:

Definition 7.72 (Quotient of a $*$-Algebra by a $*$-Ideal) Let $\mathcal{A}$ be a $*$-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a $*$-ideal. By Exercise 7.71, there is a unique operation, $(\cdot)^* : \mathcal{A}/\mathcal{I} \to \mathcal{A}/\mathcal{I}$ (conjugation) for which $[A]^* = [A^*]$ for any $A \in \mathcal{A}$. Then, $(\mathcal{A}/\mathcal{I}, +, \cdot, \circ, ^*)$ is a $*$-algebra, named the "quotient of the $*$-algebra $\mathcal{A}$ by its self-conjugate ideal $\mathcal{I}$."

Exercise 7.73

(i) Let $\mathcal{A}$ be a $*$-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a $*$-ideal. Prove that the mapping $\mathfrak{q} : A \mapsto [A]$ from the $*$-algebra $\mathcal{A}$ to the quotient $*$-algebra $\mathcal{A}/\mathcal{I}$ is a $*$-homomorphism.

(ii) Show that if $\mathcal{A}_1$ and $\mathcal{A}_2$ are two $*$-algebras, $\Theta : \mathcal{A}_1 \to \mathcal{A}_2$ a $*$-homomorphism and $\mathcal{I}_1 \subseteq \ker(\Theta) \subseteq \mathcal{A}_1$ a $*$-ideal, then the mapping $\bar{\Theta} : \mathcal{A}_1/\mathcal{I}_1 \to \mathcal{A}_2$, as defined in Exercise 7.8, is a $*$-homomorphism, which is faithful (i.e., injective) for $\mathcal{I}_1 = \ker(\Theta)$. Recall that $\ker(\Theta) \subseteq \mathcal{A}_1$ is a $*$-ideal, Θ being a $*$-homomorphism.

Exercise 7.74 Let $\mathcal{A}$ be an arbitrary $*$-algebra. Prove that there exists at least one non-zero, self-conjugate, multiplicative, linear functional $\rho : \tilde{\mathcal{A}} \to \mathbb{C}$, where $\tilde{\mathcal{A}}$ is the unitization of $\mathcal{A}$.

We now discuss properties of $*$-vector spaces and $*$-algebras that are equipped with (semi)norms, which are compatible with the algebraic structure, in a natural manner:

Definition 7.75 ($*$-Normed Spaces and Algebras) Let $(X, +, \cdot, \|\cdot\|)$ be a *complex* seminormed space (Definition 7.31) and $(\cdot)^* : X \to X$ a complex conjugation in the (complex) vector space $(X, +, \cdot)$ (Definition 7.51).

(i) The structure $(X, +, \cdot, ^*, \|\cdot\|)$ is a "$*$-seminormed space," whenever

$$\|x^*\| = \|x\| \ , \qquad x \in X \ .$$

If $\|\cdot\|$ is a norm, then the $*$-seminormed space is said to be a "$*$-normed space."

(ii) If $(\mathcal{A}, +, \cdot, \circ, ^*)$ is a $*$-algebra and $\|\cdot\|$ a (semi)norm in $(\mathcal{A}, +, \cdot)$ such that $(\mathcal{A}, +, \cdot, \|\cdot\|)$ is a (semi)normed algebra and $(\mathcal{A}, +, \cdot, ^*, \|\cdot\|)$ a $*$-(semi)normed space, then we call the structure $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ a "(semi)normed $*$-algebra."

Note, in particular, that a normed $*$-algebra is not merely an algebra which is simultaneously a normed algebra and a $*$-algebra, but it is an algebra with a complex conjugation that preserves the norm. In other words, as is usual, some mutual compatibility is required for the given operations.

Example 7.76 Let Ω be a nonempty set and $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ a normed $*$-algebra. Define the complex conjugation $(\cdot)^* : \mathcal{F}_b(\Omega; \mathcal{A}) \rightarrow \mathcal{F}_b(\Omega; \mathcal{A})$ exactly as in Example 7.61. See also Eq. (7.1). Then,

$$(\mathcal{F}_b(\Omega; \mathcal{A}), +, \cdot, \circ, ^*, \|\cdot\|_\infty)$$

is again a normed $*$-algebra.

Exercise 7.77 Let X_1 and X_2 be two $*$-normed spaces.

(i) Prove that $\mathcal{B}(X_1; X_2)$ is a self-conjugate subspace of $\mathcal{L}(X_1; X_2)$. In particular, for any $*$-normed space X, the topological dual X^{td} is a self-conjugate subspace of the dual space X'.
(ii) Show that the restriction to X_1 of mappings $X_1 \rightarrow X_2$ defines a one-to-one correspondence from $\mathrm{Re}\{\mathcal{B}(X_1; X_2)\}$ to $\mathcal{B}(\mathrm{Re}\{X_1\}; \mathrm{Re}\{X_2\})$.
(iii) Prove that, for any $\varphi \in \mathrm{Re}\{X_1'\}$, i.e., any self-conjugate linear functional on X_1, one has

$$\|\varphi\|_{\mathrm{op}} = \sup_{x \in \mathrm{Re}\{X_1\}} |\varphi(x)| \in [0, \infty] .$$

Exercise 7.78

(i) Let $(X, \|\cdot\|)$ be a seminormed $*$-vector space and $Y \subseteq X$ a self-conjugate subspace. Show that $(X/Y, \|\cdot\|_{X/Y})$ is also a seminormed $*$-vector space.
(ii) Show additionally that if $(X, \|\cdot\|)$ a normed $*$-vector space and $Y \subseteq X$ a closed self-conjugate subspace, then $(X/Y, \|\cdot\|_{X/Y})$ is also a normed $*$-vector space.
(iii) For an arbitrary seminormed $*$-algebra $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ and $*$-ideal $\mathcal{I} \subseteq \mathcal{A}$, prove that $(\mathcal{A}/\mathcal{I}, +, \cdot, \circ, ^*, \|\cdot\|_{\mathcal{A}/\mathcal{I}})$ is also a seminormed $*$-algebra.

Exercise 7.79 Let $(X, +, \cdot, ^*, \|\cdot\|)$ be a seminormed space. Show that

$$X_0 \doteq \{x \in X \ : \ \|x\| = 0\}$$

is a self-conjugate subspace of X and that $(X/X_0, \|\cdot\|_{X/X_0})$ is a normed $*$-vector. Similarly, let $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ be any seminormed $*$- algebra and show that

$$\mathcal{A}_0 \doteq \{A \in \mathcal{A} \ : \ \|A\| = 0\}$$

is a $*$-ideal of $\mathcal{A}$ and that $(\mathcal{A}/\mathcal{A}_0, \|\cdot\|_{\mathcal{A}/\mathcal{A}_0})$ is a normed $*$-algebra.

In the last part of this section on vector spaces and algebras, we briefly discuss some basic aspects of convergence in normed spaces, which will be considered again in the next section, in the more general context of metric spaces.

Definition 7.80 (Cauchy Sequences and Convergence) Let $(X, \|\cdot\|)$ be a normed space and $x_n \in X$, $n \in \mathbb{N}$, a sequence in X.

(i) $(x_n)_{n\in\mathbb{N}}$ is a "Cauchy sequence" if, for any $\varepsilon > 0$, there is $N_\varepsilon \in \mathbb{N}$ such that, for all $m, n \in \mathbb{N}$ satisfying $m, n \geq N_\varepsilon$, one has that $\|x_n - x_m\| \leq \varepsilon$.
(ii) The sequence $(x_n)_{n\in\mathbb{N}}$ "converges in X" or is "convergent" if, for some $x \in X$ and all $\varepsilon > 0$, there is $N_\varepsilon \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$, $n \geq N_\varepsilon$, one has that $\|x - x_n\| \leq \varepsilon$. In this case, the sequence "converges to x."

By the subadditivity of norms (Definition 7.31), any convergent sequence is a Cauchy sequence. If a Cauchy sequence $(x_n)_{n\in\mathbb{N}}$ converges in the normed space X to $x \in X$, then such an element, called the "limit" of the sequence, is unique. It is denoted by "$\lim_{n\to\infty} x_n$."

Exercise 7.81 Let X be a normed space and $Y \subseteq X$ a vector subspace. Show that Y is closed in the sense of Definition 7.31 (v) iff all convergent sequences $(y_n)_{n\in\mathbb{N}}$ in Y have its limit in Y.

A special and very important class of normed spaces and algebras is the following:

Definition 7.82 (Banach Spaces and Algebras) A "Banach space" is a normed space $(X, +, \cdot, \|\cdot\|)$ that is "complete," which means that all Cauchy sequences in X are convergent (in X). In this case, the norm of X is also said to be "complete." The normed algebra $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ is a "Banach algebra" if it is associative and $(\mathcal{A}, +, \cdot, \|\cdot\|)$ is a Banach space.

Exercise 7.83 Let $(X, \|\cdot\|)$ be a Banach space and $Y \subseteq X$ a closed subspace. Show that the norm $\|\cdot\|_{X/Y}$ in the quotient vector space X/Y is complete. In particular, if $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ is a Banach algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed ideal, then $(\mathcal{A}/\mathcal{I}, +, \cdot, \circ, \|\cdot\|_{\mathcal{A}/\mathcal{I}})$ is also a Banach algebra.

The following result on bounded linear operators on Banach spaces, known as the Banach-Steinhaus "uniform boundedness principle," has many important consequences and is given here (without proof) for completeness:

Theorem 7.84 (Banach-Steinhaus) *Let X_1 be a* Banach *space and X_2 any (not necessarily complete) normed space, both over the same field $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Let $\Omega \subseteq \mathcal{B}(X_1; X_2)$ be any subset of bounded linear operators $X_1 \to X_2$. If, for all $x_1 \in X_1$, the subset*

$$\{\Theta(x_1) \ : \ \Theta \in \Omega\} \subseteq X_2$$

is norm-bounded, then Ω is uniformly bounded, i.e., $\{\|\Theta\|_{\mathrm{op}} : \Theta \in \Omega\} \subseteq \mathbb{R}_0^+$ is bounded.

For a proof of this result, see, for instance, [18, Part I, Section 2].

Similarly, in the context of $*$-normed spaces and algebras, we define $*$-Banach spaces and algebras:

Definition 7.85 ($*$-Banach Spaces and C^*-Algebras)

(i) A "$*$-Banach space" is a $*$-normed space a $(X, +, \cdot, ^*, \|\cdot\|)$ such that $(X, +, \cdot, \|\cdot\|)$ is a Banach space. In the same way, a "Banach $*$-algebra" is a normed $*$-algebra $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ such that $(\mathcal{A}, +, \cdot, \circ, \|\cdot\|)$ is a "Banach algebra."

(ii) A Banach $*$-algebra $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ is a "C^*-algebra" if, for all $A \in \mathcal{A}$, one has that $\|A^*A\| = \|A\|^2$.

The most simple example of a C^*-algebra is the complex field $\mathbb{C}$, the absolute value being its norm. C^*-algebras are studied in more details in Chap. 4 and we only give below some important examples of $*$-Banach spaces and algebras:

Exercise 7.86

(i) Let X be a Banach space (over $\mathbb{K} = \mathbb{R}, \mathbb{C}$) and Ω an arbitrary nonempty set. Show that $\mathcal{F}_{\mathrm{b}}(\Omega; X)$ is a Banach space with respect to the supremum norm $\|\cdot\|_\infty$.

(ii) Let $(\mathcal{A}, +, \cdot, \|\cdot\|)$ be a Banach algebra. Show that also $\mathcal{F}_{\mathrm{b}}(\Omega; \mathcal{A})$ is a Banach algebra, with respect to the supremum norm $\|\cdot\|_\infty$. In particular, $\mathcal{F}_{\mathrm{b}}(\Omega; \mathcal{A})$ is a Banach $*$-algebra whenever $\mathcal{A}$ is such a $*$-algebra.

(iii) Let $(\mathcal{A}, +, \cdot, \|\cdot\|)$ be a C^*-algebra. Show that also $\mathcal{F}_{\mathrm{b}}(\Omega; \mathcal{A})$ is a C^*-algebra, with respect to the supremum norm $\|\cdot\|_\infty$.

(iv) Let X_1 be a normed space and X_2 a Banach space (over the same field $\mathbb{K} = \mathbb{R}, \mathbb{C}$). Show that $\mathcal{B}(X_1; X_2)$ is a Banach space with respect to the operator norm $\|\cdot\|_{\mathrm{op}}$.

(v) Let X be a Banach space. Show that $\mathcal{B}(X)$ is a Banach algebra with respect to the operator norm $\|\cdot\|_{\mathrm{op}}$.

Exercise 7.87 Let $(\mathcal{A}, +, \cdot, \circ, ^*, \|\cdot\|)$ be a Banach $*$-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed $*$-ideal. Show that $(\mathcal{A}/\mathcal{I}, +, \cdot, \circ, ^*, \|\cdot\|_{\mathcal{A}/\mathcal{I}})$ is also a Banach $*$-algebra.

By the last exercise, if $\mathcal{A}$ is a C^*-algebra and $\mathcal{I} \subseteq \mathcal{A}$ a closed $*$-ideal, then the quotient algebra $\mathcal{A}/\mathcal{I}$ is a Banach $*$-algebra. In fact, it turns out that $\mathcal{A}/\mathcal{I}$ is even a C^*-algebra. This is proven in Proposition 4.55. Moreover, in a C^*-algebra, any ideal is automatically a $*$-ideal, that is, quotients of a C^*-algebra by its closed ideals are always C^*-algebras. See Lemma 4.54.

Definition 7.88 (Completion of Normed Spaces and Algebras) Let X be a normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We say that the pair $(\overline{X}, \iota)$, where $\overline{X}$ is a Banach space over $\mathbb{K}$ and $\iota : X \to \overline{X}$ is a linear norm-preserving mapping, is the "completion" of X if it has the following universal property: For every Banach space Y and any $\Theta \in \mathcal{B}(X; Y)$, there is a unique $\bar{\Theta} \in \mathcal{B}(\overline{X}; Y)$ such that $\Theta = \bar{\Theta} \circ \iota$.

Mutatis mutandis for the completion of normed algebras, $*$-normed spaces, and normed $*$-algebras with the following additional requirements:

(i) If X is a $*$-vector space, then $\overline{X}$ and ι are, respectively, a $*$-vector space and a $*$-morphism of $*$-vector spaces.
(ii) If X is a normed $(*$-)algebra, then $\overline{X}$ and ι are, respectively, a $(*$-)Banach algebra and a homomorphism of $(*$-)algebras.

Proposition 7.89 *Every normed space X has a completion. Additionally, if $(\overline{X}_1, \iota_1)$, $(\overline{X}_2, \iota_2)$ are two completions of X, then $\overline{X}_1$ and $\overline{X}_2$ are isomorphic normed spaces. For any completion $(\overline{X}, \iota)$ of X, $\iota(X) \subseteq \overline{X}$ is a dense vector subspace. This result holds mutatis mutandis for the completions of normed algebras, $*$-normed spaces, and normed $*$-algebras.*

Proof Exercise. *Hint*: Use Theorem 7.136 and Exercise 7.137 below. □

Exercise 7.90 Let $\mathcal{A}$ be any normed $(*$-)algebra. For any completion $(\overline{\mathcal{A}}, \iota)$ of $\mathcal{A}$ as a $(*$-)normed space only, there is a unique product in $\overline{\mathcal{A}}$ such that $(\overline{\mathcal{A}}, \iota)$ is a completion of normed $(*$-)algebra, i.e., ι is a $(*$-)homomorphism of $(*$-)algebras. Show additionally that if $(\mathcal{A}, \|\cdot\|)$ is a normed $*$-algebra whose norm satisfies $\|A^*A\| = \|A\|^2$ for any $A \in A$, then its completion $\overline{\mathcal{A}}$ is a C^*-algebra.

In fact, here we define different important C^*-algebras as completions of normed $*$-algebras having norms with the C^*-property, like in the last exercise. In turn, the corresponding normed $*$-algebras are defined by choosing a convenient seminorm for a unital free $*$-algebra and taking the quotient with respect to the zero subspace of the seminorm, as explained in Exercise 7.79. See Theorem 4.134 for the whole construction.

7.2 Metric Spaces

7.2.1 Basic Notions

Definition 7.91 (Metric Space) Let M be any nonempty set. We say that the function $d : M \times M \to [0, \infty)$ is a "metric" in M if it has the following properties:

(i) *Symmetry*. For all $p, p' \in M$, $d(p, p') = d(p', p)$.
(ii) *Triangle inequality*. For all $p, p', p'' \in M$,

$$d(p, p'') \leq d(p, p') + d(p', p'') .$$

(iii) *Nondegeneracy*. For all $p, p' \in M$, $d(p, p') = 0$ iff $p = p'$.

In this case, we say that the pair (M, d) is a "metric space."

By a slight abuse of notation, for simplicity, we sometimes denote a generic metric space (M, d), which is formally a pair, by the name of the set on which the metric is defined, i.e., M.

Normed spaces (Definition 7.31) are canonically endowed with a metric and are thus special cases of metric spaces:

Exercise 7.92 Let $(X, \|\cdot\|)$ be any normed space and define the function $d_{\|\cdot\|} : X \times X \to [0, \infty)$ by

$$d_{\|\cdot\|}(x, x') \doteq \left\|x - x'\right\| \ , \qquad x, x' \in X \ .$$

Show that $(X, d_{\|\cdot\|})$ is a metric space.

We call the metric $d_{\|\cdot\|}$ defined in the last exercise the "metric associated with the norm $\|\cdot\|$." If $\|\cdot\| = \|\cdot\|_{\mathrm{e}}$, i.e., $\|\cdot\|$ is the Euclidean norm in $\mathbb{R}^D$ or $\mathbb{C}^D$, $D \in \mathbb{N}$, of Definition 7.32, then the associated metric is called the "Euclidean metric" of the corresponding space.

Cartesian products of metric spaces are naturally endowed with a metric:

Definition 7.93 Let $(M_1, d_1), \ldots, (M_n, d_n)$, $n \in \mathbb{N}$, be any finite collection of metric spaces. Then we define a metric d in the Cartesian product $M_1 \times \cdots \times M_n$ by

$$d((p_1, \ldots, p_n), (p_1', \ldots, p_n')) \doteq \max\{d_1(p_1, p_1'), \ldots, d_n(p_n, p_n')\}$$

for $(p_1, \ldots, p_n), (p_1', \ldots, p_n') \in M_1 \times \cdots \times M_n$.

By default, Cartesian products of metric spaces are endowed here with the above metric. Note that this definition extends Definition 7.33, which refers to the special case of normed spaces.

In the following, we discuss a more "exotic" example of a metric space, which is well-known and important in various application (like in image processing, for instance):

Definition 7.94 (Hausdorff Distance) Take an arbitrary metric space (M, d). Let $\mathcal{P}(M)$ be the family of all subsets of M, i.e., the so-called power set of M. We define $d_{\mathrm{H}} : \mathcal{P}(M) \times \mathcal{P}(M) \to [0, \infty]$ as follows:

$$d_{\mathrm{H}}(\emptyset, \emptyset) \doteq 0 \qquad \text{and} \qquad d_{\mathrm{H}}(\emptyset, \Omega) = d_{\mathrm{H}}(\Omega, \emptyset) \doteq \infty$$

if $\Omega \in \mathcal{P}(M)$ is nonempty, while, for every pair $\Omega, \Omega' \in \mathcal{P}(M)$ of nonempty subsets of M,

$$d_{\mathrm{H}}(\Omega, \Omega') \doteq \max\left\{\sup_{p\in\Omega}\inf_{p'\in\Omega'} d(p, p'),\ \sup_{p'\in\Omega'}\inf_{p\in\Omega} d(p, p')\right\} \ .$$

$d_{\mathrm{H}}(\Omega, \Omega')$ is called the "Hausdorff distance" between $\Omega, \Omega' \in \mathcal{P}(M)$.

Exercise 7.95 Let (M, d) be an arbitrary metric space. Show that the Hausdorff distance associated with this space is symmetric, i.e.,

$$d_{\mathrm{H}}(\Omega, \Omega') = d_{\mathrm{H}}(\Omega', \Omega) \ , \qquad \Omega, \Omega' \in \mathcal{P}(M) \ ,$$

and satisfies the triangle inequality, i.e., for every $\Omega, \Omega', \Omega'' \in \mathcal{P}(M)$,

$$d_{\mathrm{H}}(\Omega, \Omega'') \leq d_{\mathrm{H}}(\Omega, \Omega') + d_{\mathrm{H}}(\Omega', \Omega'') ,$$

with the convention

$$a \leq \infty \qquad \text{and} \qquad a + \infty = \infty , \qquad a \in [0, \infty] .$$

Observe that, in general, the Hausdorff distance defined above is degenerate: For instance, take $M \doteq \mathbb{R}$ and $d(x, x') \doteq |x - x'|$, $x, x' \in M$. In this case, we have

$$d_{\mathrm{H}}([-1, 1], (-1, 1)) = 0 ,$$

even if the subsets $[-1, 1]$, $(-1, 1)$ are not equal.

In fact, by Exercise 7.104, the Hausdorff distance becomes non-degenerate if it is restricted to "closed subsets," defined below, of the underlying metric space. Hence, except for the fact that the Hausdorff distance between two closed subsets is not necessarily finite, the family of all closed subsets of a given metric space (M, d) forms a metric space, the metric of which is the Hausdorff distance. We will see later on that if the closed subsets $\Omega, \Omega' \in \mathcal{P}(M)$ are "compact," then $d_{\mathrm{H}}(\Omega, \Omega') < \infty$, i.e., the family of all compact subsets of a given metric space endowed with the corresponding Hausdorff distance defines a metric space in the usual sense.

Definition 7.96 (ϵ-Neighborhood) Let (M, d) be an arbitrary metric space. For all $p \in M$ and $\varepsilon > 0$, the "ε-neighborhood of p" or the "open ball" of radius ε centered at p is the subset

$$B_{\varepsilon}(p) \doteq \{p' \in M \ : \ d(p', p) < \varepsilon\} \subseteq M .$$

In what follows, we introduce some important notions of the theory of metric spaces, related to properties of ε-neighborhoods.

Definition 7.97 (Open and Closed Sets) Let (M, d) be an arbitrary metric space.

(i) We say that the subset $\Omega \subseteq M$ is "open" (in (M, d)) if, for any $p \in \Omega$, one has $B_{\varepsilon}(p) \subseteq \Omega$ for some $\varepsilon > 0$. In particular, the empty set, as well as the full space M, is open (in (M, d)).
(ii) Ω is, by definition, "closed" (in (M, d)) if its complement $\Omega^c \doteq M \backslash \Omega$ is open (in (M, d)).
(iii) Ω is, by definition, a "clopen" (of (M, d)) if its simultaneously open and closed in (M, d). In particular, the empty set and M are always clopens of (M, d).
(iv) The family

$$\tau_d \doteq \{O \subseteq M \ : \ O \text{ open in } (M, d)\}$$

of all subsets of M that are open in (M, d) is called the "topology of M associated with the metric d."

Example 7.98 Let (M, d) be any metric space. The triangle inequality immediately implies the following properties:

(i) For all $p \in M$ and $\varepsilon > 0$, $B_\varepsilon(p)$ is open in (M, d).
(ii) For all $p \in M$ and $\varepsilon > 0$, the "closed ball" defined by

$$\overline{B}_\varepsilon(p) \doteq \{p' \in M \ : \ d(p', p) \leq \varepsilon\} \subseteq M$$

is closed in (M, d).

The following simple properties of open subsets are very important:

Exercise 7.99 Let (M, d) be an arbitrary metric space. Prove the following assertions:

(i) For any family of open subsets $O_i \in \tau_d$, $i \in I$, the union $\cup_{i \in I} O_i \subseteq M$ is again an open subset.
(ii) For any *finite* collection $O_1, \ldots, O_N \in \tau_d$, $N \in \mathbb{N}$, of open subsets, the intersection

$$O_1 \cap \cdots \cap O_N \subseteq M$$

is open.

In fact, the above two properties of open subsets of metric spaces are axioms of the theory of general topological spaces (in its most usual form). Recall that the closed subsets of a metric spaces are, by definition, the complement of the open subsets of this space. In particular, it follows (by using de Morgan's law) that arbitrary intersections of closed subsets of metric spaces are closed and that finite unions of open subsets are open.

Exercise 7.100 Let (M, d) be any metric space. Take an arbitrary nonempty subset $\tilde{M} \subseteq M$ and let $\tilde{d}$ be the restriction of the metric d to the subset $\tilde{M}$. Note that $(\tilde{M}, \tilde{d})$ is a new metric space. Show that

$$\tau_{\tilde{d}} = \{\tilde{M} \cap O \ : \ O \in \tau_d\} \ .$$

Definition 7.101 (Interior and Closure) Let $M \equiv (M, d)$ be any metric space and $\Omega \subseteq M$ an arbitrary subset.

(i) The subset

$$\Omega^\circ \doteq \bigcup_{p \in \Omega, \ \varepsilon > 0 \text{ with } B_\varepsilon(p) \subseteq \Omega} B_\varepsilon(p) \subseteq \Omega$$

is, by definition, the "interior" of Ω (in M).
(ii) $\overline{\Omega} \doteq ((\Omega^c)^\circ)^c \supseteq \Omega$ is the "closure" of Ω (in M).

Exercise 7.102 Let M be any metric space and $\Omega \subseteq M$ an arbitrary subset. Show that Ω is open in M iff $\Omega = \Omega^{\circ}$ and that it is closed iff $\Omega = \overline{\Omega}$.

Exercise 7.103 Let M be any metric space and $\Omega \subseteq M$ an arbitrary subset. Show that Ω° (the interior of Ω in M) is the largest subset of Ω which is open in M, whereas $\overline{\Omega}$ (the closure of Ω in M) is the smallest subset of M which contains Ω and is closed in M.

Exercise 7.104 Let (M, d) be an arbitrary metric space. Prove that, for any pair $\Omega, \Omega' \in \mathcal{P}(M)$ of subsets that are closed in (M, d), one has that $\Omega = \Omega'$ iff $d_{\mathrm{H}}(\Omega, \Omega') = 0$.

It frequently occurs that two different metrics, d_1 and d_2, define the same topology in a set M, i.e., one has $\tau_{d_1} = \tau_{d_2}$. For instance, we have the following important example for normed spaces: Let V be a vector space and $\|\cdot\|^{(1)}$, $\|\cdot\|^{(2)}$ two norms in this space. We say that $\|\cdot\|^{(1)}$ and $\|\cdot\|^{(2)}$ are "equivalent norms" if, for some $C \in [1, \infty)$,

$$\frac{1}{C}\|v\|^{(2)} \leq \|v\|^{(1)} \leq C\|v\|^{(2)}\,, \qquad v \in V\,.$$

Exercise 7.105 Let $\|\cdot\|^{(1)}$,$\|\cdot\|^{(2)}$ be two norms in a vector space, as above. Prove that these norms are equivalent iff $\tau_{\|\cdot\|^{(1)}} = \tau_{\|\cdot\|^{(2)}}$.

Still in this context, the following result is well-known:

Theorem 7.106 *Let V be any vector space of* finite *dimension. If d_1 and d_2 are two metrics in V associated with any two norms in V, then $\tau_{d_1} = \tau_{d_2}$. In particular, any two norms in a finite-dimensional vector space are equivalent to each other.*

Observe that we provide in Theorem 7.120 a more general version of the above statement, along with a reference to its proof.

We say that a given property of a metric space is "topological," whenever this property only depends on the topology associated with the metric, but not on other details of the metric. As we will see later on, this is the case for various important properties referring to metric spaces, like "convergence" of sequences, "continuity" of functions, "compactness" of subsets, and many others. In particular, by the last theorem, in normed spaces of finite dimension, such properties are completely independent of the particular choice of a norm. However, there are also many important properties of metric spaces which do depend on details of the given metric, and not only on its corresponding topology. This is the case, for instance, for the "completeness" property of metric spaces, which is crucial in many important applications of metric spaces and will be discussed below.

7.2.2 Continuous Functions

We start with the usual definition of continuous functions on metric spaces:

Definition 7.107 (Continuous Functions) Let (M_1, d_1) and (M_2, d_2) be two arbitrary metric spaces.

(i) We say that the function $f : M_1 \to M_2$ is "continuous at $p_1 \in M_1$" (with respect to the metrics d_1 and d_2) if, for any $\varepsilon > 0$, there is $\delta > 0$ (which may depend on the choice of $p_1 \in M_1$) such that

$$f(B_\delta(p_1)) \subseteq B_\varepsilon(f(p_1)) .$$

(ii) f is, by definition, "continuous" if it is continuous at every $p_1 \in M_1$.
(iii) $C(M_1; M_2)$ denotes the set of all continuous functions $M_1 \to M_2$.
(iv) $f \in C(M_1; M_2)$ is said to be "uniformly continuous" if, for any fixed $\varepsilon > 0$, there is $\delta > 0$ such that, for *all* $p_1 \in M_1$,

$$f(B_\delta(p_1)) \subseteq B_\varepsilon(f(p_1)) .$$

Example 7.108 Consider the (usual) metric in the real line, defined by $d(x, x') \doteq |x - x'|$, $x, x' \in \mathbb{R}$.

(i) The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) \doteq x^2$, $x \in \mathbb{R}$, is continuous, but not uniformly continuous.
(ii) The function $f : \mathbb{R} \to \mathbb{R}$, $f(x) \doteq \sin(x)$, $x \in \mathbb{R}$, is uniformly continuous.
(iii) For any normed space $(X, \|\cdot\|)$, the norm itself $\|\cdot\| : X \to \mathbb{R}$ is a uniformly continuous function.

Recall that if Ω is any nonempty set and V a vector space, then the set $\mathcal{F}(\Omega; V)$ of all functions $\Omega \to V$ is canonically seen as a vector space, with the pointwise vector space operations of V. If M is a metric space and X a normed space, then the continuous functions $M \to X$ form a vector subspace of $\mathcal{F}(M; X)$:

Exercise 7.109 Let M be any metric space and X an arbitrary normed space. Show that $C(M; X)$ is a vector subspace of $\mathcal{F}(M; X)$.

The continuity of functions between metric spaces is equivalent to the preservation of closedness or openness of subsets with respect to preimages through these functions:

Exercise 7.110 Let (M_1, d_1) and (M_2, d_2) be arbitrary metric spaces and f any function $M_1 \to M_2$. Show that the following three properties of f are equivalent:

(i) f is continuous (with respect to the metrics d_1 and d_2).
(ii) For every $O_2 \in \tau_{d_2}$, the preimage

$$f^{-1}(O_2) \doteq \{p_1 \in M_1 : f(p_1) \in O_2\} \subseteq M_1$$

is open in (M_1, d_1), i.e., $f^{-1}(O_2) \in \tau_{d_1}$.

(iii) For any subset $C_2 \subseteq M_2$ that is closed in (M_2, d_2), the preimage $f^{-1}(C_2) \subseteq M_1$ is closed in (M_1, d_1).

The following result is a simple but important consequence of the last exercise:

Corollary 7.111 *Let M_0, M_1, and M_2 be arbitrary metric spaces. Take two continuous functions $f_1 : M_0 \to M_1$ and $f_2 : M_1 \to M_2$. The composition $f_2 \circ f_1 : M_0 \to M_2$, where*

$$f_2 \circ f_1(p_0) \doteq f_2(f_1(p_0)) , \qquad p_0 \in M_0 ,$$

is again continuous.

We introduce now an important special case of continuous transformation between metric spaces, the so-called homeomorphisms:

Definition 7.112 (Homeomorphism) Let M_1 and M_2 be arbitrary metric spaces. The mapping $\varphi : M_1 \to M_2$ is a "homeomorphism" (between the metric spaces M_1 and M_2) if it is a continuous one-to-one correspondence, whose inverse is continuous. Two metric spaces are said to be "homeomorphic," whenever there is a homeomorphism between them.

Note that if $\varphi : M_1 \to M_2$ is a homeomorphism, then so does its inverse $\varphi^{-1} : M_2 \to M_1$. Observe also that the homeomorphisms between two metric spaces are exactly the continuous one-to-one correspondences which preserve the openness of subsets:

Exercise 7.113 Let (M_1, d_1) and (M_2, d_2) be two arbitrary homeomorphic metric spaces. Show that, for every homeomorphism $\varphi : M_1 \to M_2$, the mapping $O_1 \mapsto \varphi(O_1)$ is a one-to-one correspondence $\tau_{d_1} \to \tau_{d_2}$. Show also that, conversely, any continuous one-to-one correspondence $\varphi : M_1 \to M_2$ such that, for all $O_1 \in \tau_{d_1}$, $\varphi(O_1) \in \tau_{d_2}$, is a homeomorphism.

Because of the property stated in the first part of the last exercise, we say that homeomorphic metric spaces are "topologically equivalent."

Definition 7.114 (Isometry) Let (M_1, d_1) and (M_2, d_2) be two arbitrary metric spaces. The mapping $\varphi : M_1 \to M_2$ is an "isometry" if, for every $p_1, p_1' \in M_1$,

$$d_2(\varphi(p_1), \varphi(p_1')) = d_1(p_1, p_1') ,$$

i.e., the mapping preserves distances. (M_1, d_1) and (M_2, d_2) are said to be "equivalent metric spaces," whenever there is a one-to-one correspondence $M_1 \to M_2$ which is an isometry.

Exercise 7.115 Show that any isometry between metric spaces is continuous and that the inverse of such an isometry is again an isometry, when it exists. In particular, any one-to-one correspondence that is an isometry between metric spaces is a homeomorphism, that is, equivalent metric spaces are always homeomorphic.

Note that any two isomorphic normed spaces (see Definition 7.31) are always equivalent, in the sense of metric spaces.

In the case of normed spaces and linear functions, we have the following equivalent characterizations of continuity:

Exercise 7.116 Let X_1 and X_2 be any two normed spaces and $\Theta \in \mathcal{L}(X_1; X_2)$ an arbitrary linear function $X_1 \to X_2$. Show that the following properties are equivalent:

(i) Θ is a continuous function $X_1 \to X_2$.
(ii) $\Theta \in \mathcal{B}(X_1; X_2)$, i.e., Θ is a bounded linear operator.
(iii) $\Theta \in \mathcal{L}(X_1; X_2)$ is continuous at $0 \in X_1$.
(iv) Θ is a uniformly continuous function $X_1 \to X_2$.

7.2.3 Metric Vector Spaces and Locally Convex Spaces

In this subsection, we define so-called metric vector spaces and discuss some (here) relevant examples, as well as basic properties of these spaces. In fact, metric vector spaces are nothing else than vector spaces that are endowed with some metric that is compatible with the vector space operations, in the following sense:

Definition 7.117 (Metric and Locally Convex Vector Spaces) Let V be a vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, and d a metric in V.

(i) (V, d) is a "metric vector space" if the vector space operations of V, that is, the addition $+ : V \times V \to V$ and the scalar multiplications $\cdot : \mathbb{K} \times V \to V$, are continuous with respect to the Cartesian product metric for $V \times V$ and $\mathbb{K} \times V$ (see Definition 7.93).
(ii) The metric vector space (V, d) is a "locally convex vector space" if, for all $v \in V$, the open ball $B_\varepsilon(v) \in V$ is convex for sufficiently small $\varepsilon > 0$.
(iii) The metric d of a metric vector space (V, d) is "invariant" if

$$d(v + v'', v' + v'') = d(v, v') , \qquad v, v', v'' \in V .$$

(iv) A metric vector space (V, d) is a "Fréchet space" if it is a locally convex vector space, the metric d is invariant, and (V, d) is a complete metric space.

One simple, but very important, example of a metric vector space that is a locally convex vector spaces is given by any normed space, naturally seen as a metric space. Note in this example that Banach spaces are Fréchet spaces. Thus, the latter can be seen as some generalization of the former. Similar to the normed case (Exercise 7.137), the following assertion holds true:

Exercise 7.118 Let (V, d) be a metric vector space. Show that if (V, d) is locally convex and d is an invariant metric, then there is a completion $(\bar{V}, \mathrm{i})$ (as a metric space) of (V, d) such that $\bar{V}$ is a Fréchet space.

Exercise 7.119 Let $(V_1, d_1), \ldots, (V_n, d_n)$ be $n \in \mathbb{N}$ metric vector spaces. Show that the Cartesian product $V_1 \times \cdots \times V_n$, along with the metric of Definition 7.93, is again a metric vector space. Show additionally that $V_1 \times \cdots \times V_n$, with the same metric, is a locally convex space, if $V_1, \ldots, V_n$ are locally convex spaces.

From a purely topological point of view and in the case of finite-dimensional vector spaces, it turns out that normed spaces are already the most general example of a metric topological space, because of the following generalization of Theorem 7.106:

Theorem 7.120 *Let V be any* finite-dimensional *vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. If d_1 and d_2 are two metrics in V with respect to which V is a metric vector space, then $\tau_{d_1} = \tau_{d_2}$.*

For a (more general) version of the above theorem for general (i.e., not necessarily metric) vector space topologies, along with a complete proof, see, for instance, [18, 1.21(a) Theorem]. By contrast, in infinite dimension, not every metric vector space is topologically equivalent to a normed one. The following example is very important here, and it is not of normed type in infinite dimensions. If refers to Definition 4.80, which we recall below for convenience:

Definition 7.121 Let X be any normed space. For any sequence $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ in X, we define the mapping $d_{\mathcal{S}} : X^{\mathrm{td}} \times X^{\mathrm{td}} \to \mathbb{R}_0^+$ by

$$d_{\mathcal{S}}(\varphi, \varphi') \doteq \max_{n\in\mathbb{N}} \frac{2^{-n}|(\varphi - \varphi')(x_n)|}{1 + |(\varphi - \varphi')(x_n)|}\ , \qquad \varphi, \varphi' \in X^{\mathrm{td}}\ .$$

If X is *separable* and the sequence $\mathcal{S}$ is *dens*e in X, then $d_{\mathcal{S}}$ is an invariant metric in the (topological) dual space X^{td}. Observe that this property refers to Exercise 4.82. Thus, in this case, we have the following example:

Example 7.122 Let X be any separable normed space and $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ a dense sequence in X. Then, $(X^{\mathrm{td}}, d_{\mathcal{S}})$ is a metric vector space.

Exercise 7.123 Show that the metric vector space of the last example is a locally convex vector space.

We discuss next very general properties of continuous linear transformations between metric vector spaces. In particular, we give the general structure of continuous linear functionals on $(X^{\mathrm{td}}, d_{\mathcal{S}})$.

Exercise 7.124 Let V be any metric vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For fixed $v \in V$ and $\alpha \in \mathbb{K}\backslash\{0\}$, define the transformations $\tau_v, \vartheta_\alpha : V \to V$ by

$$\tau_v(v') = v + v' \qquad \text{and} \qquad \vartheta_\alpha(v') = \alpha v'\ , \qquad v' \in V\ .$$

Show that these transformations are homeomorphisms.

Exercise 7.125 Let V_1 and V_2 be any two metric vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Let $\Theta : V_1 \to V_2$ be any linear transformation. Show that Θ is continuous iff it is continuous at $0 \in V_1$. (See Definition 7.107.)

Hint: Use the preceding exercise.

In the following proposition, we show that every continuous linear function on the metric vector space $(X^{\mathrm{td}}, d_{\mathcal{S}})$ of Example 7.122 corresponds to the evaluation of linear functional on some element of the normed space X:

Proposition 7.126 (Continuous Linear Functionals on $(X^{\mathrm{td}}, d_{\mathcal{S}})$) *Let X be any separable normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\mathcal{S} = (x_n)_{n\in\mathbb{N}}$ a dense sequence in X. The linear functional φ on X^{td} is continuous with respect to the metric $d_{\mathcal{S}}$ iff there is some linear combination $x \in \mathrm{span}\{x_n : n \in \mathbb{N}\}$ of elements of the sequence $\mathcal{S}$, such that*

$$\varphi(\rho) = \rho(x) , \qquad \rho \in X^{\mathrm{td}}.$$

Proof It is easy to check that, for any $x \in \mathrm{span}\{x_n : n \in \mathbb{N}\}$, the linear functional φ_x on X^{td} defined by

$$\varphi_x(\rho) = \rho(x) , \qquad \rho \in X^{\mathrm{td}} ,$$

is continuous with respect to the metric $d_{\mathcal{S}}$. Thus, take now any linear functional φ on X^{td} which is continuous with respect to the metric $d_{\mathcal{S}}$. By continuity, there is some $\varepsilon > 0$ such that, for all $\rho \in X^{\mathrm{td}}$ with $d_{\mathcal{S}}(0, \rho) \leq \varepsilon$, one has that $|\varphi(\rho)| \leq 1$. Take some $N \in \mathbb{N}$ such that $\varepsilon \geq 2^{1-N}$, which yields the inequalities

$$\begin{aligned} d_{\mathcal{S}}(0, \rho) \doteq \max_{n\in\mathbb{N}} \frac{2^{-n}|\rho(x_n)|}{1 + |\rho(x_n)|} &\leq \frac{\varepsilon}{2} + \max_{n\in\{1,\ldots,N\}} \frac{2^{-n}|\rho(x_n)|}{1 + |\rho(x_n)|} \leq \frac{\varepsilon}{2} \\ &+ \max_{n\in\{1,\ldots,N\}} 2^{-n}|\rho(x_n)| . \end{aligned}$$

Therefore, if

$$\max_{n\in\{1,\ldots,N\}} |\rho(2^{-n}x_n)| \leq \frac{\varepsilon}{2} ,$$

then $d_{\mathcal{S}}(0, \rho) \leq \varepsilon$, which, by the choice of ε, in turn implies that $|\varphi(\rho)| \leq 1$. By linearity of φ, we then arrive at the bound

$$|\varphi(\rho)| \leq 2\varepsilon^{-1} \max_{n\in\{1,\ldots,N\}} |\rho(2^{-n}x_n)|$$

for all $\rho \in X^{\mathrm{td}}$. But this bound implies that there are constants $\alpha_1, \ldots, \alpha_N \in \mathbb{K}$ such that

$$\varphi(\rho) = \alpha_1 \rho(x_1) + \cdots + \alpha_n \rho(x_n) , \qquad \rho \in X^{\mathrm{td}} .$$

For a complete proof of this last step, see, for instance, [18, 3.9 Lemma]. □

The above result on the structure of general continuous linear functionals on $(X^{\mathrm{td}}, d_{\mathcal{S}})$ is pivotal in Sect. 7.5 to prove important properties of the Legendre-Fenchel transform that are, in turn, essential in Chap. 5 for the study of equilibrium states of infinite quantum lattices.

7.2.4 Convergent Sequences and Nets

In this subsection, we give a brief account of the notion of convergence in metric spaces. We generally focus on sequences in the present work, but we still shortly mention the case of more general nets, for they are sometimes more natural than sequences. We start with Cauchy sequences and their convergence by straightforwardly generalizing Definition 7.80 to the metric case, as follows:

Definition 7.127 (Cauchy Sequences and Convergence) Let (M, d) be a metric space and $p_n \in M$, $n \in \mathbb{N}$, a sequence in M.

(i) $(p_n)_{n\in\mathbb{N}}$ is a "Cauchy sequence" if, for any $\varepsilon > 0$, there is $N_\varepsilon \in \mathbb{N}$ such that, for all $m, n \in \mathbb{N}$ satisfying $m, n \geq N_\varepsilon$, one has that $d(p_{n'}, p_n) \leq \varepsilon$.
(ii) The sequence $(p_n)_{n\in\mathbb{N}}$ is "convergent" if, for some $p \in M$ and all $\varepsilon > 0$, there is $N_\varepsilon \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$, $n \geq N_\varepsilon$, one has that $d(p_n, p) \leq \varepsilon$. In this case, the point p is denoted by $\lim_{n\to\infty} p_n$, which is the "limit" of the sequence $(p_n)_{n\in\mathbb{N}}$.
(iii) Let $\Omega \subseteq M$ any nonempty subset and $p_n \in \Omega$, $n \in \mathbb{N}$, a convergent sequence. We say that $(p_n)_{n\in\mathbb{N}}$"converges in Ω," whenever

$$\lim_{n\to\infty} p_n \in \Omega .$$

(iv) Consider two subsets $\tilde{\Omega} \subseteq \Omega \subseteq M$. We say that $\tilde{\Omega}$ is "dense" in Ω if, for every $p \in \Omega$, there is a convergent sequence $p_n \in \tilde{\Omega}$, $n \in \mathbb{N}$, whose limit is p, i.e.,

$$\lim_{n\to\infty} p_n = p .$$

(v) The metric space M is "separable" if it has a countable dense subset, i.e., for some sequence $(p_n)_{n\in\mathbb{N}}$, the subset $\{p_n : n \in \mathbb{N}\} \subseteq \Omega$ is dense in Ω.

In metric spaces, one may introduce the notion of convergence and Cauchy property for nets, also called Moore-Smith sequences (see Definition 1.15), that are more general than usual sequences:

Definition 7.128 (Net Convergence) Let $(p_i)_{i\in I}$ be a net in a metric space M. The net is a "Cauchy net" if, for every $\varepsilon > 0$, there is $i_\varepsilon \in I$ such that, for all $i, i' \geq i_\varepsilon$, one has that $d(p_i, p_{i'}) \leq \varepsilon$. It converges to $p \in M$ whenever, for every $\varepsilon > 0$, there is $i_\varepsilon \in I$ such that, for all $i \geq i_\varepsilon$, one has that $d(p_i, p) \leq \varepsilon$. In this case, we write

$$\lim_{i\in I} p_i = p \ .$$

Any continuous function (Definition 7.107) is completely determined by the values it takes on any dense subset of its domain:

Lemma 7.129

(i) Let (M_1, d_1) and (M_2, d_2) be two metric spaces. Take a dense subset $\tilde{M}_1 \subseteq M_1$ in M_1, seen as a metric space by restriction of the metric d_1 to $\tilde{M}_1$. Let $\tilde{f} : \tilde{M}_1 \to M_2$ be a continuous function. Then, there is at most one continuous function $f : M_1 \to M_2$ that extends $\tilde{f}$, i.e., $\tilde{f}(p_1) = f(p_1)$ for all $p_1 \in \tilde{M}_1 \subseteq M_1$. Such a (unique) function f is called the "continuous extension" of $\tilde{f}$ to M_1.

(ii) Let X_1 be a normed space and X_2 a Banach space (Definition 7.82). Take a dense vector subspace $\tilde{X}_1 \subseteq X_1$, seen as a normed space by restriction of the norm of X_1 to $\tilde{X}_1$. Then, every $\tilde{\Theta} \in \mathcal{B}(\tilde{X}_1; X_2) \subseteq C(\tilde{X}_1; X_2)$ has a (unique) continuous extension Θ to X_1. Moreover, $\Theta \in \mathcal{B}(X_1; X_2)$ and one has $\|\Theta\|_{\text{op}} = \|\tilde{\Theta}\|_{\text{op}}$.

Proof Exercise. □

Exercise 7.130 Let M be a nonempty set and d_1,d_2 two metrics in M such that $\tau_{d_1} = \tau_{d_2}$ (Definition 7.97). Let $(p_i)_{i\in I}$ be any net in M.

(i) Show that the net converges in (M, d_1) iff it converges in (M, d_2) and that, in this case, the corresponding limits are the same.

(ii) Show that if M is a vector space and d_1,d_2 are metrics associated with norms in this space, then the net has the Cauchy property in (M, d_1) iff it has the same property in (M, d_2).

Remark 7.131

(i) By the first part of the last exercise, the convergence of a sequence in a metric space is a topological property.

(ii) The same is not true, in general, for the Cauchy property: There are sets M, metrics d_1, d_2 in M with $\tau_{d_1} = \tau_{d_2}$, and sequences $p_n \in M$, $n \in \mathbb{N}$, that are Cauchy in (M, d_1), but not in (M, d_2).

(iii) By the second part of the exercise, the Cauchy property of sequences is topological in *normed* spaces. In particular, by Theorem 7.106, for normed spaces of finite dimension, it is not necessary to specify the norm with respect to which a given sequence in such a space is Cauchy.

Exercise 7.132 Show that any convergent net in a metric space has the Cauchy property and that its limit is unique.

However, an arbitrary Cauchy sequence in a metric space is in general not convergent. This fact motivates the following definition, which is reminiscent of Definition 7.82 for normed spaces:

Definition 7.133 (Complete Spaces) A metric space is "complete" ("sequentially complete") if every net (sequence) that has the Cauchy property in this space is convergent. Sequentially complete normed spaces are called "Banach spaces."

In fact, one can easily check that the completeness and the sequential completeness are equivalent properties of metric spaces. In finite dimension, the situation is particularly simple because of the following fact:

Theorem 7.134 (Completeness of Normed Spaces of Finite Dimension) *Every Cauchy net in a normed space of finite dimension is convergent.*

Exercise 7.135 Show that the Cartesian product of a finite collection of complete metric spaces, in the sense of Definition 7.93, is again a complete metric space.

General metric spaces that are not complete can always be "completed," in the following sense:

Theorem 7.136 (Existence of the Completion of Metric Spaces) *Let (M, d) be any (not necessarily complete) metric space. Then there is a complete metric space $(\bar{M}, \bar{d})$ such that $M \subseteq \bar{M}$ is dense in $\bar{M}$ and the metric $\bar{d}$ coincides with d in M. $(\bar{M}, \bar{d})$ is called "completion" of (M, d) and is unique up to an equivalence of metric spaces: If $(\bar{M}', \bar{d}')$ is a second completion of (M, d), then there is an isometric one-to-one correspondence $\varphi : \bar{M}' \to \bar{M}$. In this case, φ can be chosen such that $\varphi(p) = p$ for all $p \in M \subseteq \bar{M}' \cap \bar{M}$.*

Idea of Proof The proof of existence of a completion for an arbitrary metric space is standard. It is a direct adaptation of Cantor's construction of the real number. Thus, we give it for completeness, but only describe its main lines and refrain from working out all details.

1. Let (M, d) be any metric space. We say that two Cauchy sequences $(p_n)_{n\in\mathbb{N}}$, $(p'_n)_{n\in\mathbb{N}}$ in M are equivalent if

$$\lim_{n\to\infty} d(p_n, p'_n) = 0 .$$

In this case, we write $(p_n)_{n\in\mathbb{N}} \sim (p'_n)_{n\in\mathbb{N}}$. $\mathcal{Y}(M, d)$ denotes the set of all Cauchy sequences in M. Note that $\sim$ defines an equivalence relation in $\mathcal{Y}(M, d)$ and that $(p_n)_{n\in\mathbb{N}} \sim (p'_n)_{n\in\mathbb{N}}$ whenever $(p'_n)_{n\in\mathbb{N}}$ is a subsequence of $(p_n)_{n\in\mathbb{N}}$. Any point $p \in M$ is seen as the constant (Cauchy) sequence $(p_n = p)_{n\in\mathbb{N}} \in \mathcal{Y}(M, d)$ and, thus, there is a natural inclusion $M \to \mathcal{Y}(M, d)$.

2. Let $\bar{M}$ be the set of equivalence classes $\mathcal{Y}(M,d)/\sim$ (named the quotient space of $\mathcal{Y}(M,d)$ by $\sim$). For all $p, p' \in M$, observe that $p \sim p'$ in $\mathcal{Y}(M,d)$ iff $p = p'$ in M. Therefore, there is again a natural inclusion $M \to \bar{M}$. Define the symmetric functions

$$\tilde{d} : \mathcal{Y}(M,d) \times \mathcal{Y}(M,d) \to [0, \infty)$$

by

$$\tilde{d}((p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}}) \doteq \limsup_{n\to\infty} d(p_n, p'_n) \ .$$

One checks that this function has the following properties:

(a) For all $(p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}} \in \mathcal{Y}(M,d)$,

$$\tilde{d}((p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}}) = 0 \qquad \text{iff} \qquad (p_n)_{n\in\mathbb{N}} \sim (p'_n)_{n\in\mathbb{N}} \ .$$

(b) For all $(p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}} \in \mathcal{Y}(M,d)$ and $(q_n)_{n\in\mathbb{N}}, (q'_n)_{n\in\mathbb{N}} \in \mathcal{Y}(M,d)$ with $(p_n)_{n\in\mathbb{N}} \sim (q_n)_{n\in\mathbb{N}}$ and $(p'_n)_{n\in\mathbb{N}} \sim (q'_n)_{n\in\mathbb{N}}$, one has

$$\tilde{d}((p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}}) = \tilde{d}((q_n)_{n\in\mathbb{N}}, (q'_n)_{n\in\mathbb{N}}) \ .$$

(c) For all $(p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}}, (p''_n)_{n\in\mathbb{N}} \in \mathcal{Y}(M,d)$,

$$\tilde{d}((p_n)_{n\in\mathbb{N}}, (p''_n)_{n\in\mathbb{N}}) \leq \tilde{d}((p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}}) + \tilde{d}((p'_n)_{n\in\mathbb{N}}, (p''_n)_{n\in\mathbb{N}}) \ .$$

(d) For all $(p_n)_{n\in\mathbb{N}} \in \mathcal{Y}(M,d)$,

$$\lim_{N\to\infty} \tilde{d}((p_n = p_N)_{n\in\mathbb{N}}, (p_n)_{n\in\mathbb{N}}) = 0 \ .$$

3. From these properties of the symmetric function $\tilde{d}$, one shows the existence of a unique metric $\bar{d}$ in $\bar{M}$ satisfying

$$\bar{d}([(p_n)_{n\in\mathbb{N}}], [(p'_n)_{n\in\mathbb{N}}]) = \tilde{d}((p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}}) \ ,$$
$$(p_n)_{n\in\mathbb{N}}, (p'_n)_{n\in\mathbb{N}} \in \mathcal{Y}(M,d) \ ,$$

where $[(p_n)_{n\in\mathbb{N}}]$ and $[(p'_n)_{n\in\mathbb{N}}]$ are, respectively, the equivalence classes of the Cauchy sequences $(p_n)_{n\in\mathbb{N}}$ and $(p'_n)_{n\in\mathbb{N}}$. The restriction of $\bar{d}$ to $M \subseteq \bar{M}$ is the original metric d, and M is a dense subset of $\bar{M}$. Finally, to show that $(\bar{M}, \bar{d})$ is a complete metric space, we take an arbitrary Cauchy sequence $([(p_k^{(n)})_{k\in\mathbb{N}}])_{n\in\mathbb{N}}$ in $(\bar{M}, \bar{d})$ and show that it converges to $[(p_n^{(n)})_{n\in\mathbb{N}}]$, which is the equivalence class of the "diagonal (Cauchy) sequence" $(p_n^{(n)})_{n\in\mathbb{N}} \in \mathcal{Y}(M,d)$.

□

The last two theorems are well-known in analysis and are not (completely) proven here.

Exercise 7.137 Show that if M is a vector space and d is the metric associated with some norm $\|\cdot\|$ in M, then the completion $(\bar{M}, \bar{d})$ of (M, d) can be chosen in such a way that also $\bar{M}$ is a vector space, M being a vector subspace $\bar{M}$, and $\bar{d}$ is the metric associated with a (unique) norm $\|\cdot\|^{\sim}$ in $\bar{M}$. In this case, we say that $(\bar{M}, \|\cdot\|^{\sim})$ is a "normed space completion" of $(M, \|\cdot\|)$. Prove also that two normed space completions of a normed space are necessarily equivalent normed spaces (or isomorphic normed spaces, in the sense of Definition 7.31 (ii)).

Recall that we had already given a notion of "completion" of a normed space, in Definition 7.88. In fact, if $(\bar{X}, \|\cdot\|^{\sim})$ is a completion of the normed space $(X, \|\cdot\|)$ and we take $\iota : X \to \bar{X}$ as being the inclusion mapping of the subset $X \subseteq \bar{X}$, then $(\bar{X}, \iota)$ is a completion of $(X, \|\cdot\|)$ in the sense of Definition 7.88, by Lemma 7.129 (ii).

Alternative to Definitions 7.97 (ii) and 7.107, respectively, it is possible to define closed subsets and continuous functions in terms of the convergence of sequences:

Exercise 7.138

(i) Let M be any metric space, $\Omega \subseteq M$ an arbitrary nonempty closed subset, and $(p_i)_{i\in I}$ a net of elements of Ω, which is convergent in M. Show that this net converges in Ω, i.e., its limit belongs to the subset Ω itself.

(ii) Let M_1 and M_2 be two metric spaces and $f : M_1 \to M_2$ a function, which is continuous in $p_1 \in M_1$. Prove that, for any net $(p'_i)_{i\in I}$ in M_1 such that $\lim_{i\in I} p'_i = p_1$, the (image) net $(f(p'_i))_{i\in I}$ converges in M_2 and

$$\lim_{i\in I} f(p'_i) = f(p_1) \ .$$

In fact, both properties stated in the last exercise are not only necessary conditions for closedness and continuity, respectively, but they are also sufficient:

Theorem 7.139

(i) Let M be a metric space and $\Omega \subseteq M$ a nonempty subset. Ω is closed in M iff, for any sequence $p_n \in \Omega$, $n \in \mathbb{N}$, that converges in M, one has $\lim_{n\to\infty} p_n \in \Omega$.

(ii) Let M_1 and M_2 be two metric spaces and $f : M_1 \to M_2$ an arbitrary function. f is continuous in $p_1 \in M_1$ iff, for any sequence $(p'_n)_{n\in\mathbb{N}}$ that converges in M_1 and whose limit is p_1, the (image) sequence $(f(p'_n))_{n\in\mathbb{N}}$ converges in M_2 and one has

$$\lim_{n\to\infty} f(p'_n) = f(p_1) \ .$$

Exercise 7.140 Let M be a metric space and $\Omega \subseteq M$ an arbitrary subset. Prove that the closure of Ω is given by

$$\overline{\Omega} = \left\{ p \in M \ : \ p = \lim_{n\to\infty} p_n \text{ for some sequence } p_n \in \Omega, n \in \mathbb{N} \right\} .$$

Conclude from this that Ω is dense in $\overline{\Omega}$. Show that, for any normed space X, the closures of its vector subspaces are new vector subspaces.

Exercise 7.141 Let (M, d) be any *complete* metric space and $\tilde{M} \subseteq M$ an arbitrary nonempty subset. Recall that the restriction of the metric d to $\tilde{M}$ defines a new metric space $(\tilde{M}, \tilde{d})$. (See Exercise 7.100.) Show that $(\tilde{M}, \tilde{d})$ is also complete, whenever $\tilde{M}$ is closed. More generally, prove that the completion of $(\tilde{M}, \tilde{d})$ can be identified with the closure of $\tilde{M}$ in (M, d), endowed with the restriction of the original metric d to this closure.

7.2.5 Real-Valued Semicontinuous Functions

In this subsection, we define so-called upper and lower semicontinuous real-valued functions on any metric space and study some of their important, albeit simple, properties. Such functions are more general than the continuous ones and are ubiquitous in maximization and minimization problems, respectively. To define these class of functions, we use the following families of open sets of real numbers:

$$\tau_{\text{inf}} \doteq \{(\alpha, \infty) \ : \ \alpha \in \mathbb{R}\} \cup \{\emptyset, \mathbb{R}\} ,$$
$$\tau_{\text{sup}} \doteq \{(-\infty, \alpha) \ : \ \alpha \in \mathbb{R}\} \cup \{\emptyset, \mathbb{R}\} .$$

Definition 7.142 (Semicontinuous Functions) Let (M, d) be any metric space. We say that the function $f : M \to \mathbb{R}$ is "upper semicontinuous" if, for all $O \in \tau_{\text{sup}}$, $f^{-1}(O) \in \tau_d$, i.e., $f^{-1}(O)$ is an open subset of M. Similarly, f is "upper semicontinuous" if, for all $O \in \tau_{\text{inf}}$, $f^{-1}(O) \in \tau_d$.

It directly follows from the above definition that f is upper-(lower-) semicontinuous iff $-f$ is lower-(upper-)semicontinuous. Clearly, by Exercise 7.110, any continuous function $M \to \mathbb{R}$ is simultaneously upper and lower semicontinuous. It turns out that the converse of this last property is also true:

Exercise 7.143 Let M be any metric space. Show that a function $M \to \mathbb{R}$ is continuous iff it is simultaneously upper and lower semicontinuous.

One very important property of semicontinuous functions is their stability with respect to arbitrary (not necessarily finite or countable) suprema and infima:

Lemma 7.144 *Let (M, d) be a metric space and $f_\omega : X \to \mathbb{R}$, $\omega \in \Omega$, any family of functions. If, for all $p \in M$, $\{f_\omega(p) : \omega \in \Omega\}$ is bounded from above (below) in $\mathbb{R}$ and, for every $\omega \in \Omega$, f_ω is lower-(upper-)semicontinuous, then the function* $\sup_{\omega\in\Omega} f_\omega : M \to \mathbb{R}$ *(*$\inf_{\omega\in\Omega} f_\omega : M \to \mathbb{R}$*) is lower-(upper-)semicontinuous.*

Proof Consider the lower semicontinuous case. Then, for all $\alpha \in \mathbb{R}$, one has that

$$\left[\sup_{\omega\in\Omega} f_\omega\right]^{-1}((\alpha,\infty)) = \bigcup_{\omega\in\Omega} f_\omega^{-1}((\alpha,\infty)) \in \tau \ .$$

In other words, the function $\sup_{\omega\in\Omega} f_\omega : M \to \mathbb{R}$ is lower semicontinuous. Mutatis mutandis for the upper semicontinuous case. □

Observe that suprema and infima of any finite collection of continuous real-valued functions on a metric space are again continuous. However, the same is generally wrong for infinite families of continuous functions, even for countable ones. Suprema and infima of countable families of continuous functions are, at least, Borel-measurable (see Definition 4.10). It turns out that semicontinuous functions are always Borel-measurable.

We give now a characterization of semicontinuity via the behavior of the image of convergent nets. With this aim, we need the following (standard) definition:

Definition 7.145 Let I be any directed set and let $(x_i)_{i\in I}$ be a net in $\mathbb{R}$ that is bounded from above. Then we define

$$\limsup_{i\in I} x_i \doteq \inf_{i\in I} \sup\{x_j \ : \ j \geq i\} = \lim_{i\in I} \sup\{x_j \ : \ j \geq i\} \ \in \mathbb{R} \ .$$

Analogously, if the net is bounded from below, then we define

$$\liminf_{i\in I} x_i \doteq \sup_{i\in I} \inf\{x_j \ : \ j \geq i\} = \lim_{i\in I} \inf\{x_j \ : \ j \geq i\} \ \in \mathbb{R} \ .$$

Proposition 7.146 (Semicontinuity and Convergent Nets) *Let M be any metric space. A function $f : M \to \mathbb{R}$ is upper-(lower-)semicontinuous iff, for all $p \in M$ and any net $(p_i)_{i\in I}$ converging to p, one has that the net $(f(p_i))_{i\in I}$ is bounded from above (below) in $\mathbb{R}$ and $\limsup_{i\in I} f(p_i) \leq f(p)$ ($\liminf_{i\in I} f(p_i) \geq f(p)$).*

Proof Exercise. □

We conclude this subsection by remarking that the lower semicontinuity of a real-valued function defined on a metric space is nothing else than the closedness of its epigraph. This observation is important, albeit simple, in various situations where semicontinuous functions appear. See, for instance, the proof of Theorem 7.353, which is a central result for the study of equilibrium states in Chap. 5. See in this context Theorem 5.31, along with its proof.

Definition 7.147 (Epigraph) Let Ω be any nonempty set and $f : \Omega \to \mathbb{R}$ an arbitrary function. We define the "epigraph" of f to be the following subset of the Cartesian product $\Omega \times \mathbb{R}$:

$$\mathrm{epi}(f) \doteq \{(p, s) \ : \ p \in \Omega \ , \ s \geq f(p)\} \subseteq \Omega \times \mathbb{R} \ .$$

Lemma 7.148 *Let (M, d) be any metric space. The function $f : M \to \mathbb{R}$ is lower semicontinuous iff* $\mathrm{epi}(f)$ *is a closed subset of $M \times \mathbb{R}$ with respect to the usual metric for the Cartesian product of metric spaces (see Definition 7.93).*

Proof Exercise. □

7.2.6 Compactness

We now discuss the important topological notion of compactness in general metric spaces. Compact sets in a metric space refer to a particular type of closed and bounded sets in this spaces. In fact, it turns out that in some simple cases, like the (finite-dimensional) Euclidean spaces, the compact sets are exactly the closed and bounded ones (cf. Theorem 7.169), what is not true in general metric spaces. To define compactness, we use the notion of subsequences:

Definition 7.149 (Subsequences) Let M be a nonempty set and consider two arbitrary sequences $p_n, p'_n \in M$, $n \in \mathbb{N}$. We say that $(p'_n)_{n\in\mathbb{N}}$ is a "subsequence" of the sequence $(p_n)_{n\in\mathbb{N}}$ if there is a (not necessarily unique) *strictly increasing* function $k \mapsto n_k$ such that, for all $k \in \mathbb{N}$, one has $p'_k = p_{n_k}$.
Notation: "$(p_{n_k})_{k\in\mathbb{N}}$" denotes a generic subsequence of the sequence $(p_n)_{n\in\mathbb{N}}$.

In other words, $(p'_n)_{n\in\mathbb{N}}$ is said to be a subsequence of $(p_n)_{n\in\mathbb{N}}$ if the first is the sequence obtained from the latter, by "erasing" all points p_n, whose index $n \in \mathbb{N}$ is not in the set

$$\{n_k : k \in \mathbb{N}\} \subseteq \mathbb{N} .$$

More generally, one may introduce the notion of a subnet:

Definition 7.150 (Subnets) Let M be a nonempty set and consider two arbitrary nets $(p_i)_{i\in I}$ and $(p'_j)_{j\in J}$ in M. We say that $(p'_j)_{j\in J}$ is a "subnet" of the net $(p_i)_{i\in I}$ if there is a (not necessarily unique) *strictly increasing* function $j \mapsto i_j$ from J to I such that, for all $j \in J$, one has $p'_j = p_{i_j}$ and, for all $i \in I$, there is $j_i \in J$ such that $i_{j_i} \geq i$. Notation: "$(p_{i_j})_{j\in J}$" denotes a generic subnet of the net $(p_i)_{i\in I}$.

The following simple properties of subnets are frequently used:

Exercise 7.151 Let M be any metric space.

(i) Show that any subnet of a Cauchy net in M is itself a Cauchy net.
(ii) Prove that any subnet of a net that converges in M is itself convergent and its limit is the same one of the original net.
(iii) Show that if a Cauchy net in M has a subnet that converges, then it converges itself. (In this case, by (ii), the net has the same limit as its convergent subnet.)

We are now in a position to define compact subsets of metric spaces:

Definition 7.152 (Compact Subsets) Let M be any metric space. The subset $\Omega \subseteq M$ is "compact (sequentially compact)" if any net (sequence) in Ω has a subnet (subsequence) that converges in Ω. If any net (sequence) has a net (subsequence) converging in M (but not necessarily in Ω), then Ω is, by definition, "relatively compact (relatively sequentially compact)."

In turns out that compactness and sequential compactness are equivalent properties in metric spaces. Note also that the compactness of a metric space may hold true locally, but not globally. This refers to the notion of local compactness:

Definition 7.153 (Local Compactness) We say that the metric space M is "locally compact" if, for any $p \in M$, there is $\varepsilon_p > 0$ such that the closed ball $\overline{B_{\varepsilon_p}}(p) \subseteq M$ is compact.

Clearly, any compact metric space is locally compact.

The compactness, the relative compactness, as well as the local compactness of a subset of a metric space (M, d) are topological properties, i.e., they only depend on the topology τ_d, since they refer to the convergence of subsequences, which is a topological property, as already discussed.

Exercise 7.154 Prove that the local compactness is a topological property of metric spaces.

Example 7.155 The subset $\Omega \doteq (0, 1) \subseteq \mathbb{R}$ is not compact (with respect to the usual metric in the real line): The (Cauchy) sequence $x_n \doteq 1/n \in \Omega$, $n \in \mathbb{N}$, has no subsequences converging in Ω. In fact, any subsequence of $(x_n)_{n\in\mathcal{N}}$ converges to $0 \notin \Omega$.

Note however that $(0, 1) \subseteq \mathbb{R}$ is relatively compact. This property follows from the following exercise:

Exercise 7.156 Show that $\Omega \doteq [0, 1] \subseteq \mathbb{R}$ is compact in $(\mathbb{R}, d)$, where d is the usual metric in the real line, defined by $d(x, x') \doteq |x - x'|$, $x, x' \in \mathbb{R}$.

Hint: For any fixed sequence $x_n \in \Omega$, $n \in \mathbb{N}$, consider intervals $\mathcal{I}_k \subseteq \Omega$, $k \in \mathbb{N}$, with the following properties:

(i) $\mathcal{I}_{k+1} \subseteq \mathcal{I}_k$ for all $k \in \mathbb{N}$.
(ii) $\mathcal{I}_k$, $k \in \mathbb{N}$, has length 2^{-k}.
(iii) For any $k \in \mathbb{N}$, there are *infinitely many* $n \in \mathbb{N}$ for which $x_n \in \mathcal{I}_k$.

By construction, there is a subsequence $(x_{n_k})_{k\in\mathbb{N}}$ such that, for all $k \in \mathbb{N}$, $x_{n_k} \in \mathcal{I}_k$. Note that this sequence is Cauchy.

In fact, any subset of a compact set is clearly relatively compact. Conversely, we have the following statement:

Exercise 7.157 Prove that the closure of any relatively compact subset of a metric space is compact.

There is an important relationship between the compactness and the completeness of metric spaces:

Exercise 7.158 Let (K, d) be any compact metric space. Show that this space is necessarily complete.

In what follows, we give an equivalent definition for compactness in metric spaces, which does not make any direct reference to the given metric, but only to its associated topology. Recall the compactness is a topological property.

Let (M, d) be any metric space and $\Omega \subseteq M$ an arbitrary subset. We say that the family of subsets $O_i \subseteq M$, $i \in I$, is an "open covering" of Ω, if

$$\Omega \subseteq \bigcup_{i \in I} O_i$$

and $O_i \in \tau_d$ (i.e., O_i is open) for all $i \in I$. Compactness can be completely characterized via the following property of open coverings:

Exercise 7.159 Let (M, d) be any metric space and $\Omega \subseteq M$ an arbitrary subset. Assume that, for any open covering of Ω, denoted by $O_i \in \tau_d$, $i \in I$, there is *finite* subset $J \subseteq I$ such that

$$\Omega \subseteq \bigcup_{j \in J} O_j \ .$$

Show that Ω is compact.

Hint: Use the following adaptation of the hint of the previous exercise:

1. For simplicity and without loss of generality, assume that the metric space (M, d) is complete. In fact, if (M, d) were not complete, then we could consider the completion of this metric space: By Exercise 7.100, Ω has the above open covering property, also in the completion of (M, d).
2. Show that Ω is closed.
3. Consider, for all $m \in \mathbb{N}$, the open covering of Ω given by $B_{1/m}(p)$, $p \in \Omega$.
4. Show that, for any fixed sequence $p_n \in \Omega$, $n \in \mathbb{N}$, there are subsequences $(p_{n_k}^{(m)})_{k \in \mathbb{N}}$ and points $p_m \in \Omega$ such that $p_{n_k}^{(m)} \in B_{1/m}(p_m)$, $k \in \mathbb{N}$, and, for all $m \in \mathbb{N}$, $(x_{n_k}^{(m+1)})_{k \in \mathbb{N}}$ is a subsequence of $(x_{n_k}^{(m)})_{k \in \mathbb{N}}$.
5. Take the "diagonal sequence" $(x_{n_k}^{(k)})_{k \in \mathbb{N}}$.

By the last exercise, the open covering property given above is a sufficient condition for compactness in metric spaces. We will show below, in the present subsection, that it is also a necessary condition. In fact, in the theory of general topological spaces, the property is frequently used as the *definition* compactness, whereas Definition 7.152 refers to the "sequential compactness," which is not equivalent to the open covering property, in general topological spaces.

The above open covering property is equivalent to the following property of closed subsets:

Lemma 7.160 *Let* (M, d) *be any metric space. The following properties are equivalent:*

(i) Open covering property. *For any open covering of* M*,* $O_i \in \tau_d$*,* $i \in I$*, there is a* finite *subset* $J \subseteq I$ *such that*

$$M = \bigcup_{j \in J} O_j \ .$$

(ii) Finite intersection property. *For any family* $F_i \subseteq M$*,* $i \in I$*, of* closed *subsets such that, for every* finite *subset* $J \subseteq I$*,*

$$\bigcap_{j \in J} F_j \neq \emptyset \ ,$$

one has that

$$\bigcap_{i \in I} F_i \neq \emptyset \ .$$

Proof

1. Assume that (i) is true and consider any family of closed subsets with the finite intersection property of (ii). If

$$\bigcap_{i \in I} F_i = \emptyset$$

was not true, then we would have that

$$M = \emptyset^c = (\cap_{i \in I} F_i)^c = \bigcup_{i \in I} F_i^c \ .$$

In other words, $F_i^c \in \tau_d$, $i \in I$, would be an open covering of M. Recall that the complement of any closed subset is open.

2. In this case, by (i), there is a finite $J \subseteq I$ such that

$$M = \bigcup_{j \in J} F_j^c$$

and we arrive at

$$\emptyset = (\cup_{j \in J} F_j^c)^c = \bigcap_{j \in J} F_j \ .$$

3. This contradicts the finite intersection property of the family $F_i \subseteq M$, $i \in I$. Thus, (i) implies (ii).
4. The converse implication is proven in a similar manner.

□

In various applications, the compactness is used under the form of Lemma 7.160 (ii).

In the following, we discuss important, albeit simple, properties of subsets and continuous functions in metric spaces, related to compactness.

Definition 7.161 (Bounded Subsets) Let (M, d) be any metric space. The subset $\Omega \subseteq M$ is "bounded" in (M, d) if there is some constant $C < \infty$ such that, for all $p, p' \in \Omega$, $d(p, p') \leq C$. For the sake of clearness, we sometimes call bounded subsets in normed spaces "norm-bounded."

Equivalently, $\Omega \subseteq M$ is bounded if, for all $p \in M$, there is $R_p < \infty$ such that $\Omega \subseteq B_{R_p}(p)$.

Exercise 7.162 Show that being bounded is not a topological property of subsets of metric spaces. In other words, prove that there are a set M, a subset $\Omega \subseteq M$, and two metrics, d and d', in M such that $\tau_d = \tau_{d'}$ and Ω is bounded in (M, d), but not in (M, d').

Exercise 7.163 Let (M, d) be any metric space.

(i) Show that any compact subset $K \subseteq M$ is closed and bounded. *Hint*: In order to prove that a compact subset is always closed, use Theorem 7.139 (i), as well as Exercise 7.151 (iii).
(ii) Let $K \subseteq M$ be any compact subset. If $\Omega \subseteq K$ is closed, then show that Ω is also compact. *Hint*: Use Theorem 7.139 (i).

The first part of the last exercise has the following important consequence for the Hausdorff metric:

Exercise 7.164 Let (M, d) any metric space. The set of all compact nonempty subsets of this space is denoted by $\mathcal{K}(M)$. Prove that, for all, $\Omega, \Omega' \in \mathcal{K}(M)$, one has $d_{\mathrm{H}}(\Omega, \Omega') < \infty$. In particular, by Exercise 7.104 and the second part of the last exercise, $(\mathcal{K}(M), d_{\mathrm{H}})$ is a metric space.

Moreover, it turns out that the compactness is preserved by the Hausdorff construction of metric spaces. By Exercise 7.163 (ii), note that if M is itself compact, then $\mathcal{K}(M) = \mathcal{C}(M)$, where $\mathcal{C}(M)$ denotes the set of all closed nonempty subsets of the metric space M. In particular, in this case, we have the following result:

Theorem 7.165 *If (K, d) is a compact metric space, then so does $(\mathcal{C}(K), d_{\mathrm{H}})$. In particular, $(\mathcal{C}(K), d_{\mathrm{H}})$ is complete in this case.*

The last theorem is well-known in the theory of metric spaces and will not be proven here. It is stated here only for completeness and is not important in the sequel.

Because any compact subset of a metric space is bounded, we define functions "decaying at infinity" as follows:

Definition 7.166 (Functions Decaying at Infinity) Let M be any metric space and $(X, \|\cdot\|)$ any normed space. We say that the function $f : M \to X$ "decays at infinity" if, for all $\varepsilon > 0$, there exists a *compact* subset $K_\varepsilon \subseteq M$ such that

$$\|f(p)\| < \varepsilon, \qquad \text{for all } p \notin K_\varepsilon \ .$$

The space of all functions $f : M \to X$ which are continuous and decay at infinity is denoted by $C_0(M; X)$.

In the special case of a compact domain K, observe that $C_0(K; X) = C(K; X)$, i.e., all continuous functions $f : K \to X$ are decaying at infinity, in the above sense.

Definition 7.167 (Heine-Borel Property) A metric space has the "Heine-Borel property" if every bounded closed subset is compact in this space, i.e., the converse of Exercise 7.163 (i) holds true.

In general, metric spaces do not have this property, as it is demonstrated via the following exercise:

Exercise 7.168 Let the metric space (M, d) be defined by $M \doteq \mathbb{Q}$ and $d(x, x') \doteq |x - x'|$, $x, x' \in M$. Show that this space has subsets which are closed and bounded, but that are not compact, i.e., the Heine-Borel property does not hold for (M, d).

By Exercise 7.163, note that any compact metric space has the Heine-Borel property. The following well-known theorem gives another important example of metric spaces having such a property:

Theorem 7.169 (Bolzano-Weierstrass) *Every (real or complex) normed space of finite dimension has the Heine-Borel property.*

The proof of this result will be done below, as an exercise.

Note that any metric space with the Heine-Borel property is necessarily locally compact (Definition 7.153). In particular, finite-dimensional normed spaces are locally compact, but they are not (globally) compact, being not bounded.

Another important property of compactness is that it is preserved by continuous functions:

Lemma 7.170 *Let M_1 and M_2 be two metric spaces and $f : M_1 \to M_2$ an arbitrary continuous function. For every subset $K_1 \subseteq M_1$ which is compact in M_1, the image $f(K_1) \subseteq M_2$ is compact in M_2.*

Proof Let $(q_n)_{n\in\mathbb{N}}$ be an arbitrary sequence in the image $f(K_1) \subseteq M_2$. Take any sequence $(p_n)_{n\in\mathbb{N}}$ in $K_1 \subseteq M_1$ such that $q_n = f(p_n)$, $n \in \mathbb{N}$. For K_1 is compact, there is a subsequence $(p_{n_k})_{k\in\mathbb{N}}$ converging in K_1. By Theorem 7.139 (ii), the subsequence $q_{n_k} \doteq f(p_{n_k}) \in f(K_1)$, $k \in \mathbb{N}$, converges in M_2 and

$$\lim_{k\to\infty} q_{n_k} = \lim_{k\to\infty} f(p_{n_k}) = f\left(\lim_{k\to\infty} p_{n_k}\right) \in f(K_1) .$$

□

Observe however that the converse of the last lemma does not hold, i.e., there are functions between metric spaces, which are not continuous, but preserve compactness. Such an example can be easily construct for the real line with its usual metric. The last lemma has various important consequences, as, for instance, the following corollary:

Corollary 7.171 (Weierstrass Theorem) *Let K be any compact metric space and consider the real line, $\mathbb{R}$, as being a metric space with its usual metric. Every continuous function $f : K \to \mathbb{R}$ has at least one minimizer and one maximizer in K, i.e., there are points $p_{\min}, p_{\max} \in M$ such that, for all $p \in K$,*

$$f(p_{\min}) \leq f(p) \leq f(p_{\max}) .$$

Additionally, the set of all minimizers (maximizers) of f is closed.

Proof Exercise. *Hint:* Show first that any compact subset of the real line has a minimum and a maximum, i.e., the subset contains its infimum and its supremum, respectively. □

The following proposition is a version of the above corollary for semicontinuous functions:

Proposition 7.172 *Let (M, d) be any compact metric space and $f : X \to \mathbb{R}$ an arbitrary lower-(upper-)semicontinuous functions. Let $\Omega \subseteq M$ denote the set of all minimizers (maximizers) of f. Then, Ω is a nonempty closed (and thus compact) subset of M.*

Proof Assume that f is lower semicontinuous, and take any minimizing sequence $(p_n)_{n\in\mathbb{N}}$ for f, that is,

$$\lim_{n\to\infty} f(x_n) = \inf f(M) .$$

As M is compact, we can assume that this sequence converges in M to some $p \in M$. By Proposition 7.146,

$$\lim_{n\to\infty} f(p_n) = \liminf_{n\to\infty} f(p_n) = \inf f(M) \geq f(p) .$$

Thus, $p \in \Omega$ and, hence, Ω is nonempty. Take now any sequence $(p_n)_{n\in\mathbb{N}}$ in Ω that converges to some $p \in M$. Again by Proposition 7.146, $\lim_{n\to\infty} f(p_n) \geq f(p)$. As, trivially, $f(p_n) = \inf f(M)$, one has that $p \in \Omega$. Thus, by Exercise 7.138 (ii), $\Omega \subseteq M$ is a closed subset of M. The upper semicontinuous case is proven exactly in the same way. □

Notice that functions, even when they are bounded, may not have minimizers or maximizers: Take, for instance, $\Omega = (0, 1)$ and $f(x) = x$. Then $\inf_{x\in\Omega} f(x) = 0$ and $\sup_{x\in\Omega} f(x) = 0$, but there is no $x_{\min}, x_{\max} \in \Omega$ such that $f(x_{\min}) = 0$ and $f(x_{\max}) = 1$.

The following important property of continuous functions with compact domain is also a direct consequence of Lemma 7.170:

Exercise 7.173 Let K_1 and M_2 be two metric spaces, K_1 compact, and $f : K_1 \to M_2$ a continuous one-to-one correspondence. (This implies that M_2 is also compact, by Lemma 7.170.) Show that the inverse of f is continuous. In particular, the continuous one-to-one correspondences $K_1 \to M_2$ are exactly the homeomorphisms $K_1 \to M_2$.

Hint: Recall that a function between metric spaces is continuous iff the preimage of any closed subset through this function is closed.

Exercise 7.174 Prove the Bolzano-Weierstrass theorem for $(\mathbb{R}, |\cdot|)$.

Hint: Proceed in the following way:

1. Using Exercise 7.156 and Lemma 7.170 for conveniently chosen continuous functions $f : \mathbb{R} \to \mathbb{R}$, show that, for all $L > 0$, $[-L, L] \subseteq \mathbb{R}$ is compact.
2. From Exercise 7.163 (ii), conclude that every subset $\Omega \subseteq \mathbb{R}$ which is bounded and closed is compact.

Exercise 7.175 By combining the last exercise with Theorem 7.106, prove the Bolzano-Weierstrass for $(\mathbb{R}^D, \|\cdot\|)$, $D \in \mathbb{N}$, where $\|\cdot\|$ is any norm in $\mathbb{R}^D$.

Note, again by Theorem 7.106, that the last exercise implies the (full) Bolzano-Weierstrass theorem (Theorem 7.169).

Another convenient property of compact metric spaces is the fact that the continuity of functions on such spaces is equivalent to their uniform continuity:

Proposition 7.176 *Let (M_1, d_1) and (M_2, d_2) be two metric spaces. If M_1 is compact, then any continuous mapping $f : M_1 \to M_2$ is uniformly continuous (Definition 7.107).*

Proof

1. Let $f : M_1 \to M_2$ be continuous and assume, by contradiction, that this function is not uniformly continuous. Then there is $\varepsilon > 0$ such that, for all $\tilde{\delta} > 0$, there are $p_{\tilde{\delta}}, p'_{\tilde{\delta}} \in M_1$ satisfying $d_1(p_{\tilde{\delta}}, p'_{\tilde{\delta}}) < \tilde{\delta}$ for which
$$d_2(f(p_{\tilde{\delta}}), f(p'_{\tilde{\delta}})) > \varepsilon \, . \tag{7.2}$$
2. Let $\tilde{\delta}_n \doteq 1/n$, $n \in \mathbb{N}$, and consider the sequences
$$p_n \doteq p_{\tilde{\delta}_n} \in M_1 \, , \quad p'_n \doteq p'_{\tilde{\delta}_n} \in M_1 \, ,$$

$n \in \mathbb{N}$, where $p_{\tilde{\delta}_n}$ and $p'_{\tilde{\delta}_n}$ refer to $p_{\tilde{\delta}}$ and $p'_{\tilde{\delta}}$ of the previous item, for $\tilde{\delta} = \tilde{\delta}_{n_k}$. As M_1 is compact, there is a subsequence $(p_{n_k})_{k\in\mathbb{N}}$ converging in (M_1, d_1). Note that the (sub)sequence $(p'_{n_k})_{k\in\mathbb{N}}$ also converges, by construction, and has the same limit as $(p_{n_k})_{k\in\mathbb{N}}$:

$$p \doteq \lim_{k\to\infty} p_{n_k} = \lim_{k\to\infty} p'_{n_k} .$$

See Exercise 7.151.

3. For f is continuous, Theorem 7.139 (ii) thus implies that

$$\lim_{k\to\infty} f(p_{n_k}) = \lim_{k\to\infty} f(p'_{n_k}) = f(p) .$$

In particular, one has that

$$\lim_{k\to\infty} d_2(f(p_{n_k}), f(p'_{n_k})) = 0 ,$$

which contradicts (7.2) with $\tilde{\delta} = \tilde{\delta}_{n_k}$, for a sufficiently large $k \in \mathbb{N}$. Therefore, f is necessarily uniformly continuous.

□

Now, we prove a few more properties of (sequentially) compact subsets of metric spaces, which will be eventually used to show that such subsets always have the open covering property stated in Exercise 7.159.

Definition 7.177 (Totally Bounded Subsets) Let (M, d) be any metric space. The subset $\Omega \subseteq M$ is "totally bounded" in (M, d) if, for every $\varepsilon > 0$, there is a *finite* subset $\Omega_\varepsilon \subseteq M$ such that

$$\Omega \subseteq \bigcup_{p\in\Omega_\varepsilon} B_\varepsilon(p) .$$

Note that totally bounded subsets in metric spaces are necessarily bounded, but the converse is generally not true.

Lemma 7.178 *Every compact subset of a metric space is totally bounded.*

Proof Let $K \subseteq M$ be any compact subset of the metric space (M, d).

1. Assume, by contradiction, that this subset is not totally bounded. Then, for some $\varepsilon > 0$ and all finite subsets $\Omega \subseteq K$, there is $p \in K$,

$$p \notin \bigcup_{p\in\Omega} B_\varepsilon(p) .$$

2. Hence, for some $\varepsilon > 0$, by induction, we can find a sequence $p_n \in K$, $n \in \mathbb{N}$, for which, for all $n \in \mathbb{N}$, one has

$$d(p_n, p_{n-1}), d(p_n, p_{n-2}), \ldots, d(p_n, p_2), d(p_n, p_1) \geq \varepsilon \ .$$

3. By construction, such a sequence has no convergent subsequence (in fact, it has no Cauchy subsequence). This would imply that K is not compact. Hence, K must be totally bounded.

□

Definition 7.179 (Lebesgue Numbers of an Open Covering) Let (M, d) be any metric space. Let $O_i \in \tau_d$, $i \in I$, be an open covering of a given subset $\Omega \subseteq M$. We say that the constant $\lambda > 0$ is a "Lebesgue number" for this covering if, for all $p \in \Omega$, there is $i_p \in I$ such that $B_\lambda(p) \subseteq O_{i_p}$.

Lemma 7.180 *Any open covering of a compact subset of a metric space has a Lebesgue number* $\lambda > 0$.

Proof Let $K \subseteq M$ be an arbitrary nonempty compact set in a metric space (M, d).

1. Let $O_i \in \tau_d$, $i \in I$, be any open covering of K and assume, by contradiction, that this covering has no Lebesgue number $\lambda > 0$. Then, for any $n \in \mathbb{N}$, there is $p_n \in K$ such that, for all $i \in I$, one has $B_{1/n}(p_n) \nsubseteq O_i$.
2. For Ω is compact, there is a subsequence $(p_{n_k})_{k\in\mathbb{N}}$ converging in K. Let $p \in K$ be its limit. As $O_i \in \tau_d$, $i \in I$, is a covering of K, there is $i_p \in I$ for which $p \in O_{i_p}$. For O_{i_p} is open, there is $\varepsilon > 0$ such that $B_\varepsilon(p) \subseteq O_{i_p}$.
3. As both $d(p_{n_k}, p)$ and $1/n_k$ tend to zero as $k \to \infty$, it follows, by the triangle inequality, that $B_{1/n_k}(p_{n_k}) \subseteq B_\varepsilon(p)$ for any sufficiently large $k \in \mathbb{N}$. But this contradicts the property defining the sequence $(p_n)_{n\in\mathbb{N}}$. Therefore, there exists a Lebesgue number $\lambda > 0$ for the open covering.

□

Proposition 7.181 *Let (M, d) be any metric space. Any subset $\Omega \subseteq M$ is compact iff, for every open covering of Ω, $O_i \in \tau_d$, $i \in I$, there is a finite subset $J \subseteq I$ such that*

$$\Omega \subseteq \bigcup_{j\in J} O_j \ .$$

Proof It was already proven via Exercise 7.159 that any subset of a metric space with the above open covering property is (sequentially) compact. Thus, assume that $\Omega \subseteq M$ is (sequentially) compact and take any open covering $O_i \in \tau_d$, $i \in I$, of Ω. We proved in Lemma 7.180 that such a covering always has a Lebesgue number $\lambda > 0$. For compact subsets in metric spaces are totally bounded (Lemma 7.178), there is a finite subset $\Omega_\lambda \subseteq \Omega$ for which

$$\Omega \subseteq \bigcup_{p\in\Omega_\lambda} B_\lambda(p) \ .$$

By the definition of Lebesgue numbers, for all $p \in \Omega_\lambda$, there is $i_p \in I$ such that $B_\lambda(p) \subseteq O_{i_p}$. In particular, $J \doteq \{i_p : p \in \Omega_\lambda\} \subseteq I$ is finite, and one has

$$\Omega \subseteq \bigcup_{j \in J} O_j .$$

□

Exercise 7.182 Let (M, d) be any metric space. Show that every *finite* union and any arbitrary intersection of compact subsets of this space are again compact.

7.2.7 Uniform Convergence

Let M be any metric space and X any normed space. Recall that $C(M; X)$ denotes the set of all continuous functions $M \to X$. Let $C_{\mathrm{b}}(M; X)$ be the set of all *bounded* continuous functions $M \to X$, i.e.,

$$C_{\mathrm{b}}(M; X) \doteq C(M; X) \cap \mathcal{F}_{\mathrm{b}}(M; X) ,$$

with

$$\mathcal{F}_{\mathrm{b}}(M; X) \doteq \left\{ f : M \to X \text{ such that } \|f\|_\infty \doteq \sup_{x \in M} \|f(x)\| < \infty \right\} .$$

Note that $C(M; X)$ is a vector subspace of the space $\mathcal{F}(M; X)$ of functions $M \to X$, while $C_{\mathrm{b}}(M; X)$ is a vector subspace of $\mathcal{F}_{\mathrm{b}}(M; X)$. Similarly, $C(M; \mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$, is a subalgebra of $\mathcal{F}(M; \mathbb{K})$, while $C_{\mathrm{b}}(M; \mathbb{K})$ is a subalgebra of $C(M; \mathbb{K})$. Observe also that $C_{\mathrm{b}}(M; X) = C(M; X)$, whenever M is compact.

Recall that

$$(\mathcal{F}_{\mathrm{b}}(M; X), \|\cdot\|_\infty)$$

is a Banach space, whenever X is a Banach space. It is in particular a Banach algebra. See Sect. 7.1. In what follows, we show that the same holds true for the space of bounded continuous functions

$$(C_{\mathrm{b}}(M; \mathbb{K}), \|\cdot\|_\infty).$$

For simplicity and without loss of generality, in the present subsection, we restrict ourselves to sequences and do not discuss general nets.

Definition 7.183 (Uniform Convergence of Functions) Let $(X, \|\cdot\|)$ be any normed space and Ω an arbitrary nonempty set. We say that the sequence of functions $f_n : \Omega \to X$, $n \in \mathbb{N}$, "uniformly converges" if there is a function $f : \Omega \to X$ such that

$$\lim_{n\to\infty} \|f - f_n\|_\infty \doteq \lim_{n\to\infty} \sup_{x\in M} \|f(x) - f_n(x)\| = 0 .$$

In this case, f is called the "uniform limit" of the sequence of functions.

Remark 7.184 If a sequence of functions $f_n : \Omega \to X, n \in \mathbb{N}$, uniformly converges to the function $f : \Omega \to X$, then, for every fixed $p \in \Omega$, the sequence $f_n(p) \in X$, $n \in \mathbb{N}$, converges (in the normed space $(X, \|\cdot\|)$) to $f(p) \in X$. This convergence property of a sequence of functions is called "pointwise" convergence. Thus, the uniform convergence of a sequence of functions implies its pointwise convergence. However, in general, the converse is not true: For instance, let $\Omega \doteq (0, 1)$ and $(X, \|\cdot\|) \doteq (\mathbb{R}, |\cdot|)$. The sequence $f_n : \Omega \to X, n \in \mathbb{N}$, where $f_n(x) \doteq x^n$, $x \in (0, 1)$, converges pointwise to the constant function $f(x) \doteq 0, x \in (0, 1)$, but does not converge uniformly to this function, as, for all $n \in \mathbb{N}$, one has that $\|f - f_n\|_\infty = 1$.

One crucial property of uniform convergence is that the uniform limit of continuous functions is necessarily a continuous function:

Lemma 7.185 (Uniform Limits of Continuous Functions) *Let M be any metric space and X any normed space. Let $f_n : M \to X, n \in \mathbb{N}$, be a sequence of continuous functions converging uniformly.*

(i) The uniform limit $f = \lim_{n\to\infty} f_n$ is a continuous function.
(ii) If the functions f_n are uniformly continuous, then also f is a uniformly continuous function.

Proof Exercise. □

Proposition 7.186 *If M is a metric space and $(X, \|\cdot\|)$ a* Banach *space, then $(C_b(M; X), \|\cdot\|_\infty)$ is also a Banach space.*

Proof

1. $C_b(M; X)$ is clearly a vector subspace of $C(M; X)$. In particular, $C_b(M; X)$ is a vector space. For $C_b(M; X) \subseteq \mathcal{F}_b(M; X)$, $\|\cdot\|_\infty$ defines a norm in $C_b(M; X)$. Therefore, $(C_b(M; X), \|\cdot\|_\infty)$ is a normed space.
2. In order to show that $(C_b(M; X), \|\cdot\|_\infty)$ is a complete normed space, consider an arbitrary Cauchy sequence $f_n \in C_b(M; X), n \in \mathbb{N}$. As $(X, \|\cdot\|)$ is complete, for all $p \in M$, the (Cauchy) sequence $f_n(p) \in X, n \in \mathbb{N}$, is convergent. Thus, define the function $f : M \to X$ by
$$f(p) \doteq \lim_{n\to\infty} f_n(p) , \qquad p \in M .$$
3. For $(f_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(C_b(M; X), \|\cdot\|_\infty)$, there exists a subsequence $(f_{n_k})_{k\in\mathbb{N}}$ such that
$$\|f_{n_{k+1}} - f_{n_k}\|_\infty < 2^{-k}, \quad k \in \mathbb{N} .$$

4. For all $k \in \mathbb{N}$ and $p \in M$,

$$f(p) = f_{n_k}(p) + \lim_{K\to\infty}(f_{n_K}(p) - f_{n_k}(p))$$
$$= f_{n_k}(p) + \lim_{K\to\infty}\sum_{k'=k}^{K-1}(f_{n_{k'+1}}(p) - f_{n_{k'}}(p)) .$$

Hence, by the subadditivity and continuity of the norm $\|\cdot\|$, for all $k \in \mathbb{N}$,

$$\|f - f_{n_k}\|_\infty \le \sum_{k'=k}^{\infty} 2^{-k'} = 2^{1-k}.$$

5. In particular, for any fixed $K \in \mathbb{N}$,

$$\|f\|_\infty = \|f_{n_K} + (f - f_{n_K})\|_\infty \le \|f_{n_K}\|_\infty + \|f - f_{n_K}\|_\infty < \infty ,$$

i.e., $f \in \mathcal{F}_b(M; X)$, and one has that

$$\lim_{k\to\infty}\|f - f_{n_k}\|_\infty = 0 ,$$

i.e., f is the uniform limit of the sequence $(f_n)_{n\in\mathbb{N}}$. (Recall that a Cauchy sequence with a convergent subsequence is itself convergent, with the same limit.)

6. Finally, by Lemma 7.185 (i), $f \in C_b(M; X)$, and, thus, the normed space $(C_b(M; X), \|\cdot\|_\infty)$ is complete.

□

Corollary 7.187 *If K is a* compact *metric space and X a Banach space, then $(C(K; X), \|\cdot\|_\infty)$ is a Banach space.*

Corollary 7.188 *If M is any metric space and $\mathcal{A}$ a Banach algebra, then $(C_b(M; \mathcal{A}), \|\cdot\|_\infty)$ is a Banach algebra. It is a C^*-algebra, whenever $\mathcal{A}$ is a C^*-algebra.*

Exercise 7.189 Let M be any metric space and X any normed space.

(i) Show that $C_0(M; X)$ is a vector subspace of $C_b(M; X)$.
(ii) Show that if $(X, \|\cdot\|)$ is a Banach space, then so does $(C_0(M; X), \|\cdot\|_\infty)$.
(iii) Show that $(C_0(M; \mathcal{A}), \|\cdot\|_\infty)$ is a Banach algebra, when $\mathcal{A}$ is any Banach algebra.
(iv) Show that $(C_0(M; \mathcal{A}), \|\cdot\|_\infty)$ is a (not necessarily unital) C^*-algebra, where $\mathcal{A}$ is any C^*-algebra.
(v) For M compact, show that $(C(M; \mathcal{A}), \|\cdot\|_\infty)$ is a unital C^*-algebra, when $\mathcal{A}$ is any unital C^*-algebra.

Note from the last exercise that, if the metric space M is compact, then

$$C_{\mathrm{b}}(M; X) = C(M; X) = C_0(M; X) ,$$

because, in this case,

$$C(M; X) = C_0(M; X) \subseteq C_{\mathrm{b}}(M; X) \subseteq C(M; X) .$$

By the Bolzano-Weierstrass theorem (Theorem 7.169), recall that any normed space of finite dimension has the Heine-Borel property. In particular, any closed ball $\overline{B}_R(0)$, $R \in (0, \infty)$, is compact in such a space. This property can be easily seen to be generally false in infinite dimensions, by using the normed space of continuous functions introduced above:

Example 7.190 Consider the metric space $([0, 1], d)$, where $d(x, x') \doteq |x - x'|$, $x, x' \in [0, 1]$, and the normed space $(\mathbb{R}, |\cdot|)$. Let the functions $f_n : [0, 1] \to \mathbb{R}$, $n \in \mathbb{N}$, be defined by

$$f_n(x) \doteq \cos(2\pi n \cdot x) , \qquad x \in [0, 1] .$$

The sequence $f_n \in C_{\mathrm{b}}([0, 1]; \mathbb{R})$, $n \in \mathbb{N}$, has no convergent subsequence. In fact,

$$\begin{aligned} \int_0^1 f_n(x) f_{n'}(x) \mathrm{d}x &= \int_0^1 f_n(x) f_n(x) \mathrm{d}x + \int_0^1 f_n(x)(f_{n'}(x) - f_n(x)) \mathrm{d}x \\ &= \frac{1}{2} + \int_0^1 f_n(x)(f_{n'}(x) - f_n(x)) \mathrm{d}x . \end{aligned}$$

Hence,

$$\int_0^1 f_n(x) f_{n'}(x) \mathrm{d}x \geq \frac{1}{2} - \|f_n\|_\infty \|f_n - f_{n'}\|_\infty = \frac{1}{2} - \|f_n - f_{n'}\|_\infty . \tag{7.3}$$

On the other hand,

$$\int_0^1 f_n(x) f_{n'}(x) \mathrm{d}x = 0$$

for all $n, n' \in \mathbb{N}$, $n \neq n'$. Thus, (7.3) implies that $\|f_n - f_{n'}\|_\infty \geq 1/2$, whenever $n \neq n'$. In other words, the closed ball $\overline{B}_1(0)$ is not compact in the normed (Banach) space $(C_{\mathrm{b}}([0, 1]; \mathbb{R}), \|\cdot\|_\infty)$.

To conclude this subsection, in the following, we state a criterion of algebraic nature, the Stone-Weierstrass theorem [64, Chap. V, §8], which is very useful to prove the density of subalgebras of algebras of continuous functions $(C(M; \mathbb{R}), \|\cdot\|_\infty)$.

Theorem 7.191 (Stone-Weierstrass—Metric Space Version) *Let M be any locally compact metric space and $\mathcal{A}$ a subalgebra of $C_0(M;\mathbb{R})$, where the real line $\mathbb{R}$ is (canonically) seen as a normed space, the absolute value $|\cdot|$ being its norm. Assume that the following properties are satisfied:*

(i) For all $p \in M$, there is $f \in \mathcal{A}$ such that $f(p) \neq 0$.
(ii) For all $p_1, p_2 \in M$, $p_1 \neq p_2$, there is $f \in \mathcal{A}$ such that $f(p_1) \neq f(p_2)$. (In this case, we say that $\mathcal{A}$ "separates points" in M.)

Then $\mathcal{A}$ is dense in $(C_0(M;\mathbb{R}), \|\cdot\|_\infty)$.

Let $K \subseteq \mathbb{R}$ be any compact subset. (For instance, $K = [a,b]$ with $a < b$.) Let $\mathrm{Pol}(K;\mathbb{R}) \subseteq C(K;\mathbb{R})$ be the subspace of continuous functions on K which are restrictions to K of polynomial functions $\mathbb{R} \to \mathbb{R}$, with real coefficients. Note that $\mathrm{Pol}(K;\mathbb{R}) \subseteq C(K;\mathbb{R})$ contains all constant functions and is a subalgebra of $C(K;\mathbb{R}) = C_0(K;\mathbb{R})$. Let $\mathrm{id}_K : K \to \mathbb{R}$ be the identity function $\mathrm{id}_K(x) = x$. Clearly, $\mathrm{id}_K \in \mathrm{Pol}(K;\mathbb{R})$ and

$$\mathrm{id}_K(x) \neq \mathrm{id}_K(x')$$

for all $x, x' \in K$, $x \neq x'$. Therefore, from the Stone-Weierstrass theorem (Theorem 7.191), we obtain the density of the subspace of polynomial functions:

Corollary 7.192 $\mathrm{Pol}(K;\mathbb{R})$ *is dense in* $(C(K;\mathbb{R}), \|\cdot\|_\infty)$.

One example of the use of the Stone-Weierstrass theorem for functions on a non-compact domain is given via the subspace

$$\begin{aligned}\mathcal{A}_0 \doteq \mathrm{span}\{f : f(x) = \mathrm{e}^{-cx^2}\mathcal{P}(x),\ c > 0,\ \mathcal{P} \text{ is a polynomial function}\}\\ \subseteq C_0(\mathbb{R};\mathbb{R})\ .\end{aligned}$$

Note that $\mathbb{R}$ is locally compact (by the Bolzano-Weierstrass theorem, i.e., Theorem 7.169) and that $\mathcal{A}_0$ is a subalgebra of $C_0(\mathbb{R};\mathbb{R})$ separating points of $\mathbb{R}$. For $\mathcal{A}_0$ contains nowhere-vanishing functions, the Stone-Weierstrass theorem yields the following assertion:

Corollary 7.193 $\mathcal{A}_0$ *is a dense subspace of* $(C_0(\mathbb{R};\mathbb{R}), \|\cdot\|_\infty)$.

7.3 Hilbert Spaces

7.3.1 Hilbert Spaces as Generalized Euclidean Spaces

Let $(X, \|\cdot\|)$ be any normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We say that two vectors $x, x' \in X$ are "orthogonal" to each other (notation: $x \perp x'$) if their multiples satisfy the "Pythagorean theorem," that is, for all $\alpha, \alpha' \in \mathbb{K}$,

$$\left\|\alpha x+\alpha' x'\right\|^2=\|\alpha x\|^2+\left\|\alpha' x'\right\|^2 \ .$$

In particular, for all $x \in X$, one has $0 \perp x$. Clearly, from this definition of orthogonality, for arbitrary $x, x' \in X$, if $x \perp x'$, then $\alpha x \perp \alpha x'$ for all $\alpha \in \mathbb{K}$.

Two subsets $\Omega, \Omega' \subseteq X$ are said to be "orthogonal" to each other (notation: $\Omega \perp \Omega'$) if, for all $x \in \Omega$ and all $x' \in \Omega'$, $x \perp x'$. For an arbitrary subset $\Omega \subseteq X$, $\Omega^{\perp}$ denotes the largest subset of X that is orthogonal to Ω. This subset is called the "orthogonal complement" of Ω in X. Directly from the definition of orthogonal complements, it follows that, for every subset $\Omega, \Omega' \subseteq X$, we have (a) $\Omega^{\perp} \subseteq \Omega'^{\perp}$ whenever $\Omega' \subseteq \Omega$, (b) $\Omega \subseteq \Omega^{\perp\perp} \doteq (\Omega^{\perp})^{\perp}$, and (c) $\Omega \cap \Omega^{\perp} \subseteq \{0\}$. By an argument that is similar to the one proving Lemma 7.279, note additionally that, for every nonempty $\Omega \subseteq X$, one has $(\Omega^{\perp})^{\perp\perp} = \Omega^{\perp}$.

Note that the "Pythagorean theorem" stated above implies the "parallelogram identity" for any $x, x' \in X$, $x \perp x'$:

$$\left\|x-x'\right\|^2+\left\|x+x'\right\|^2=2\|x\|^2+2\left\|x'\right\|^2 \ .$$

This motivates the following definition:

Definition 7.194 (Hilbert Space) Let $(X, \|\cdot\|)$ be a normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. We say that X is a "pre-Hilbert space" if, for all $x, x' \in X$, the norm $\|\cdot\|$ satisfies the above parallelogram identity. The pre-Hilbert space X is a "Hilbert space" or "generalized Euclidean space" if it is a Banach space.

In general, arbitrary pairs of vectors in a normed space do not satisfy the parallelogram identity. In particular, a general Banach space is not a Hilbert space. As we will see below, as a consequence of the parallelogram identity, in a pre-Hilbert space $(X, \|\cdot\|)$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, the vectors $x, x' \in X$ are orthogonal to each other iff, for all $\alpha \in \mathbb{K}$, $|\alpha| = 1$,

$$\left\|x+\alpha x'\right\|=\left\|x+x'\right\| \ .$$

Observe that this condition is only necessary for orthogonality in general normed spaces.

Simple examples of real (complex) Hilbert spaces are the usual Euclidean spaces, i.e., $\mathbb{R}^D$ ($\mathbb{C}^D$), $D \in \mathbb{N}$, equipped with the norm $\|\cdot\|_{\mathrm{e}}$ of Definition 7.32. In fact, the theory of Hilbert spaces can be seen as an abstract generalization of these simple spaces. This was, by the way, the original motivation that lead Hilbert to introduce and study these spaces. In many presentations of the theory of Hilbert spaces, the notion of scalar product is taken as primary and enters in the definition itself of the spaces. Here, we present (in the next subsection) scalar products as secondary objects, which are defined via the given norm (satisfying the parallelogram identity).

For an example of pre-Hilbert space in infinite dimension, consider the space $C([0,1]; \mathbb{K})$ of $\mathbb{K}$-valued functions on the interval $[0,1] \subseteq \mathbb{R}$, with $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For all $f \in C([0,1]; \mathbb{K})$, define

$$\|f\|_2 \doteq \sqrt{\int_0^1 |f(s)|^2 \, \mathrm{d}s} \in \mathbb{R}_0^+ .$$

$\|\cdot\|_2$ is a norm in X (called "L^2-norm"). It is easy to check that it fulfills the parallelogram identity.

Exercise 7.195 Show that the pre-Hilbert space $(C([0,1];\mathbb{R}), \|\cdot\|_2)$ is not a Hilbert space.

Exercise 7.196 Prove that the completion of a pre-Hilbert space is a Hilbert space.

The parallelogram identity and norm completeness imply the following crucial geometrical property of Hilbert spaces:

Proposition 7.197 (Distance-Minimizing Vectors) *Let $(H, \|\cdot\|)$ be any Hilbert space. For any* closed *vector subspace $G \subseteq H$ and any vector $x \in H$, there is a unique $x_G \in G$ such that*

$$\|x - x_G\| = \inf_{x' \in G} \|x - x'\| .$$

In other words, there is a unique $x_G \in G$ minimizing the distance from $x \in H$ to the subspace $G \subseteq H$.

Proof

1. Choose any sequence $x_n' \in G$, $n \in \mathbb{N}$, such that, for all $n \in \mathbb{N}$,

$$\|x - x_n'\| \leq \frac{1}{n} + \inf_{x' \in G} \|x - x'\| .$$

By the parallelogram identity, for all $m, n \in \mathbb{N}$,

$$\begin{aligned} &\left\|(x - x_n') - (x - x_m')\right\|^2 + \left\|(x - x_n') + (x - x_m')\right\|^2 \\ &\qquad = 2\left\|x - x_n'\right\|^2 + 2\left\|x - x_m'\right\|^2 . \end{aligned}$$

In particular, one has

$$\begin{aligned} \left\|x_n' - x_m'\right\|^2 \leq\ & 2\left(\inf_{x' \in G}\left\|x - x'\right\| + \tfrac{1}{m}\right)^2 + 2\left(\inf_{x' \in G}\left\|x - x'\right\| + \tfrac{1}{n}\right)^2 \\ & -4\left\|x - \tfrac{x_m' + x_n'}{2}\right\|^2 . \end{aligned}$$

2. Since $\left(x_m' + x_n'\right)/2 \in G$,

$$\left\|x - \frac{x_m' + x_n'}{2}\right\|^2 \geq \left(\inf_{x' \in G}\left\|x - x'\right\|\right)^2 .$$

Thus, combining the last two inequalities, we arrive at

$$\left\|x_n' - x_m'\right\|^2 \leq 4 \inf_{x' \in G} \left\|x - x'\right\| \left(\frac{1}{m} + \frac{1}{n}\right) + 2\left(\frac{1}{m^2} + \frac{1}{n^2}\right) .$$

Hence, any distance-minimizing sequence $x_n' \in G$, $n \in \mathbb{N}$, as above, is a Cauchy sequence.

3. As G is closed and H complete, such a sequence has a limit x_G, which lies in G. By the continuity of the norm,

$$\|x - x_G\| = \inf_{x' \in G} \left\|x - x'\right\| .$$

This proves the existence of a distance-minimizing vector $x_G \in G$, as required.

4. In order to prove uniqueness of x_G, assume that we have a second minimizing vector $x_G' \in G$ (i.e., $x_G' \in G$ also satisfies the last equality). Then, by similar arguments as above, one finds that

$$\left\|x_G' - x_G\right\|^2 = 2\left\|x - x_G'\right\|^2 + 2\|x - x_G\|^2 - 4\left\|x - \frac{x_G' + x_G}{2}\right\|^2 \leq 0 .$$

□

In the following subsection, we show that norms satisfying the parallelogram identity are exactly those that are defined from scalar products.

7.3.2 Scalar Products

As already mentioned, Hilbert spaces H are often defined via a particular type of sesquilinear form $H \times H \to \mathbb{C}$ called "scalar product," instead of being defined as normed spaces with the parallelogram identity, as done in Definition 7.194. The precise definition of scalar products is as follows:

Definition 7.198 (Scalar Product) Let V be any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For $\mathbb{K} = \mathbb{R}$ ($\mathbb{K} = \mathbb{C}$), we say that $\langle \cdot, \cdot \rangle : V \times V \to \mathcal{K}$ is a "scalar semiproduct" in V if it has the following properties:

(i) *Positivity.* $\langle v, v \rangle \geq 0$ for any $v \in V$.
(ii) *Symmetry (Hermiticity in the complex case).* $\langle v, v' \rangle = \langle v', v \rangle$ ($\langle v, v' \rangle = \overline{\langle v', v \rangle}$ when $\mathbb{K} = \mathbb{C}$) for any $v, v' \in V$.
(iii) *Additivity in the second argument.* $\langle v, v' + v'' \rangle = \langle v, v' \rangle + \langle v, v' \rangle$ for any $v, v', v'' \in V$.
(iv) *Homogeneity in the second argument.* $\langle v, \alpha v' \rangle = \alpha \langle v, v' \rangle$ for any $v, v' \in V$ and $\alpha \in \mathbb{K}$.

The scalar semiproduct $\langle\cdot,\cdot\rangle$ is said to be a "scalar product" if it is non-degenerate, i.e., for all $v \in V$, $\langle v, v\rangle = 0$ iff $v = 0$.

The following fact about scalar products is frequently useful in proofs:

Lemma 7.199 *Let V be any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ a scalar* product. *For any two fixed mappings $\Phi, \Phi' : V \to V$, $\Phi = \Phi'$ iff*

$$\langle x, \Phi(y)\rangle = \left\langle x, \Phi'(y)\right\rangle , \qquad x, y \in V .$$

Proof Exercise. □

Scalar products in real (complex) vector spaces are special cases of bilinear (sesquilinear) forms:

Definition 7.200 (Sesquilinear and Bilinear Forms) Let V be any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.

(i) $[\cdot,\cdot] : V \times V \to \mathbb{K}$ is a "bilinear form" if it has the following properties:

- (i.a) *Additivity in the first argument.* $\left[v' + v'', v'\right] = \left[v', v\right] + \left[v'', v\right]$ for all $v, v', v'' \in V$.
- (i.b) *Additivity in the second argument.* $\left[v, v' + v''\right] = \left[v, v'\right] + \left[v, v''\right]$ for all $v, v', v'' \in V$.
- (i.c) *Homogeneity in both arguments.* $\left[\alpha v, v'\right] = \left[v, \alpha v'\right] = \alpha \left[v, v'\right]$ for all $v, v' \in V$ and $\alpha \in \mathcal{K}$.

(ii) In the *complex* case, we say that $[\cdot,\cdot] : V \times V \to \mathbb{C}$ is a "sesquilinear form" if it has the following properties:

- (ii.a) *Additivity in the first argument.* $\left[v' + v'', v\right] = \left[v', v\right] + \left[v'', v\right]$ for all $v, v', v'' \in V$.
- (ii.b) *Additivity in the second argument.* $\left[v, v' + v''\right] = \left[v, v'\right] + \left[v, v''\right]$ for all $v, v', v'' \in V$.
- (ii.c) *(Anti)Homogeneity in the (first) second argument.* $\left[\overline{\alpha} v, v'\right] = \left[v, \alpha v'\right] = \alpha \left[v, v'\right]$ for all $v, v' \in V$ and $\alpha \in \mathbb{C}$.

(iii) The bilinear form $[\cdot,\cdot]$ is said to be "symmetric" if $\left[v, v'\right] = \left[v', v\right]$ for all $v, v' \in V$.

(iv) The sesquilinear form $[\cdot,\cdot]$ is "Hermitian" if $\left[v, v'\right] = \overline{[v', v]}$ for all $v, v' \in V$.

Note that any real scalar semiproduct is a symmetric bilinear form, whereas complex scalar semiproducts are Hermitian sesquilinear forms.

Scalar products can be used to define seminorms on vector spaces as follows:

Definition 7.201 (Seminorm Associated with a Scalar Semiproduct) Given a vector space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and a scalar semiproduct $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$, we define the associated seminorm to be

$$\|v\|_{\langle\cdot,\cdot\rangle} \doteq \sqrt{\langle v, v\rangle} , \qquad v \in V .$$

It is a priori not clear that we properly defined above a seminorm, in the sense of Definition 7.31. In fact, this is a direct consequence of the celebrated Cauchy-Schwarz inequality:

Proposition 7.202 (Cauchy-Schwarz Inequality) *Let V be a vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$, with a scalar semiproduct denoted by $\langle \cdot, \cdot \rangle$. Then, for all $v, v' \in V$,*

$$\left|\left\langle v, v' \right\rangle\right| \leq \|v\|_{\langle \cdot, \cdot \rangle} \left\|v'\right\|_{\langle \cdot, \cdot \rangle} .$$

Proof

1. Fix arbitrary vectors $v, v' \in V$. For all $\alpha \in \mathbb{R}$,

$$0 \leq \left\langle v + \alpha v', v + \alpha v' \right\rangle = \alpha^2 \left\|v'\right\|^2_{\langle \cdot, \cdot \rangle} + 2\alpha \mathrm{Re}\{\left\langle v, v' \right\rangle\} + \|v\|^2_{\langle \cdot, \cdot \rangle} .$$

2. If $\left\|v'\right\|_{\langle \cdot, \cdot \rangle} = 0$, by taking the limits $\alpha \to \pm\infty$, we conclude that

$$\pm \mathrm{Re}\{\left\langle v, v' \right\rangle\} \geq 0 .$$

   Hence, in this case, $\mathrm{Re}\{\left\langle v, v' \right\rangle\} = 0$. In particular,

$$\left|\mathrm{Re}\{\left\langle v, v' \right\rangle\}\right| \leq \|v\|_{\langle \cdot, \cdot \rangle} \left\|v'\right\|_{\langle \cdot, \cdot \rangle} .$$

3. Assume now that $\left\|v'\right\|_{\langle \cdot, \cdot \rangle} > 0$. By positivity (see Point 1), the polynomial equation

$$\left\|v'\right\|^2_{\langle \cdot, \cdot \rangle} \alpha^2 + 2\mathrm{Re}\{\left\langle v, v' \right\rangle\}\alpha + \|v\|^2_{\langle \cdot, \cdot \rangle} = 0$$

   has at most one solution in the real line. Thus,

$$4 \left(\mathrm{Re}\{\left\langle v, v' \right\rangle\}\right)^2 - 4 \|v\|^2_{\langle \cdot, \cdot \rangle} \left\|v'\right\|^2_{\langle \cdot, \cdot \rangle} \leq 0 .$$

   Hence, for all $v, v' \in X$, we have again that

$$\left|\mathrm{Re}\{\left\langle v, v' \right\rangle\}\right| \leq \|v\|_{\langle \cdot, \cdot \rangle} \left\|v'\right\|_{\langle \cdot, \cdot \rangle}$$

4. If V is a real vector space, then $\left\langle v, v' \right\rangle \in \mathbb{R}$, and the last inequality is nothing else than the (Cauchy-Schwarz) inequality stated in the proposition.
5. If V is a complex vector space, for any fixed $v, v' \in V$, we can choose a constant $c \in \mathbb{C}$, $|c| = 1$, such that

$$\left\langle v, cv' \right\rangle = c \left\langle v, v' \right\rangle = \left|\left\langle v, v' \right\rangle\right| . \tag{7.4}$$

Thus, by the last inequality,

$$\begin{aligned}\left|\left\langle v, v'\right\rangle\right| &= \left|\mathrm{Re}\{c\left\langle v, v'\right\rangle\}\right| = \left|\mathrm{Re}\{\left\langle v, cv'\right\rangle\}\right| \\ &\leq \|v\|_{\langle\cdot,\cdot\rangle}\left\|cv'\right\|_{\langle\cdot,\cdot\rangle} = \|v\|_{\langle\cdot,\cdot\rangle}\left\|v'\right\|_{\langle\cdot,\cdot\rangle} \ . \end{aligned} \tag{7.5}$$

□

One important consequence of the Cauchy-Schwarz inequality is the fact that $\|\cdot\|_{\langle\cdot,\cdot\rangle}$ is a seminorm in V:

Lemma 7.203 *Let $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ be any scalar semiproduct in the (real or complex) vector space V. $\|\cdot\|_{\langle\cdot,\cdot\rangle} : V \to \mathbb{R}_0^+$ is a seminorm in V. If $\langle\cdot,\cdot\rangle$ is a scalar product, then $\|\cdot\|_{\langle\cdot,\cdot\rangle}$ is a norm.*

Proof Exercise. □

Seminorms associated with scalar semiproducts always satisfy the parallelogram identity: Let V be a vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. In the real (complex) case $\mathbb{K} = \mathbb{R}$ ($\mathbb{K} = \mathbb{C}$), let $[\cdot,\cdot] : V \times V \to \mathbb{K}$ be any bilinear (sesquilinear) form. Then, for all $v, v' \in V$,

$$\left[v + v', v + v'\right] + \left[v - v', v - v'\right] = 2\left([v, v] + \left[v', v'\right]\right) \ .$$

In particular, if $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ is a scalar semiproduct in the vector space V, then, for all $v, v' \in V$,

$$\left\|v + v'\right\|^2_{\langle\cdot,\cdot\rangle} + \left\|v - v'\right\|^2_{\langle\cdot,\cdot\rangle} = 2\left(\|v\|^2_{\langle\cdot,\cdot\rangle} + \left\|v'\right\|^2_{\langle\cdot,\cdot\rangle}\right) \ .$$

It turns out that any norm satisfying the parallelogram identity is a norm associated with a unique scalar product:

Theorem 7.204 (Jordan and von Neumann) *Let V be any vector space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\|\cdot\|$ a norm in V. If this norm satisfies the parallelogram identity, i.e.,*

$$\left\|v - v'\right\|^2 + \left\|v + v'\right\|^2 = 2\|v\|^2 + 2\left\|v'\right\|^2 \ , \qquad v, v' \in V \ ,$$

then there is a unique scalar product $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ such that $\|\cdot\| = \|\cdot\|_{\langle\cdot,\cdot\rangle}$. In this case, if $\mathbb{K} = \mathbb{R}$, then

$$\left\langle v, v'\right\rangle = \frac{1}{4}\sum_{n=1}^{2}(-1)^n\left\|v + (-1)^n v'\right\|^2 \ , \qquad v, v' \in V \ ,$$

while if $\mathbb{K} = \mathbb{C}$, then

$$\left\langle v, v'\right\rangle = \frac{1}{4}\sum_{n=1}^{4}(-i)^n\left\|v + i^n v'\right\|^2 \ , \qquad v, v' \in V \ .$$

The above expression for the scalar product as a function of the corresponding norms is known under the name (real or complex) "polarization identity." Because of the last theorem, any pre-Hilbert space is canonically endowed with a scalar product, the one associated uniquely with its norm. See Definition 7.194.

Consider again the pre-Hilbert space $(C([0,1];\mathbb{K})), \|\cdot\|_2)$, given as an example after Definition 7.194. From the polarization equality, in the real case $\mathbb{K}=\mathbb{R}$,

$$\langle f, f'\rangle = \int_0^1 f(s)f'(s)\mathrm{d}s \ , \qquad f, f' \in C([0,1];\mathbb{R}) \ ,$$

while in the complex case $\mathbb{K}=\mathbb{C}$,

$$\langle f, f'\rangle = \int_0^1 \overline{f(s)}f'(s)\mathrm{d}s \ , \qquad f, f' \in C([0,1];\mathbb{C}) \ .$$

Recall that the space of bounded (linear) operators acting on an pre-Hilbert space H is denoted by $\mathcal{B}(H)$ with norm equal to

$$\|A\|_{\mathrm{op}} \doteq \sup\{\|A\varphi\| : \varphi \in H,\ \|\varphi\| = 1\} \ , \qquad A \in \mathcal{B}(H) \ .$$

Using again the Cauchy-Schwarz inequality, this norm can be rewritten in a way which is frequently used in proofs:

Lemma 7.205 *Let $(H, \|\cdot\|)$ be any (real or complex) pre-Hilbert space with its (canonically) associated scalar product denoted by $\langle\cdot,\cdot\rangle$. For any $A \in \mathcal{B}(H)$,*

$$\|A\|_{\mathrm{op}} = \sup_{\substack{x,x'\in X\\ \|x\|=\|x'\|=1}} |\langle x, A(x')\rangle| = \sup_{\substack{x,x'\in X\\ \|x\|=\|x'\|=1}} |\langle A(x), x'\rangle| \ .$$

Proof Exercise. □

Because of the Jordan-von Neumann theorem (Theorem 7.204) and the polarization identity, orthogonality can be characterized in terms of vanishing scalar products:

Lemma 7.206 (Orthogonality and Vanishing Scalar Products) *Let H be any Hilbert space with its associated scalar product denoted by $\langle\cdot,\cdot\rangle$.*

(i) Two vectors $x, x' \in H$ are orthogonal iff $\langle x, x'\rangle = 0$.
(ii) For any nonempty $\Omega \subseteq H$, $\Omega^{\perp} \subseteq H$ is a closed subspace.
(iii) For any nonempty $\Omega \subseteq H$,

$$\left(\overline{\mathrm{span}(\Omega)}\right)^{\perp} = \Omega^{\perp}.$$

Proof Exercise. □

Notice that part (ii) of the above lemma is analogous to Exercise 7.296, which is a result on bands of Riesz spaces.

Using the last lemma, we may characterize the distance-minimizing vectors of Proposition 7.197 via an orthogonality condition:

Proposition 7.207 (Distance-Minimizing Property as Orthogonality) *Let* $(H, \|\cdot\|)$ *be any pre-Hilbert space,* $G \subseteq H$ *an arbitrary vector subspace, and* $x \in H$*. For every* $x' \in G$*,*

$$\left\|x - x'\right\| = \inf_{x'' \in G} \left\|x - x''\right\|$$

iff $\left\{x - x'\right\} \perp G$*. In other words,* $x' \in G$ *minimizes the distance from* G *to* x *iff the difference* $x - x'$ *is orthogonal to any vector in the subspace* $G \subseteq H$*.*

Proof Let $G \subseteq H$ be any subspace and fix two vectors $x \in H$ and $x' \in G$.

1. Assume that $\left\{x - x'\right\} \perp G$. Then, for any $x'' \in G$,

$$\left\|x - (x' + x'')\right\|^2 = \left\|x - x'\right\|^2 + \left\|x''\right\|^2 \ .$$

Thus, the distance from x to $x' + x'' \in G$ reaches a minimum at $x'' = 0$. As any vector of G has the form $x' + x''$ for some $x'' \in G$, we conclude that

$$\left\|x - x'\right\| = \inf_{x'' \in G} \left\|x - x''\right\| \ .$$

2. Suppose now that the above equality is satisfied for $x' \in G$. Then, for any $x'' \in G$, the function $\mathbb{R} \to \mathbb{R}$, $t \mapsto \left\|x - x' - tx''\right\|^2$ takes its minimum at $t = 0$. Observe that this function is differentiable and its derivative is

$$2(t \left\|x''\right\|^2 - \mathrm{Re}\left\langle x - x', x''\right\rangle)$$

for any $t \in \mathbb{R}$. As the derivatives vanish at critical points (in particular at minima), we have that

$$\mathrm{Re}\left\langle x - x', x''\right\rangle = 0$$

for all $x'' \in G$. Finally, as G is a vector (sub)space, it follows that, for all $x'' \in G$,

$$\left\langle x - x', x''\right\rangle = 0 \ .$$

(In the complex case, use similar arguments as those done in Eqs. (7.4)–(7.5).)

□

Notice that, among various other important things, the last proposition is behind the famous “least squares method” for finding the best fit for a set of data.

Corollary 7.208 *Let H be any Hilbert space. If $\Omega \subseteq H$ is a vector* subspace*, then* $\overline{\Omega} = \Omega^{\perp\perp}$*. In particular,* $\Omega = \Omega^{\perp\perp}$*, whenever Ω is closed.* *

Proof From Lemma 7.206, one has that $\overline{\Omega} \subseteq \overline{\Omega}^{\perp\perp} = \Omega^{\perp\perp}$. Thus, suppose, by contradiction, that there is $x \in \Omega^{\perp\perp} \backslash \overline{\Omega}$, that is, $\Omega^{\perp\perp} \neq \overline{\Omega}$. Let $x_{\overline{\Omega}} \in \overline{\Omega}$ be the unique vector of the closed vector subspace $\overline{\Omega} \subseteq H$ (see Exercise 7.140) minimizing the distance $\|x - \cdot\|$. See Proposition 7.197. Clearly, $\left\|x - x_{\overline{\Omega}}\right\| > 0$, because x would be otherwise a vector of (the closed subset) $\overline{\Omega}$. As $\overline{\Omega} \subseteq \Omega^{\perp\perp}$ and $\Omega^{\perp\perp}$ is a vector subspace (Lemma 7.206), $x - x_{\overline{\Omega}} \in \Omega^{\perp\perp}$. But, by Proposition 7.207, $x - x_{\overline{\Omega}} \in \overline{\Omega}^{\perp} = \Omega^{\perp}$. In other words, $x - x_{\overline{\Omega}} \in \Omega^{\perp} \cap \Omega^{\perp\perp} = \{0\}$. Thus, we arrive at $x - x_{\overline{\Omega}} = 0$, which contradicts $\left\|x - x_{\overline{\Omega}}\right\| > 0$. □

Notice that the above corollary is analogous to Proposition 7.298, which is a result on bands of Riesz spaces.

We discuss now the equivalence of Hilbert spaces, which corresponds to the following definition:

Definition 7.209 (Unitary Transformation) Let H_1 and H_2 be two Hilbert spaces. A "unitary transformation" $U \in \mathcal{B}(H_1; H_2)$ is a one-to-one correspondence which is isometric, i.e., $U : H_1 \to H_2$ is an isomorphism of normed spaces. The Hilbert spaces are "equivalent" or "unitarily equivalent," whenever there exists a unitary transformation $H_1 \to H_2$, i.e., when they are equivalent as normed spaces.

Because of the polarization identity, unitary transformations "preserve angles," i.e., they preserve scalar products:

Lemma 7.210 *Let $(H_1, \|\cdot\|^{(1)})$ and $(H_2, \|\cdot\|^{(2)})$ be two equivalent Hilbert spaces with the respective associated scalar products denoted by $\langle\cdot,\cdot\rangle^{(1)}$ and $\langle\cdot,\cdot\rangle^{(2)}$. Then, for any unitary transformation $U : H_1 \to H_2$,*

$$\left\langle U(x_1), U(x_1')\right\rangle^{(2)} = \left\langle x_1, x_1'\right\rangle^{(1)}, \qquad x_1, x_1' \in H_1 .$$

Conversely, any surjective linear transformation $U : H_1 \to H_2$ satisfying the above equality is a unitary transformation.

Proof Exercise. □

7.3.3 Orthogonal Decompositions

We discuss now the orthogonal decompositions of Hilbert spaces, which requires the notion of (Hilbert) direct sums of Hilbert (sub)spaces:

Definition 7.211 (Direct Sums in a Hilbert Space) A Hilbert space H is the "(Hilbert) direct sum"

$$H \equiv H_1 \oplus_2 \cdots \oplus_2 H_N$$

of $N \in \mathbb{N}$ vector subspaces $H_1, \ldots, H_N \subseteq H$ if the following conditions are satisfied:

(i) *Mutual orthogonality*. For all $m, n \in \{1, \ldots, N\}$ with $m \neq n$, $H_m \perp H_n$.
(ii) *Algebraic direct sum*. For all $x \in H$, there are unique vectors $x_1 \in H_1, \ldots, x_N \in H_N$ such that

$$x = x_1 + \cdots + x_N \ .$$

Note that direct sums in arbitrary vector spaces (i.e., in vector spaces that are not necessarily Hilbert spaces) are defined by property (ii) of the above definition. Recall that, in this case, the usual notation for the direct sum is $H_1 \oplus \cdots \oplus H_N$. We make use of the symbol "$\oplus_2$" to refer to property (i), i.e., the mutual orthogonality of the subspaces $H_1, \ldots, H_N$, which only makes sense in a Hilbert space.

Proposition 7.212 (Orthogonal Decomposition Theorem) *Let H be a Hilbert space. Then, for any* closed *vector subspace $G \subseteq H$, $H = G \oplus_2 G^{\perp}$.*

Proof Fix any closed vector subspace $G \subseteq H$. Clearly, by the mere definition of the orthogonal complement $G^{\perp}$, $G \perp G^{\perp}$. Take any $x \in H$ and let x_G be the (unique) vector in G minimizing the distance $\|\cdot - x\|$. Note that x_G exists, by Proposition 7.197. Define further $x_{G^{\perp}} \doteq x - x_G$. By Proposition 7.207, $x_{G^{\perp}} \in G^{\perp}$. In particular, $x = x' + x''$ for some $x' \in G$ and $x'' \in G^{\perp}$. The uniqueness of this decomposition of x follows from $G \cap G^{\perp} = \{0\}$. □

Notice that the last proposition is completely analogous to Proposition 7.300, which is a result on the disjoint decomposition of (order-)complete Riesz spaces by their bands.

The above proposition has various important consequences. The first one to be discussed is the existence of so-called orthogonal projectors on arbitrary closed subspaces of any Hilbert space H: Let $G \subseteq H$ be any closed subspace and, for any $x \in H$, let $x = x' + x''$ be the unique decomposition of x for which $x' \in G$ and $x'' \in G^{\perp}$. The condition

$$P_G(x) = x' \ , \qquad x \in H \ ,$$

uniquely defines a linear transformation $P_G \in \mathcal{L}(H)$. Such mappings have the following properties:

Lemma 7.213 *Let $G \subseteq H$ be any closed subspace of a Hilbert space H.*

(i) $P_G \in \mathcal{L}(H)$ *is a projector, i.e.,* $P_G \circ P_G = P_G$.
(ii) $P_G \in \mathcal{B}(H)$ *and* $\|P_G\|_{\mathrm{op}} \leq 1$.
(iii) $\|P_G\|_{\mathrm{op}} = 0$ *iff* $G = \{0\}$. *Otherwise,* $\|P_G\|_{\mathrm{op}} = 1$.
(iv) $P_G + P_{G^{\perp}} = \mathrm{id}_H$.

Proof Exercise. □

A second important consequence of the orthogonal decomposition theorem is the "Riesz-Fréchet representation theorem" proven below. Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $x \in H$ an arbitrary fixed vector. Define the linear functional $\varphi_x : H \to \mathbb{K}$ by

$$\varphi_x(x') \doteq \langle x, x' \rangle , \qquad x' \in H .$$

From the Cauchy-Schwarz inequality, one has

$$|\varphi_x(x')| \leq \|x\| \, \|x'\| , \qquad x' \in H .$$

Thus, $\|\varphi_x\|_{\mathrm{op}} \leq \|x\| < \infty$ and, hence, $\varphi_x \in H^{\mathrm{td}} \doteq \mathcal{B}(H; \mathbb{K})$. We show next that any element of the (topological) dual space H^{td} is of this form:

Theorem 7.214 (Riesz-Fréchet) *Let H be a Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For all $\varphi \in H^{\mathrm{td}}$, there is a unique $x_\varphi \in H$ such that $\varphi = \varphi_{x_\varphi}$. Additionally, for any $\varphi \in H^{\mathrm{td}}$, $\|x_\varphi\| = \|\varphi\|_{\mathrm{op}}$, and, for all $\varphi, \varphi' \in H^{\mathrm{td}}$, $\alpha \in \mathbb{K}$,*

$$x_{\varphi+\varphi'} = x_\varphi + x_{\varphi'} \qquad \textit{and} \qquad x_{\alpha\varphi} = \bar{\alpha} x_\varphi .$$

In other words, the mapping $\varphi \mapsto x_\varphi$ from H^{td} to H is "antilinear."

Proof

1. Fix any $\varphi \in H^{\mathrm{td}}$ and let $x_\varphi, x'_\varphi \in H$ be two vectors for which

$$\langle x_\varphi, x \rangle = \langle x'_\varphi, x \rangle = \varphi(x)$$

   for all $x \in H$. In particular,

$$\langle x_\varphi, x_\varphi - x'_\varphi \rangle = \langle x'_\varphi, x_\varphi - x'_\varphi \rangle .$$

   Hence,

$$\langle x_\varphi - x'_\varphi, x_\varphi - x'_\varphi \rangle = 0 ,$$

   from which we conclude that $x_\varphi - x'_\varphi = 0$. This proves the uniqueness of x_φ.
2. Let $\varphi, \varphi' \in H^{\mathrm{td}}$ be, respectively, represented by the vectors $x_\varphi, x_{\varphi'} \in H$. Thus, for all $x \in H$, one has

$$(\varphi + \varphi')(x) = \langle x_\varphi + x_{\varphi'}, x \rangle .$$

   By the uniqueness of the vectors representing continuous linear functionals (see Point 1), we arrive at

$$x_{\varphi+\varphi'} = x_\varphi + x_{\varphi'} .$$

3. In a similar way, we show that, for all $\varphi \in H^{\mathrm{td}}$ and $\alpha \in \mathbb{K}$,

$$x_{\alpha\varphi} = \overline{\alpha} x_\varphi \ .$$

4. If $x_\varphi \in H$ represents $\varphi \in H^{\mathrm{td}}$, then, by the Cauchy-Schwarz inequality (Proposition 7.202),

$$\|\varphi\|_{\mathrm{op}} = \sup_{x \in H,\ \|x\|=1} \left|\langle x_\varphi, x\rangle\right| \leq \left\|x_\varphi\right\| \ .$$

Moreover, if $\varphi \neq 0$, then $x_\varphi \neq 0$ and one has

$$\left\|x_\varphi\right\| = \frac{\langle x_\varphi, x_\varphi\rangle}{\left\|x_\varphi\right\|} = \left\langle x_\varphi, \frac{x_\varphi}{\left\|x_\varphi\right\|}\right\rangle \leq \sup_{x \in H,\ \|x\|=1} \left|\langle x_\varphi, x\rangle\right| = \|\varphi\|_{\mathrm{op}} \ .$$

Hence, $\left\|x_\varphi\right\| = \|\varphi\|_{\mathrm{op}}$ for any continuous linear functional $\varphi \in H^{\mathrm{td}}$ that can be represented by a vector $x_\varphi \in H$.

5. Now we prove the existence of the representing vector $x_\varphi \in H$ for any $\varphi \in H^{\mathrm{td}}$. If $\varphi = 0$, then, clearly, $x_\varphi = 0$. Suppose thus that $\varphi \neq 0$. By linearity and continuity of φ, the kernel

$$\ker(\varphi) \doteq \{x \in H : \varphi(x) = 0\} \subseteq H$$

is a closed vector subspace.

6. $\ker(\varphi)^\perp \subseteq H$ has dimension at most equal 1: Observe that, for any $x \in \ker(\varphi)^\perp$ with $x \neq 0$ (if it exists at all), one has that $\varphi(x) \neq 0$, because $\ker(\varphi) \cap \ker(\varphi)^\perp = \{0\}$. Let $x, x' \in \ker(\varphi)^\perp$ with $x, x' \neq 0$. By construction,

$$\frac{1}{\varphi(x)} x - \frac{1}{\varphi(x')} x' \in \ker(\varphi) \cap \ker(\varphi)^\perp = \{0\} \ .$$

Hence, x and x' are linearly dependent.

7. Take any $y \in H$ such that $\varphi(y) = 1$. Such a vector must exist, as φ is linear and $\varphi \neq 0$, by assumption. Define

$$y_\varphi \doteq P_{\ker(\varphi)^\perp}(y) \ .$$

As $y - y_\varphi \in \ker(\varphi)$ (see Lemma 7.213 (iv)), $\varphi(y_\varphi) = 1$. In particular, $y_\varphi \neq 0$ and, hence, $\dim \ker(\varphi)^\perp = 1$, provided that $\varphi \neq 0$.

8. Finally, let $x_\varphi \doteq \left\|y_\varphi\right\|^{-2} y_\varphi \in \ker(\varphi)^\perp$. For all $x \in H$,

$$\begin{aligned}
\langle x_\varphi, x\rangle &= \langle x_\varphi, P_{\ker(\varphi)^\perp}(x) + P_{\ker(\varphi)}(x)\rangle \\
&= \langle x_\varphi, P_{\ker(\varphi)^\perp}(x)\rangle + 0 \\
&= \langle x_\varphi, P_{\ker(\varphi)^\perp}(x)\rangle + \varphi(P_{\ker(\varphi)}(x)) \ .
\end{aligned}$$

By the last equality, one has only to prove that

$$\varphi(x) = \langle x_\varphi, x \rangle$$

for all $x \in \ker(\varphi)^\perp$. Recalling that $\dim \ker(\varphi)^\perp = 1$, by linearity with respect to x in both sides of the above equation, this follows from

$$\varphi(y_\varphi) = 1 = \frac{\langle y_\varphi, y_\varphi \rangle}{\|y_\varphi\|^2} = \langle x_\varphi, y_\varphi \rangle \ .$$

□

From the above proposition, any Hilbert space H and its (topological) dual space H^{td} can be canonically identified as sets (but not as vector spaces, in the complex case). Such an identification is isometric (i.e., norm preserving) and antilinear. In particular, in the real case, H and H^{td} are canonically isomorphic normed spaces.

Exercise 7.215 Let H be any (not necessarily real) Hilbert space. Show that $(H^{\mathrm{td}}, \|\cdot\|_{\mathrm{op}})$ is again a Hilbert space and derive an expression for its scalar product in terms of the scalar product of H.

Another important consequence of the Riesz-Fréchet representation theorem is the Riesz theorem on the representation of bounded sesquilinear forms:

Corollary 7.216 (Riesz) *Let $(H, \|\cdot\|)$ be a Hilbert space over $\mathbb{K} = \mathbb{R}$ ($\mathbb{K} = \mathbb{C}$). Take a bilinear (sesquilinear) form $[\cdot, \cdot] : H \times H \to \mathbb{K}$ and assume its boundedness, i.e.,*

$$|[x, x']| \leq C \|x\| \|x'\| \ , \qquad x, x' \in H \ ,$$

for some constant $C < \infty$. Then, there exists a unique $A_{[\cdot,\cdot]} \in \mathcal{B}(H)$ such that, for all $x, x' \in H$,

$$[x, x'] = \langle A_{[\cdot,\cdot]}(x), x' \rangle \ .$$

Additionally, $\|A_{[\cdot,\cdot]}\|_{\mathrm{op}} \leq C$.

Proof We only prove the corollary for the complex case $\mathbb{K} = \mathbb{C}$, the real one being even simpler.

1. Clearly, $A_{[\cdot,\cdot]}$ has to be unique, if it exists. Recall Lemma 7.205: For any $A \in \mathcal{B}(H)$,

$$\|A\|_{\mathrm{op}} = \sup_{\substack{x, x' \in H \\ \|x\| = \|x'\| = 1}} |\langle A(x), x' \rangle| \ .$$

Thus, $\|A_{[\cdot,\cdot]}\|_{\mathrm{op}} \leq C$, if the operator $A_{[\cdot,\cdot]}$ exists.

2. We now show the existence of an operator $A_{[\cdot,\cdot]} \in \mathcal{B}(H)$ representing the bounded sesquilinear form $[\cdot,\cdot]$. For all $x \in H$, define the linear functional $\varphi_x \in H'$ by

$$\varphi_x(x') \doteq [x, x'] , \qquad x' \in H .$$

As the sesquilinear form $[\cdot,\cdot]$ is bounded, $\varphi_x \in H^{\mathrm{td}}$. Thus, by the Riesz-Fréchet theorem (Theorem 7.214), for all $x \in H$, there exists a unique $\psi_x \in H$ such that, for all $x' \in H$,

$$[x, x'] = \langle \psi_x, x' \rangle .$$

3. By uniqueness of ψ_x and antilinearity of $[\cdot,\cdot]$ and $\langle\cdot,\cdot\rangle$ with respect to their left arguments, we conclude that the mapping $A_{[\cdot,\cdot]} : H \to H$ defined by

$$A_{[\cdot,\cdot]}(x) = \psi_x , \qquad x \in H ,$$

is linear. By construction, for all $x, x' \in H$,

$$\langle A_{[\cdot,\cdot]}(x), x' \rangle = [x, x'] .$$

4. Because $[\cdot,\cdot]$ is a bounded sesquilinear form, for some $C < \infty$ and all $x \in H$,

$$\begin{aligned} \left\| A_{[\cdot,\cdot]}(x) \right\|^2 &= \langle A_{[\cdot,\cdot]}(x), A_{[\cdot,\cdot]}(x) \rangle \\ &= [x, A_{[\cdot,\cdot]}(x)] \leq C \left\| A_{[\cdot,\cdot]}(x) \right\| \|x\| . \end{aligned}$$

In other words, for all $x \in H$,

$$\left\| A_{[\cdot,\cdot]}(x) \right\| \leq C \|x\|$$

and $A_{[\cdot,\cdot]}$ is thus a bounded operator.

□

7.3.4 *Hilbert Bases*

In finite dimension, the notion of bases of vector spaces is well-known, appearing in any elementary course on linear algebra. In infinite dimension, one has to be a little bit more cautious. As in the finite-dimensional case, in any vector space, there is always a Hamel basis, that is, a set of linearly independent vectors whose linear combinations give the whole vector space. If the vector space is normed, one also considers so-called Schauder bases, which are again families of linearly independent vectors of the given space that gives all other vectors not necessarily as linear combinations, but, at least, as limits of sequences of this combinations. In Hilbert spaces, it is natural to consider Schauder bases whose elements are *orthonormal*. Such bases are known as "Hilbert bases":

Definition 7.217 (Orthonormal Families and Hilbert Bases) Let H be any Hilbert space and $\Omega \subseteq H$.

(i) Ω is an "orthogonal family" if $x \perp x'$ for all $x, x' \in \Omega$, $x \neq x'$.
(ii) An orthogonal family Ω is an "orthonormal family" if $\|x\| = 1$ for all $x \in \Omega$.
(iii) An orthonormal family Ω is a "Hilbert basis" of H if it is maximal with respect to inclusion, that is, Ω is not contained in a second orthonormal family strictly larger than Ω.

A Hilbert basis generates the whole Hilbert space in the following sense:

Lemma 7.218 *If $\Omega \subseteq H$ is a Hilbert basis, then $\overline{\mathrm{span}(\Omega)} = H$.*

Proof Let Ω be a maximal orthonormal family and assume that $\overline{\mathrm{span}(\Omega)} \neq H$. Let $x \in H \backslash \overline{\mathrm{span}(\Omega)}$. By the orthogonal decomposition theorem, $P_{\overline{\mathrm{span}(\Omega)}^{\perp}}(x) \neq 0$. Thus,

$$\Omega \cup \left\{ \left\| P_{\overline{\mathrm{span}(\Omega)}^{\perp}}(x) \right\|^{-1} P_{\overline{\mathrm{span}(\Omega)}^{\perp}}(x) \right\}$$

is an orthonormal family. But this contradicts the maximality of Ω. □

The following result is a consequence of Zorn's lemma (or, equivalently, the axiom of choice of ZFC):

Lemma 7.219 *Any orthonormal family Ω_0 in an arbitrary Hilbert space H is contained in some maximal orthonormal family of this space. In particular, any Hilbert space of non-zero dimension has a Hilbert basis.*

An important class of Hilbert spaces is given by the separable ones:

Definition 7.220 (Separable Hilbert Spaces) We say that the Hilbert space H is "separable" if it possesses a countable Hilbert basis.

The following conditions are equivalent to the above definition:

Theorem 7.221 (Separability of Hilbert Spaces) *Let H be a Hilbert space. The following conditions are equivalent:*

(i) H is separable.
(ii) Any Hilbert basis of H is countable.
(iii) H has a dense countable subset.

Proof For simplicity and without loss of generality, we only consider the case of a real Hilbert space.

1. Clearly, Assertion (ii) implies (i).

2. Assume (i) and let $\{e_n\}_{n=1}^N$, $N \in \mathbb{N} \cup \{\infty\}$, be a Hilbert basis of H. The subset

$$\bigcup_{N \in \mathbb{N}} \left\{ \frac{k_1}{m_1} e_{n_1} + \cdots + \frac{k_N}{m_N} e_{n_N} \ : k_1, \ldots, k_N \in \mathbb{Z},\ m_1, \ldots, m_N \in \mathbb{N},\ n_1, \ldots, n_N \in \mathbb{N} \right\} \subseteq H$$

is countable, because countable unions of countable sets are itself countable. Additionally, this set is dense in H, as, by Lemma 7.218,

$$H = \overline{\mathrm{span}\{e_n : n = 1, \ldots, N\}} \,.$$

Hence, if H is separable, then Assertion (iii) holds true.

3. Assume Assertion (iii), i.e., H has a dense subset $D \subseteq H$, which is countable. Suppose the existence of some non-countable Hilbert basis $\{e_i\}_{i \in I}$. For all $i \in I$, let $x_i \in D$ be any vector such that $\|e_i - x_i\| < 1/2$. Such vectors must exist, by density of D in H. Then, for all $i, i' \in I$,

$$\|e_i - e_{i'}\| = \|(e_i - x_i) - (e_{i'} - x_{i'}) + (x_i - x_{i'})\| \leq 1 + \|x_i - x_{i'}\| \ .$$

In particular, for $i \neq i'$,

$$\|x_i - x_{i'}\| \geq \|e_i - e_{i'}\| - 1 = \sqrt{2} - 1 > 0$$

and, hence, $x_i \neq x_{i'}$ for any $i, i' \in I$, $i \neq i'$. From this, D must contain a non-enumerable subset

$$\{x_i \ : \ i \in I\} \subseteq D \,.$$

This contradicts the fact that D is countable and we conclude that any Hilbert basis of H is countable, whenever H possesses a dense countable subset. In other words, Condition (ii) follows from (iii).

□

For technical simplicity and without restriction on applications to quantum statistical mechanics, we frequently assume that Hilbert spaces are separable.

Theorem 7.222 (Bessel Inequality) *Let* $(H, \|\cdot\|)$ *be any Hilbert space. Take any orthonormal family* $\{e_n\}_{n=1}^N \subseteq H$, $N \in \mathcal{N} \cup \{\infty\}$. *Then, for all* $x \in H$,

$$\sum_{n=1}^{N} |\langle e_n, x\rangle|^2 \leq \|x\|^2 \,.$$

Proof Fix any $x \in H$ and note that, for all $M \in \mathbb{N}$ so that $M \leq N$,

$$\left(x - \sum_{n=1}^{M} \langle \mathrm{e}_n, x \rangle \, \mathrm{e}_n\right) \perp \mathrm{e}_m$$

for all $m \in \{1, \ldots, M\}$. In particular,

$$\left(x - \sum_{n=1}^{M} \langle \mathrm{e}_n, x \rangle \, \mathrm{e}_n\right) \perp \left(\sum_{n=1}^{M} \langle \mathrm{e}_n, x \rangle \, \mathrm{e}_n\right)$$

and one has

$$\sum_{n=1}^{M} | \langle \mathrm{e}_n, x \rangle |^2 + \left\| x - \sum_{n=1}^{M} \langle \mathrm{e}_n, x \rangle \, \mathrm{e}_n \right\|^2 = \|x\|^2 \ .$$

From this, we arrive at

$$\sum_{n=1}^{N} | \langle \mathrm{e}_n, x \rangle |^2 \doteq \lim_{M \to N} \sum_{n=1}^{M} | \langle \mathrm{e}_n, x \rangle |^2 \leq \|x\|^2 \ .$$

□

The following notions of series convergence are relevant in the theory of Hilbert spaces, in particular in the context of basis decompositions:

Definition 7.223 (Absolute and Unconditional Convergence) Let X be any normed space and $\{x_i\}_{i \in I} \subseteq X$ a countable family of vectors.

(i) We say that the sum of this family "unconditionally converges" if, for any enumeration $\{i_n\}_{n=1}^{N}$, $N \in \mathbb{N} \cup \{\infty\}$, of I, the limit

$$\lim_{M \to N} \sum_{n=1}^{M} x_{i_n} \in X$$

exists (in X) and is independent of the chosen enumeration. In this case, this limit is denoted by $\sum\limits_{i \in I} x_i$.

(ii) We say that the sum of the family $\{x_i\}_{i \in I}$ "absolutely converges" (notation: $\sum\limits_{i \in I} \|x_i\| < \infty$) if

$$\sum_{n=1}^{N} \left\| x_{i_n} \right\| < \infty$$

for some enumeration $\{i_n\}_{n=1}^{N}$, $N \in \mathbb{N} \cup \{\infty\}$, of I.

The following proposition gathers some important well-known facts about these two notions of convergence for sums in normed spaces:

Proposition 7.224 *Let X be a Banach space and $\{x_i\}_{i\in I} \subseteq X$ any countable family of vectors.*

(i) If the sum of $\{x_i\}_{i\in I}$ absolutely converges, then it unconditionally converges.
(ii) If X is finite dimensional, then the sum of $\{x_i\}_{i\in I}$ absolutely converges iff it unconditionally converges.
(iii) If the dimension of X is not finite, then there is a family $\{x_i\}_{i\in I} \subseteq X$ whose sum unconditionally converges, but not absolutely (Dvoretzky-Rogers theorem).
(iv) For $X = \mathbb{R}$, if the sum of $\{x_i\}_{i\in I}$ does not absolutely converge, but the limit

$$\lim_{N\to\infty} \sum_{n=1}^{N} x_{i_n} \in \mathbb{R}$$

exists for some enumeration $\{i_n\}_{n=1}^{\infty}$ of I, then, for all $\alpha \in \mathbb{R}$, there is an enumeration $\{i_n^{(\alpha)}\}_{n=1}^{\infty}$ of I such that

$$\lim_{N\to\infty} \sum_{n=1}^{N} x_{i_n^{(\alpha)}} = \alpha \ .$$

The following important result is a consequence of the Bessel inequality and the above proposition:

Theorem 7.225 *Let H be a separable Hilbert space and $\Omega \subseteq H$ any orthonormal family. Then, for all $x \in H$, the sum of $\{\langle \mathrm{e}, x\rangle\, \mathrm{e}\}_{\mathrm{e}\in\Omega}$ unconditionally converges with*

$$\left\| \sum_{\mathrm{e}\in\Omega} \langle \mathrm{e}, x\rangle\, \mathrm{e} \right\|^2 = \sum_{\mathrm{e}\in\Omega} |\, \langle \mathrm{e}, x\rangle \,|^2 \qquad \textit{and} \qquad \sum_{\mathrm{e}\in\Omega} \langle \mathrm{e}, x\rangle\, \mathrm{e} = P_{\overline{\mathrm{span}(\Omega)}}(x) \ .$$

Proof Exercise. *Hint*: Prove the assertion first for $x \in \mathrm{span}(\Omega)$ and then generalize it by representing $x \in H$ as the limit of a sequence in $\mathrm{span}(\Omega)$. Proceeding in this way, the Bessel inequality is useful to prove convergence of the terms appearing in the above equalities. □

Observe that, in general, the sum of $\{\langle \mathrm{e}, x\rangle\, \mathrm{e}\}_{\mathrm{e}\in\Omega}$ does not absolutely converge. Recall that $\overline{\mathrm{span}(\Omega)} = H$ when $\Omega \subseteq H$ is a Hilbert basis. In this case, one has:

Corollary 7.226 *Let H be any separable Hilbert space and $\Omega \subseteq H$ a Hilbert basis. For all $x \in H$, the sum of $\{\langle \mathrm{e}, x\rangle\, \mathrm{e}\}_{\mathrm{e}\in\Omega}$ unconditionally converges with*

$$\|x\|^2 = \sum_{\mathrm{e}\in\Omega} |\, \langle \mathrm{e}, x\rangle \,|^2 \qquad \textit{and} \qquad \sum_{\mathrm{e}\in\Omega} \langle \mathrm{e}, x\rangle\, \mathrm{e} = x \ .$$

The first equality of the above corollary is called "Parseval's identity."

Exercise 7.227 Let $(H, \|\cdot\|)$ be any separable Hilbert space and $\{\mathrm{e}_n\}_{n=1}^N, \{\tilde{\mathrm{e}}_n\}_{n=1}^N \subseteq H$, $N \in \mathbb{N} \cup \{\infty\}$, two Hilbert bases. Show that the condition $U(\mathrm{e}_n) = \tilde{\mathrm{e}}_n$, $n \in \{1, \ldots, N\}$, defines a unique unitary operator $U \in \mathcal{B}(H)$.

We show below that the dimension classifies the equivalence classes of separable Hilbert spaces:

Definition 7.228 **(ℓ^2 Spaces)** Let Ω be any countable set and define $\ell^2(\Omega) \subseteq \mathcal{F}(\Omega; \mathbb{K})$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$, as being the space of the functions $f : \Omega \to \mathbb{C}$ for which the sum of (non-negative numbers) $\{|f(\omega)|^2\}_{\omega \in \Omega}$ absolutely converges. For $f \in \ell^2(\Omega)$, define

$$\|f\|_2 \doteq \sqrt{\sum_{\omega \in \Omega} |f(\omega)|^2} \in \mathbb{R}_0^+ .$$

Exercise 7.229 Let Ω and $\tilde{\Omega}$ be two countable sets and assume that there is a one-to-one correspondence between them.

(i) Show that $(\ell^2(\Omega), \|\cdot\|_2)$ and $(\ell^2(\tilde{\Omega}), \|\cdot\|_2)$ are unitarily equivalent separable Hilbert spaces.
(ii) Prove that $\{\delta_{\omega,\cdot}\}_{\omega \in \Omega}$ is a Hilbert basis of $\ell^2(\Omega)$. Here, $\delta_{\cdot,\cdot}$ is the Kronecker delta, i.e., $\delta_{\omega,\omega'} = 1$ if $\omega = \omega'$ and $\delta_{\omega,\omega'} = 0$ else.
(iii) Show that, for all $f, f' \in \ell^2(\Omega)$, the family $\{\overline{f(\omega)} f'(\omega)\}_{\omega \in \Omega}$ of numbers absolutely converges and

$$\langle f, f' \rangle_2 = \sum_{\omega \in \Omega} \overline{f(\omega)} f'(\omega) ,$$

where $\langle \cdot, \cdot \rangle_2$ is the scalar product of the Hilbert space $\ell^2(\Omega)$.

Let H be any separable Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $\Omega \subseteq H$ some Hilbert basis. For all $x \in H$, define the function $f_x : \Omega \to \mathbb{K}$ by

$$f_x(\mathrm{e}) \doteq \langle \mathrm{e}, x \rangle , \qquad \mathrm{e} \in \Omega .$$

Note that, by Parseval's identity, $\|f_x\|_2 = \|x\|$ (see Corollary 7.226). More precisely, we obtain a unitary transformation:

Proposition 7.230 *The mapping $x \mapsto f_x$ from any separable Hilbert space H to $\ell^2(\Omega)$ is unitary.*

Proof Exercise. □

Corollary 7.231 *Two separable Hilbert spaces are unitarily equivalent iff they have the same dimension.*

7.3.5 The Space of Bounded Operators on a Hilbert Space as a ∗-Vector Space

We introduce now the notion of adjoint operator of a bounded operator on a Hilbert space. Thus, let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and $A \in \mathcal{B}(H)$. By the Cauchy-Schwarz inequality, note that for all $x' \in H$, $\langle x', A(\cdot)\rangle$ defines a bounded linear functional $\varphi_{A,x'} \in H^{\mathrm{td}}$. Hence, for any fixed $A \in \mathcal{B}(H)$, by the Riesz-Fréchet theorem (Theorem 7.214), for all $x' \in H$, there is a unique $x_{A,x'} \in H$ such that, for all $x'' \in H$,

$$\langle x', A(x'')\rangle = \langle x_{A,x'}, x''\rangle \ .$$

From the uniqueness of such vectors $x_{A,x'} \in H$, $x' \in H$, and antilinearity of the canonical identification $H \to H^{\mathrm{td}}$, it follows that the mapping $A^* : H \to H$, $x' \mapsto x_{A,x'}$ is *linear*. The linear transformation $A^* \in \mathcal{L}(H)$ is called the "adjoint" operator of $A \in \mathcal{B}(H)$. By construction,

$$\langle x, A(x')\rangle = \langle A^*(x), x'\rangle \qquad \text{and} \qquad \langle A(x), x'\rangle = \langle x, A^*(x')\rangle$$

for all $x, x' \in H$ and $A \in \mathcal{B}(H)$. Observe that these two identities can be satisfied by at most one element $A^* \in \mathcal{L}(H)$ and can thus be used as the definition of the adjoint operator A^*. The discussion above shows that such a linear map must exist for all $A \in \mathcal{B}(H)$, because of the Riesz-Fréchet theorem (Theorem 7.214).

The following proposition gathers important properties of adjoint operators:

Proposition 7.232 *Let H be a Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Take any two operators $A, A' \in \mathcal{B}(H)$.*

(i) C^*-property. $A^* \in \mathcal{B}(H)$ *with* $\|A^*\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}$ *and* $\|A^*A\|_{\mathrm{op}} = \|A\|_{\mathrm{op}}^2$.
(ii) $*$-algebra property. $(AA')^* = A'^*A^*$.
(iii) Antilinearity. $(A + A')^* = A^* + A'^*$ *and* $(\alpha A)^* = \overline{\alpha} A^*$ *for all* $\alpha \in \mathbb{K}$.
(iv) Involution. $(A^*)^* = A$.
(v) "Polarization identity for operators." *If* $\mathbb{K} = \mathbb{C}$,

$$A^*A' = \frac{1}{4}\sum_{n=1}^{4}(-i)^n(A + i^n A')^*(A + i^n A') \ .$$

For $\mathbb{K} = \mathbb{R}$, *one only has*

$$\frac{A^*A' + A'^*A}{2} = \frac{1}{4}\sum_{n=1}^{2}(-1)^n(A + (-1)^n A')^*(A + (-1)^n A') \ .$$

Proof (ii–iv) are immediate consequences of the uniqueness of adjoint operators, and (v) directly follows from (iii). Thus, we prove (i): Take any $A \in \mathcal{B}(H)$, $A \neq 0$. Note from Lemma 7.205 that

$$\begin{aligned}\|A^*\|_{\text{op}} &= \sup_{\substack{x,x'\in X\\ \|x\|=\|x'\|=1}} \left|\langle x, A^*(x')\rangle\right| = \sup_{\substack{x,x'\in X\\ \|x\|=\|x'\|=1}} \left|\langle A^*(x), x'\rangle\right| \\ &= \sup_{\substack{x,x'\in X\\ \|x\|=\|x'\|=1}} \left|\langle x, A(x')\rangle\right| = \|A\|_{\text{op}} \ .\end{aligned}$$

Additionally, by the Cauchy-Schwarz inequality,

$$\begin{aligned}\|A\|_{\text{op}} &= \sup_{x\in H,\ \|x\|=1} \|A(x)\| = \sup_{x\in H,\ \|x\|=1} \sqrt{\langle A(x), A(x)\rangle} \\ &= \sup_{x\in H,\ \|x\|=1} \sqrt{|\langle x, A^*A(x)\rangle|} \le \sqrt{\|A^*A\|_{\text{op}}} \ .\end{aligned}$$

Therefore,

$$\|A\|_{\text{op}}^2 \le \left\|A^*A\right\|_{\text{op}} \le \left\|A^*\right\|_{\text{op}} \|A\|_{\text{op}} = \|A\|_{\text{op}}^2 \ .$$

□

Note from parts (iii) and (iv) of the above proposition that the operation $(\cdot)^*$ of taking the adjoint of a bounded operator is a complex conjugation in $\mathcal{B}(H)$, where H is any *complex* Hilbert space. In particular, $\mathcal{B}(H)$ is canonically a $*$-vector space (Definition 7.52). By part (ii), it is even a $*$-algebra (Definition 7.60). Finally, recalling that $\mathcal{B}(X)$ is a Banach algebra for any Banach space X (see Exercise 7.86 (v)) and that Hilbert spaces are (by definition) complete, by part (i), $\mathcal{B}(H)$ is a C^*-algebra with respect to the operator norm. See Definition 7.85. In particular, any self-conjugate subalgebra of $\mathcal{B}(H)$, which is closed with respect to the operator norm, is again a C^*-algebra. Such C^*-algebras, i.e., C^*-algebras that are operator algebras on Hilbert spaces, are called "concrete" C^*-algebras. Thanks to the celebrated Gelfand-Naimark theorem (Theorem 4.89), it turns out that any (abstract) C^*-algebra is equivalent, as a $*$-algebra, to a concrete one, that is, a self-adjoint subalgebra of $\mathcal{B}(H)$ for some complex Hilbert space H.

In the theory of operators on complex Hilbert spaces H, real elements of $\mathcal{B}(H)$ are called "self-adjoint" elements, and, frequently, the notation $\mathcal{B}(H)^{\mathbb{R}}$ is used for the real subspace $\text{Re}\{\mathcal{B}(H)\} \subseteq \mathcal{B}(H)$. Similarly, self-conjugate linear functionals $\mathcal{B}(H) \to \mathbb{C}$, i.e., real elements of the $*$-vector space $\mathcal{B}(H)' \doteq \mathcal{L}(\mathcal{B}(H); \mathbb{C})$, are called "Hermitian" linear functionals on $\mathcal{B}(H)$. (By definition, such Hermitian functionals φ satisfy $\varphi(A^*) = \overline{\varphi(A)}$ for $A \in \mathcal{B}(H)$.)

In the case of *real* Hilbert spaces, the operation $(\cdot)^*$ of taking the adjoint of a bounded operator is a *linear* involution $\mathcal{B}(H) \to \mathcal{B}(H)$. Operators $A \in \mathcal{B}(H)$, where H is any *real* Hilbert space, that are invariant under this operation, i.e., $A^* = A$, are said to be "self-adjoint," like in the complex case.

By Exercise 7.56 (iv), note that, for any complex Hilbert space H, the (real) space $\text{Re}\{\mathcal{B}(H)'\}$ of Hermitian functionals on $\mathcal{B}(H)$ can be canonically identified (by restriction of functionals) with $(\mathcal{B}(H)^{\mathbb{R}})' = \mathcal{L}(\mathcal{B}(H)^{\mathbb{R}}; \mathbb{R})$. As C^*-algebras are $*$-normed spaces, observe also that $\mathcal{B}(H)^{\text{td}} \doteq \mathcal{B}(\mathcal{B}(H); \mathbb{C})$ is a self-conjugate subspace of (the $*$-vector space) $\mathcal{B}(H)'$. See Exercise 7.77 (i). In particular, by Exercise 7.77 (ii), $\mathrm{Re}\{\mathcal{B}(H)^{\mathrm{td}}\}$ can be canonically identified (by restriction of functionals) with $(\mathcal{B}(H)^{\mathbb{R}})^{\mathrm{td}} = \mathcal{B}(\mathcal{B}(H)^{\mathbb{R}}; \mathbb{R})$.

The following proposition is a stronger version of Lemma 7.205, for the special case of self-adjoint operators:

Proposition 7.233 *Let H be any (real or complex) Hilbert space. For any self-adjoint $A \in \mathcal{B}(H)$,*

$$\|A\|_{\mathrm{op}} = \sup_{x\in H,\ \|x\|=1} |\langle x, A(x)\rangle| \ .$$

Proof

1. Take any self-adjoint $A \in \mathcal{B}(H)$. Observe from Lemma 7.205 that

$$\|A\|_{\mathrm{op}} = \sup_{x\in H,\ \|x\|=1}\ \sup_{x'\in H,\ \|x'\|=1} \left|\langle x, A(x')\rangle\right| \geq \sup_{x\in H,\ \|x\|=1} |\langle x, A(x)\rangle| \ .$$

Hence, it is enough to show that

$$\sup_{x\in H,\ \|x\|=1}\ \sup_{x'\in H,\ \|x'\|=1} \left|\langle x, A(x')\rangle\right| \leq \sup_{x\in H,\ \|x\|=1} |\langle x, A(x)\rangle| \doteq M \ .$$

2. As $A = A^*$, one has that

$$\begin{aligned}\left\langle A(x+x'), x+x'\right\rangle - \left\langle A(x-x'), x-x'\right\rangle &= 2\left\langle A(x), x'\right\rangle + 2\left\langle A(x'), x\right\rangle \\ &= 2\left\langle A(x), x'\right\rangle + 2\left\langle x', A(x)\right\rangle \\ &= 4\mathrm{Re}\{\left\langle A(x'), x\right\rangle\} \ .\end{aligned}$$

By the definition of the positive constant M along with the parallelogram identity, we obtain that

$$4\mathrm{Re}\{\left\langle A(x'), x\right\rangle\} \leq M\left(\left\|x+x'\right\|^2 + \left\|x-x'\right\|^2\right) = 2M(\|x\|^2 + \left\|x'\right\|^2) \ .$$

3. Hence,

$$\sup_{x\in H,\ \|x\|=1}\ \sup_{x'\in H,\ \|x'\|=1} \mathrm{Re}\left\{\left\langle A(x'), x\right\rangle\right\} \leq M = \sup_{x\in H,\ \|x\|=1} |\langle x, A(x)\rangle| \ .$$

Since $A = A^*$, observe finally that

$$\sup_{x\in H,\ \|x\|=1}\ \sup_{x'\in H,\ \|x'\|=1} \left|\langle x', A(x)\rangle\right| = \sup_{x\in H,\ \|x\|=1}\ \sup_{x'\in H,\ \|x'\|=1} \mathrm{Re}\left\{\left\langle A(x'), x\right\rangle\right\} \ ,$$

using in the complex case that $|\mathrm{Re}\,(z)| \leq |z|$ for $z \in \mathbb{C}$ and the fact that, for any $x, x' \in H$, there is $c \in \mathbb{C}$, $|c| = 1$, such that $\left\langle cx', A(x)\right\rangle = \left|\langle x', A(x)\rangle\right| \in \mathbb{R}_0^+$.

□

Corollary 7.234 *Let H be any* complex *Hilbert space. $A \in \mathcal{B}(H)^{\mathbb{R}}$, i.e., A is self-adjoint, iff $\langle x, A(x)\rangle \in \mathbb{R}$ for all $x \in H$. In particular, for any (possibly non-self-adjoint) $A \in \mathcal{B}(H)$, $\langle x, A(x)\rangle = 0$ for all $x \in H$ only if $A = 0$.*

Proof Take any $A \in \mathcal{B}(H)$. If $A \in \mathcal{B}(H)^{\mathbb{R}}$, then, for all $x \in H$,

$$\langle x, A(x)\rangle = \langle A(x), x\rangle = \overline{\langle x, A(x)\rangle} \in \mathbb{R} .$$

Now suppose that $\langle x, A(x)\rangle \in \mathbb{R}$ for all $x \in H$. Then, for all $x \in H$,

$$\langle x, \mathrm{Re}\{A\}(x)\rangle + i \, \langle x, \mathrm{Im}\{A\}(x)\rangle \in \mathbb{R} .$$

As $\mathrm{Im}\{A\} \in \mathcal{B}(H)^{\mathbb{R}}$, it follows that $\langle x, \mathrm{Im}\{A\}(x)\rangle = 0$ for all $x \in H$. By the above theorem, $\mathrm{Im}\{A\} = 0$. This implies that $A = \mathrm{Re}\{A\} \in \mathcal{B}(H)^{\mathbb{R}}$. □

We introduce now an important class of operators that is larger than the class of self-adjoint ones. It refers to the operators that commute with their adjoint:

Definition 7.235 (Normal Operators) Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. An operator $A \in \mathcal{B}(H)$ is "normal" if $AA^* = A^*A$.

Note that any self-adjoint $A \in \mathcal{B}(H)$ is (trivially) normal. Another important example of normal operators are the unitary ones, as defined in Definition 7.209 for $H_1 = H_2 = H$:

Lemma 7.236 *Let H be any (real or complex) Hilbert space. For any unitary operator $U \in \mathcal{B}(H)$, $U^{-1} = U^*$, i.e., $U^*U = UU^* = \mathrm{id}_H$. In particular, unitary operators are normal.*

Proof Let $U \in \mathcal{B}(H)$ be a unitary operator. Then, for all $x, x' \in H$,

$$\left\langle x, x'\right\rangle = \left\langle U(x), U(x')\right\rangle = \left\langle x, U^*U(x)\right\rangle ,$$

thanks to Lemma 7.210. Hence, for $x, x' \in H$,

$$\left\langle x, x'\right\rangle = \left\langle x, \mathrm{id}_H(x')\right\rangle = \left\langle x, U^*U(x')\right\rangle$$

and one thus has $U^*U = \mathrm{id}_H$; see, for instance, Lemma 7.205. As $U : H \to H$ is by definition a one-to-one correspondence, we conclude that $U^{-1} = U^*$. □

Kernels of normal operators have the following properties, which will be useful later on:

Lemma 7.237 *Let H be a Hilbert space and $A \in \mathcal{B}(H)$.*

(i) $\ker(A) = A^*(H)^{\perp}$ *and* $\ker(A^*) = A(H)^{\perp}$.
(ii) *If A is a normal operator, then* $\ker(A^*) = \ker(A)$.
(iii) *If A is a normal operator, then* $\ker(A) = A(H)^{\perp}$.

Proof Observe that $x \in \ker(A)$ iff

$$\langle x', A(x)\rangle = \langle A^*(x'), x\rangle = 0$$

for all $x' \in H$. But the second equality yields $x \in A^*(H)^\perp$. Hence, $\ker(A) = A^*(H)^\perp$. Using the identity $A^{**} = A$, we also conclude that $\ker(A^*) = A(H)^\perp$. This proves (i). Let $A \in \mathcal{B}(H)$ be a normal operator. Then, for all $x \in H$,

$$0 = \langle x, (A^*A - AA^*)(x)\rangle = \langle A(x), A(x)\rangle - \langle A^*(x), A^*(x)\rangle \ ,$$

that is,

$$\|A(x)\| = \|A^*(x)\|$$

for all $x \in H$. In particular, $\ker(A^*) = \ker(A)$. This proves (ii). Finally, (iii) is a direct consequence of (i) and (ii). □

Recalling that the double orthogonal complement of a subspace of any Hilbert space H is the same as its closure, by the third part of the last lemma, observe that, for any normal operator $A \in \mathcal{B}(H)$,

$$H = \ker(A) \oplus_2 \overline{A(H)} \ .$$

See Definition 7.211 and Proposition 7.212. In particular, any normal operator $A \in \mathcal{B}(H)$ can be seen, by restriction, as a one-to-one correspondence $\overline{A(H)} \to A(H)$.

In the next proposition, we show that orthogonal projectors (see Lemma 7.213) are a special case of self-adjoint operators:

Proposition 7.238 *Let H be any (real or complex) Hilbert space. For any closed subspace $G \subseteq H$, $P_G = P_G^*$. Conversely, if $P \in \mathcal{B}(H)$ is a self-adjoint projector, i.e., $P \circ P = P = P^*$, then $P(H) \subseteq H$ is a closed subspace, and P is the orthogonal projector associated with it. In particular, $G \mapsto P_G$ is a one-to-one correspondence between closed subspaces of H and self-adjoint projectors in $\mathcal{B}(H)$.*

Proof

1. Let G be any closed subspace of H. Then, for all $x, x' \in H$,

$$\begin{aligned}\langle P_G^*(x), x'\rangle &= \langle x, P_G(x')\rangle \\ &= \langle P_G(x) + P_{G^\perp}(x), P_G(x')\rangle \\ &= \langle P_G(x), P_G(x')\rangle \\ &= \langle P_G(x), P_G(x') + P_{G^\perp}(x')\rangle = \langle P_G(x), x'\rangle \ .\end{aligned}$$

Hence, by the uniqueness of adjoint operators, $P_G^* = P_G$.

2. Let now $P \in \mathcal{B}(H)$ be any projector, $(x_n)_{n\in\mathbb{N}}$ a Cauchy sequence in the subspace $P(H) \subseteq H$, and $x \in H$ its limit. As P is a projector, for all $n \in \mathbb{N}$, $P(x_n) = x_n$. As P is continuous, one thus has that $P(x) = x$. In particular, $x \in P(H)$, and we conclude that $P(H) \subseteq H$ is a closed subspace.
3. If P is self-adjoint, then $\ker(P) = P(H)^\perp$, thanks to Lemma 7.237 (iii). Note from Corollary 7.208 that $\ker(P)^\perp = P(H)^{\perp\perp} = P(H)$, because $P(H)$ is a closed subspace. Hence, for all $x \in H$,

$$\begin{aligned} P(x) &= P(P_{\ker(P)}(x) + P_{\ker(P)^\perp}(x)) \\ &= P(P_{\ker(P)^\perp}(x)) \\ &= P(P_{P(H)}(x)) = P_{P(H)}(x) \; . \end{aligned}$$

The last equality follows from the fact that $P(x) = x$ for all $x \in P(H)$, P being a projector.

□

Exercise 7.239 Let H be any (real or complex) Hilbert space. Prove the following properties of orthogonal projectors:

(i) For any closed subspaces $G_1, G_2 \subseteq H$, $G_2 \subseteq G_1$,

$$P_{G_2} P_{G_1} = P_{G_1} P_{G_2} = P_{G_2} \; .$$

(ii) For any two self-adjoint projectors $P, P' \in \mathcal{B}(H)$, $P(H) \perp P'(H)$ iff

$$P P' = P' P = 0 \; .$$

7.3.6 The Spectrum of Bounded Operators on Hilbert Spaces

Given a normed space X, recall that a bounded operator $A \in \mathcal{B}(X)$ is "invertible" in $\mathcal{B}(X)$ if there are $A_L^{-1}, A_R^{-1} \in \mathcal{B}(X)$ such that

$$A A_R^{-1} = A_L^{-1} A = \mathrm{id}_X \; .$$

In this case, note that $A_R^{-1} = A_L^{-1} \doteq A^{-1}$. The "inverse element," or the "inverse," $A^{-1} \in \mathcal{B}(X)$ is unique when it exists; see, for instance, Lemma 4.17 (i), which refers to general unital associative algebras. It is used to define the spectrum of any bounded operator on a normed space:

Definition 7.240 (Spectrum of a Bounded Operator) Let X be any normed space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For all $A \in \mathcal{B}(X)$, we define the "resolvent set" by

$$R(A) \doteq \{z \in \mathbb{K} \; : \; (A - z\mathrm{id}_X) \text{ has an inverse in } \mathcal{B}(X)\} \subseteq \mathbb{K} \; .$$

The "spectrum" of $A \in \mathcal{B}(X)$ is defined as being the complement in $\mathbb{K}$ of its resolvent:

$$\sigma(A) \doteq \mathbb{K} \backslash R(A) .$$

By the (Banach-Schauder) open mapping theorem [18, Sections 2.11 and 2.12], if X is a Banach space, then every linear continuous operator $A \in \mathcal{B}(X)$ that is bijective has a continuous inverse, that is, $A^{-1} \in \mathcal{B}(X)$. Thus, if X is complete (in particular, if X is a Hilbert space), then, equivalently, for all $A \in \mathcal{B}(X)$, we may define the resolvent set of A by

$$R(A) = \{z \in \mathbb{K} \ : \ (A - z\mathrm{id}_X) \text{ is a one-to-one correspondence}\} .$$

In other words, in this case, $z \in \sigma(A)$ iff $(A - z\mathrm{id}_X)$ fails to be surjective or injective. We mention this fact for the sake of completeness, but this is not really needed below.

In quantum physics, given some complex Hilbert space H, recall that operators $A \in \mathcal{B}(H)$ are related to quantities associated with some quantum system. The corresponding spectra $\sigma(A)$ of $A \in \mathcal{B}(H)$ are interpreted as being the sets of all values those quantities can effectively take.

We say that $z \in \mathbb{K}$ is an "eigenvalue" of a linear mapping $A \in \mathcal{L}(V)$ on a vector space V over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ if, for some $v \in V$, $v \neq 0$, one has $A(v) = zv$. Clearly, if X is any normed space and $A \in \mathcal{B}(X)$, such a number z is in the spectrum $\sigma(A)$, because $(A - z\mathrm{id}_X)$ is not injective in this case. The set of all eigenvalues of $A \in \mathcal{B}(X)$ is called the "pure point spectrum" of A and is denoted by $\sigma_{\mathrm{pp}}(A) \subseteq \sigma(A)$. If X has finite dimension, then $\sigma(A) = \sigma_{\mathrm{pp}}(A)$ for any $A \in \mathcal{B}(X) = \mathcal{L}(X)$. In infinite dimension, however, for a given $A \in \mathcal{B}(X)$, in general, not any $z \in \sigma(A)$ is an eigenvalue of A, i.e., $\sigma_{\mathrm{pp}}(A) \subsetneq \sigma(A)$.

The spectrum of an operator $A \in \mathcal{B}(H)$ acting on a Hilbert space H is directly related to its so-called numerical range:

Definition 7.241 (Numerical Range of an Operator) Let H be a Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For any operator $A \in \mathcal{B}(H)$, we define its "numerical range" by

$$N(A) \doteq \{\langle x, A(x)\rangle \ : \ x \in H \ , \ \ \|x\| = 1\} \subseteq \mathbb{K} .$$

It turns out that the spectrum of any bounded operator is a subset of the closure of its numerical range:

Proposition 7.242 (The Spectrum Is a Subset of the Numerical Range) *Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For all $A \in \mathcal{B}(H)$,*

$$\sigma(A) \subseteq \overline{N(A)} .$$

Proof

1. Take any fixed $A \in \mathcal{B}(H)$ and $\tilde{z} \in \mathbb{K}\backslash\overline{N(A)}$. Then, as $\overline{N(A)}$ is a closed subset,

$$\inf_{z \in N(A)} |z - \tilde{z}| = d > 0 \, .$$

In particular, for all $x \in H$, $\|x\| = 1$, one has

$$0 < d \leq |\tilde{z} - \langle x, A(x) \rangle | = | \langle x, (A - \tilde{z}\mathrm{id}_H)(x) \rangle | \leq \|(A - \tilde{z}\mathrm{id}_H)(x)\| \ , \tag{7.6}$$

using the Cauchy-Schwarz inequality. From this estimate, we conclude that $A - \tilde{z}\mathrm{id}_H$ is one-to-one (i.e., injective) and that the image $(A - \tilde{z}\mathrm{id}_H)(H) \subseteq H$ is closed: In fact, for any Cauchy sequence

$$y_n \in (A - \tilde{z}\mathrm{id}_H)(H) \, , \qquad n \in \mathbb{N} \, ,$$

by injectivity, there is a (unique) sequence $x_n \in H$, $n \in \mathbb{N}$, such that

$$y_n = (A - \tilde{z}\mathrm{id}_H)(x_n) \, .$$

Moreover, one gets from the above estimate that

$$\|x_n - x_m\| \leq d^{-1} \, \|y_n - y_m\|$$

for all $m, n \in \mathbb{N}$. Thus, $(x_n)_{n \in \mathbb{N}}$ is also a Cauchy sequence, and, by continuity of $(A - \tilde{z}\mathrm{id}_H)$ and completeness of H, $(y_n)_{n \in \mathbb{N}}$ converges to

$$(A - \tilde{z}\mathrm{id}_H)(\lim_{n \to \infty} x_n) \in (A - \tilde{z}\mathrm{id}_H)(H) \, .$$

2. Hence, keeping in mind Proposition 7.212, we deduce that

$$H = (A - \tilde{z}\mathrm{id}_H)(H) \oplus (A - \tilde{z}\mathrm{id}_H)(H)^{\perp} \, .$$

Assume that $H \neq (A - \tilde{z}\mathrm{id}_H)(H)$. Then, there is

$$x \in (A - \tilde{z}\mathrm{id}_H)(H)^{\perp} \subseteq H$$

satisfying $\|x\| = 1$. In particular,

$$\langle x, (A - \tilde{z}\mathrm{id}_H)(x) \rangle = 0$$

and thus $\tilde{z} = \langle x, A(x) \rangle \in N(A)$, which contradicts $\tilde{z} \in \mathbb{K}\backslash\overline{N(A)}$. Hence, $A - \tilde{z}\mathrm{id}_H$ is bijective and, hence, has an inverse $(A - \tilde{z}\mathrm{id}_H)^{-1} \in \mathcal{L}(H)$. Additionally, note that, for any $B \in \mathcal{B}(X)$ with inverse,

$$\left\|B^{-1}\right\|_{\mathrm{op}} = \sup\left\{\frac{\|B^{-1}x\|}{\|x\|} : x \in X,\ x \neq 0\right\}$$
$$= \sup\left\{\frac{\|y\|}{\|By\|} : y \in X,\ y \neq 0\right\}$$
$$= \frac{1}{\inf_{y\in X:\|y\|=1} \|By\|} ,$$

which combined with Eq. (7.6) yields

$$\left\|(A - \tilde{z}\mathrm{id}_H)^{-1}\right\|_{\mathrm{op}} \leq d^{-1} < \infty ,$$

i.e., $(A - \tilde{z}\mathrm{id}_H)^{-1} \in \mathcal{B}(H)$. In other words, for all $\tilde{z} \in \mathbb{K}\backslash\overline{N(A)}$, $\tilde{z} \in R(A)$. Equivalently, $\sigma(A) \subseteq \overline{N(A)}$.

□

Corollary 7.243 *Let H be a Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For any self-adjoint $A \in \mathcal{B}(H)$,*

$$\sigma(A) \subseteq \left[\inf_{x\in H,\ \|x\|=1} \langle x, A(x)\rangle ,\ \sup_{x\in H,\ \|x\|=1} \langle x, A(x)\rangle\right] \subseteq [-\|A\|_{\mathrm{op}} , \|A\|_{\mathrm{op}}] \subseteq \mathbb{R} .$$

Note that the spectrum of an operator can a priori be empty. In the following assertion, we show that this is never the case for self-adjoint operators on Hilbert spaces:

Proposition 7.244 *Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. For any self-adjoint $A \in \mathcal{B}(H)$,*

$$\inf_{x\in H,\ \|x\|=1} \langle x, A(x)\rangle ,\ \sup_{x\in H,\ \|x\|=1} \langle x, A(x)\rangle \in \sigma(A) .$$

In particular, the spectrum of self-adjoint elements of $\mathcal{B}(H)$ is not empty.

Proof

1. For any fixed $A \in B(H)$, $\alpha, \beta \in \mathbb{K}$, $\alpha \neq 0$, one has that $\lambda \in \sigma(A)$ iff $\alpha\lambda + \beta \in \sigma(\alpha A + \beta\mathrm{id}_H)$. To show this, one can use the so-called Neumann series given in Lemma 4.24. Thus, it suffices to prove that $0 \in \sigma(A)$ for any self-adjoint $A \in \mathcal{B}(H)$ for which

$$\inf_{x\in H,\ \|x\|=1} \langle x, A(x)\rangle = 0 .$$

2. Note that, for such an $A \in \mathcal{B}(H)$, the mapping $\langle\cdot,\cdot\rangle_A : H \times H \to \mathbb{K}$ defined by

$$\left\langle x, x'\right\rangle_A \doteq \left\langle x, A(x')\right\rangle , \qquad x, x' \in H ,$$

is a scalar semiproduct. Thus, by the Cauchy-Schwarz inequalities for both $\langle\cdot,\cdot\rangle_A$ and $\langle\cdot,\cdot\rangle$, one has that, for all $x, x' \in H$,

$$\left|\langle x, A(x')\rangle\right| \le \sqrt{\langle x, A(x)\rangle}\sqrt{\langle x', A(x')\rangle} \le \|x\| \sqrt{\|A\|_{\mathrm{op}}}\sqrt{\langle x', A(x')\rangle} .$$

3. As $\inf_{x\in H,\ \|x\|=1} \langle x, A(x)\rangle = 0$, there is a sequence $x_n \in H$, $\|x_n\| = 1$, $n \in \mathbb{N}$, such that

$$\lim_{n\to\infty} \langle x_n, A(x_n)\rangle = 0 .$$

   From the above inequality, it then follows that

$$\begin{aligned}\lim_{n\to\infty} \|A(x_n)\|^2 &= \lim_{n\to\infty} \langle A(x_n), A(x_n)\rangle \\ &\le \lim_{n\to\infty} \|A(x_n)\| \sqrt{\|A\|_{\mathrm{op}}}\sqrt{\langle x_n, A(x_n)\rangle} \\ &\le \sqrt{\|A\|^3_{\mathrm{op}}} \lim_{n\to\infty} \sqrt{\langle x_n, A(x_n)\rangle} = 0 .\end{aligned}$$

4. Assume that $0 \notin \sigma(A)$. Then, there is $A^{-1} \in \mathcal{B}(H)$ such that $A^{-1}(A(x_n)) = x_n$ for all $n \in \mathbb{N}$. By the continuity of A^{-1} and the above estimate for $\|A(x_n)\|$, it would then follow that

$$\lim_{n\to\infty} x_n = 0 .$$

   But this contradicts $\|x_n\| = 1$ and, thus, 0 must be in the spectrum $\sigma(A)$.

□

One important consequence of the last proposition and Proposition 7.233 is the following, purely algebraic, equivalent definition of the operator norm in $\mathcal{B}(H)$:

Corollary 7.245 *Let H be any (real or complex) Hilbert space. For every self-adjoint $A \in \mathcal{B}(H)$, there is a spectral point $\lambda \in \sigma(A)$ such that $|\lambda| = \|A\|_{\mathrm{op}}$. In particular, in this case,*

$$\|A\|_{\mathrm{op}} = \sup_{\lambda\in\sigma(A)} |\lambda| = \max_{\lambda\in\sigma(A)} |\lambda| .$$

The last corollary can be extended to the (more general) case of normal operators. However, in general, for a non-normal $A \in \mathcal{B}(H)$, it may happen that

$$\|A\|_{\mathrm{op}} > \sup_{\lambda\in\sigma(A)} |\lambda| .$$

Observing that, for all $A \in \mathcal{B}(H)$, the operator $A^*A \in \mathcal{B}(H)$ is self-adjoint and $\|A^*A\|_{\text{op}} = \|A\|_{\text{op}}^2$ (by Proposition 7.232 (i)), we can give the norm of any (not necessarily self-adjoint) $A \in \mathcal{B}(H)$ in terms of spectral values:

Corollary 7.246 *Let H be any (real or complex) Hilbert space. For all $A \in \mathcal{B}(H)$,*

$$\|A\|_{\text{op}} = \sup_{\lambda \in \sigma(A^*A)} \sqrt{\lambda}\,.$$

(Notice from Corollary 7.243 that every $\lambda \in \sigma(A^*A)$ is a non-negative real number.)

By the last corollary, the operator norm of any $A \in \mathcal{B}(H)$ can be defined in purely algebraic terms, without any reference to the underlying Hilbert space. This fact is one of the motivations for the abstract theory of C^*-algebras that can be seen as a generalization of the algebras of operators on (complex) Hilbert spaces. In fact, we show in Sect. 4.2 that the above corollary also holds true for general C^*-algebras. See the proof of Corollary 4.32.

7.3.7 Weak and Strong Operator Convergence, von Neumann Algebras, and Commutants

The space of bounded operators on a Hilbert space is canonically a normed space (with the operator norm). Thus, there is a canonical notion of convergence for nets of operator on any Hilbert space. In this section, we introduce two weaker notions of convergence of operators on Hilbert spaces, which turn out to be very important in various parts of the theory of C^*-algebras.

Definition 7.247 (Strong Operator Convergence) Let X be any normed space. A net $(A_i)_{i\in I}$ of bounded operators on X, i.e., a net in $\mathcal{B}(X)$, converges to $A \in \mathcal{B}(X)$ in the "strong operator topology" if, for all $x \in X$, the net $(A_i(x))_{i\in I}$ converges to $A(x)$ in X. $A \in \mathcal{B}(X)$ is called, in this case, the strong operator limit of $(A_i)_{i\in I}$. A subset $\Omega \subseteq \mathcal{B}(X)$ is "closed in the strong operator topology" if, for any net $(A_i)_{i\in I}$ in Ω converging in the strong operator topology to $A \in \mathcal{B}(X)$, one has that $A \in \Omega$.

Note that the strong operator convergence has the Hausdorff property, i.e., the net $(A_i)_{i\in I}$ converges in the strong operator topology to $A \in \mathcal{B}(X)$ and to $B \in \mathcal{B}(X)$ only if $A = B$. Observe also that if the net $(A_i)_{i\in I}$ converges to $A \in \mathcal{B}(X)$ in the sense of the operator norm, then it also converges in the strong operator topology to A. The converse of this property is generally false. In other words, the strong operator convergence is a weaker notion of convergence than the (operator-)norm convergence in $\mathcal{B}(X)$. In particular, any subset $\Omega \subseteq \mathcal{B}(X)$ is closed in the strong operator topology only if it is closed with respect to the operator norm.

Definition 7.248 (von Neumann Algebra) Given a complex Hilbert space H, the self-conjugate subalgebras of $\mathcal{B}(H)$ that contain $\text{id}_H \in \mathcal{B}(H)$ and are closed in the strong operator topology are called (concrete) "von Neumann algebras." In particular, $\mathcal{B}(H)$ itself is a von Neumann algebra.

von Neumann algebras defined as operator algebras, that is, as subalgebras of $\mathcal{B}(H)$ for some complex Hilbert H, are sometimes called "concrete von Neumann algebras." It is also possible to define such algebras without any reference to a Hilbert space. In this case, we call them "abstract von Neumann algebras." It turns out that, similar to the general C^*-algebras, any abstract von Neumann algebra is equivalent, as a $*$-algebra, to some concrete one. In contrast to general C^*-algebras, by various technical reasons, in most recent books on the subject, only concrete von Neumann algebras are considered, like in von Neumann's seminal works. For a detailed explanation of the theory of operator algebras, see, for instance, [60–62]. Notice additionally that it is possible to use order theoretical aspects of the space $\mathcal{B}(H)$ to characterize von Neumann algebras in an abstract way, independent of any (explicit) Hilbert space structure. See Corollary 2.18 and the discussion following it.

Recall that any subset of operators that is closed in the strong operator topology is also norm-closed. Therefore, von Neumann algebras are special cases of (unital) C^*-algebras. Observe however that, for an arbitrary complex Hilbert space H, not every self-conjugate subalgebra of $\mathcal{B}(H)$ which is norm-closed and contains $\mathrm{id}_H \in \mathcal{B}(H)$ is a von Neumann algebra. That is, such an algebra is not necessarily closed in the strong operator topology.

Exercise 7.249 Let X be any Banach space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Take any convergent net $(\alpha_i)_{i\in I}$ in $\mathbb{K}$, and two nets (with the same index set) $(A_i)_{i\in I}$ and $(B_i)_{i\in I}$ in $\mathcal{B}(X)$, both converging in the strong operator topology. Their limits are denoted by $\alpha \in \mathbb{K}$, $A \in \mathcal{B}(X)$ and $B \in \mathcal{B}(X)$, respectively. Prove the following statements:

(i) $(\alpha_i A_i)_{i\in I}$ converges in the strong operator topology to $\alpha A \in \mathcal{B}(X)$.
(ii) $(A_i + B_i)_{i\in I}$ converges in the strong operator topology to $A + B \in \mathcal{B}(X)$.
(iii) $(A_i B_i)_{i\in I}$ converges in the strong operator topology to $AB \in \mathcal{B}(X)$, if the net $(A_i)_{i\in I}$ is norm-bounded.

Remark 7.250 If X is a Banach space, then any strongly convergent *sequence* in $\mathcal{B}(X)$ is norm-bounded. This is a consequence of the (Banach-Steinhaus) "uniform boundedness principle" (Theorem 7.84). Thus, in the case of sequences, one can remove the norm boundedness condition of the third part of the last exercise.

In contrast to the operator-norm convergence, for general Hilbert spaces H, the operation of taking adjoints is generally *not* continuous with respect to the strong operator convergence. That is, a net $(A_i)_{i\in I}$ in $\mathcal{B}(H)$ may converge in the strong operator topology to some $A \in \mathcal{B}(H)$, while $(A_i^*)_{i\in I}$ does not converge (in the strong operator topology) to $A^* \in \mathcal{B}(H)$. In other words, the strong operator convergence is only compatible with the algebraic structure of $\mathcal{B}(H)$, but not with the (whole) $*$-vector space structure.

Definition 7.251 (Weak Operator Convergence) Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. A net $(A_i)_{i\in I}$ of bounded operators on H, i.e., a net in $\mathcal{B}(H)$, converges to $A \in \mathcal{B}(H)$ in the "weak operator topology" if, for all $x, x' \in H$, the net $(\langle x, A_i(x')\rangle)_{i\in I}$ in $\mathbb{K}$ converges to $\langle x, A(x')\rangle \in \mathbb{K}$. $A \in \mathcal{B}(H)$ is called, in this case, the weak operator limit of $(A_i)_{i\in I}$. A subset $\Omega \subseteq \mathcal{B}(H)$ is "closed in the weak operator topology" if, for any net $(A_i)_{i\in I}$ in Ω converging in the weak operator topology to $A \in \mathcal{B}(H)$, one has that $A \in \Omega$.

The weak operator convergence has again the Hausdorff property, that is, the net $(A_i)_{i\in I}$ converges in the weak operator topology to $A \in \mathcal{B}(H)$ and to $B \in \mathcal{B}(H)$ only if $A = B$. Remark additionally that if the net $(A_i)_{i\in I}$ converges to $A \in \mathcal{B}(H)$ in the strong operator topology, then it also converges in the weak operator topology to A. This is a simple consequence of the Cauchy-Schwarz inequality. The converse of this property is generally false. In other words, the weak operator convergence is a notion of convergence that is strictly weaker than the strong operator convergence in $\mathcal{B}(H)$. Thus, any subset $\Omega \subseteq \mathcal{B}(X)$ is closed in the weak operator topology only if it is closed with respect to the strong operator topology. In particular, given a complex Hilbert space H, any self-conjugate algebra of $\mathcal{B}(H)$ that contains $\mathrm{id}_H \in \mathcal{B}(H)$ and is closed in the weak operator topology is a von Neumann algebra. In fact, it turns out that any (concrete) von Neumann is closed in the weak (and not only in the strong) operator topology, thanks to (part of) the "von Neumann's density theorem." See, for instance, [51, Theorem 2.4.11] (also known as the bicommutant theorem).

Exercise 7.252 Let H be any Hilbert space over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Take any convergent net $(\alpha_i)_{i\in I}$ in $\mathbb{K}$, and two nets (with the same index set) $(A_i)_{i\in I}$ and $(B_i)_{i\in I}$ in $\mathcal{B}(H)$, both converging in the weak operator topology. Their limits is denoted by $\alpha \in \mathbb{K}$, $A \in \mathcal{B}(H)$ and $B \in \mathcal{B}(H)$, respectively. Prove the following statements:

(i) $(\alpha_i A_i)_{i\in I}$ converges in the weak operator topology to $\alpha A \in \mathcal{B}(H)$.
(ii) $(A_i + B_i)_{i\in I}$ converges in the weak operator topology to $A + B \in \mathcal{B}(H)$.
(iii) $(A_i^*)_{i\in I}$ converges in the weak operator topology to $A^* \in \mathcal{B}(X)$.

In contrast to the operator-norm and strong convergence, for general Hilbert spaces H, the product of operators is not continuous with respect to the weak operator convergence, not even sequentially continuous. That is, sequences $(A_n)_{n\in\mathbb{N}}$ and $(B_n)_{n\in\mathbb{N}}$ in $\mathcal{B}(H)$ may converge in the weak operator topology to some $A, B \in \mathcal{B}(H)$, respectively, while $(A_n B_n)_{n\in\mathbb{N}}$ does not converge (in the weak operator topology) to $AB \in \mathcal{B}(H)$ (even if, like for the strong operator convergence, sequences converging in the weak operator topology are norm-bounded). However, this time, the operation of taking adjoints is continuous, like for the norm convergence, with respect to the weak operator convergence. In other words, in the case of complex Hilbert spaces, the weak operator convergence is only compatible with the $*$-vector space structure of $\mathcal{B}(H)$, but not with its algebraic structure. This also follows from the uniform boundedness principle (Theorem 7.84), combined with the Cauchy-Schwarz inequality.

In what follows, we discuss an important example of spaces of operators that are closed in the weak operator topology, the so-called commutants. We also briefly discuss their important relation to the von Neumann algebras.

Definition 7.253 (Commutant) Let X be any normed space. We define the "commutant" of any subset $\Omega \subseteq \mathcal{B}(X)$ by

$$\Omega' \doteq \{B \in \mathcal{B}(X) \ : \ [A, B] = 0 \text{ for all } A \in \Omega\} .$$

Here,

$$[A, B] \doteq AB - BA \in \mathcal{B}(X)$$

is the so-called commutator of $A, B \in \mathcal{B}(X)$.

For any nonempty $\Omega_1, \Omega_2 \subseteq \mathcal{B}(X)$ satisfying $\Omega_2 \subseteq \Omega_1$, one has $\Omega_1' \subseteq \Omega_2'$. Moreover, any $\Omega \subseteq \mathcal{B}(X)$ is a subset of its own "bicommutant" $\Omega'' \doteq (\Omega')'$. From these two relations, by a similar argument as the one that proves Lemma 7.279, we conclude in particular that, for every nonempty $\Omega \subseteq \mathcal{B}(X)$, one has $(\Omega'')' = \Omega'$.

For any nonempty $\Omega \subseteq \mathcal{B}(X)$, the commutant $\Omega' \subseteq \mathcal{B}(X)$ is clearly a vector subspace of $\mathcal{B}(X)$ that contains the identity operator $\mathrm{id}_X \in \mathcal{B}(X)$. If H is a complex Hilbert space and $\Omega \subseteq \mathcal{B}(H)$ a self-conjugate subset, then Ω' is a self-conjugate vector subspace of $\mathcal{B}(H)$. Additionally, by the "Leibniz rule" for commutators, that is, the identity

$$[A, BC] = [A, B]C + B[A, C] , \qquad A, B, C \in \mathcal{B}(X) ,$$

it follows that commutants are always subalgebras of $\mathcal{B}(X)$. If H is a complex Hilbert space and $\Omega \subseteq \mathcal{B}(H)$ a self-conjugate subset, then Ω' is a $*$-subalgebra of $\mathcal{B}(H)$.

It turns out that commutants are always closed in the weak operator topology:

Lemma 7.254 *Let H be any Hilbert space and $\Omega \subseteq \mathcal{B}(H)$ any nonempty subset $\Omega \subseteq \mathcal{B}(H)$. Let $(A_i)_{i \in I}$ be any net in the commutant $\Omega' \subseteq \mathcal{B}(H)$ converging in the weak operator topology to some $A \in \mathcal{B}(H)$. Then $A \in \Omega'$, that is, Ω' is closed in the weak operator topology.*

Proof Take any net $(A_i)_{i \in I}$ in the commutant $\Omega' \subseteq \mathcal{B}(H)$, converging in the weak operator topology to some $A \in \mathcal{B}(H)$. For all $B \in \Omega$ and all $x, x' \in H$,

$$\begin{aligned} 0 &= \lim_{i \in I} \langle x, [A_i, B]x' \rangle \\ &= \lim_{i \in I} \left[\langle x, A_i B x' \rangle - \langle B^* x, A_i x' \rangle \right] \\ &= \langle x, ABx' \rangle - \langle B^* x, Ax' \rangle \\ &= \langle x, ABx' \rangle - \langle x, BAx' \rangle = \langle x, [A, B]x' \rangle \ . \end{aligned}$$

Thus, $[A, B] = 0$ for all $B \in \Omega$, i.e., $A \in \Omega'$. □

Corollary 7.255 *Let H be any complex Hilbert space. For any nonempty self-conjugate subset $\Omega \subseteq \mathcal{B}(H)$, the commutant $\Omega' \subseteq \mathcal{B}(H)$ is a von Neumann algebra.*

If H is an arbitrary complex Hilbert space and $\mathfrak{M} \subseteq \mathcal{B}(H)$ is a von Neumann algebra that is a commutant (like in the above corollary), then $\mathfrak{M}'' = \mathfrak{M}$, because, as shown above, the tricommutant of any subset is the same as its single commutant. It turns out that any von Neumann algebra is of this form, thanks again to the "von Neumann's density theorem." See, for instance, [51, Theorem 2.4.11] (also known as the bicommutant theorem). For this reason, the identity $\mathfrak{M}'' = \mathfrak{M}$ is frequently used as *the definition* of (concrete) von Neumann algebras.

7.3.8 Tensor Product Hilbert Spaces

We conclude this section devoted to Hilbert spaces by discussing their tensor product. For simplicity, we restrict ourselves to the case of separable Hilbert spaces, but all results in the present subsection hold true also in the general case. We start by defining "bounded multiantilinear forms":

Definition 7.256 (Bounded Multiantilinear Forms) Let $(H_1, \|\cdot\|^{(1)}), \ldots, (H_N, \|\cdot\|^{(N)})$ be $N \in \mathbb{N}$ Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. The mapping $\Phi : H_1 \times \cdots \times H_N \to \mathbb{K}$ is a "bounded multiantilinear form" if it is antilinear with respect to each of its arguments and, for some $C < \infty$ and all $x_1 \in H_1, \ldots, x_N \in H_N$,

$$|\Phi(x_1, \ldots, x_N)| \leq C \|x_1\|^{(1)} \cdots \|x_N\|^{(N)} .$$

For fixed $x_1 \in H_1, \ldots, x_N \in H_N$, define the mapping

$$x_1 \otimes \cdots \otimes x_N : H_1 \times \cdots \times H_N \to \mathbb{K}$$

to be

$$x_1 \otimes \cdots \otimes x_N (x'_1, \ldots, x'_N) \doteq \langle x'_1, x_1 \rangle^{(1)} \cdots \langle x'_N, x_N \rangle^{(N)}$$

for all $x'_1 \in H_1, \ldots, x'_N \in H_N$. Clearly, by the Cauchy-Schwarz inequality, such mappings are always bounded multiantilinear forms.

In fact, recall from the Riesz-Fréchet theorem (Theorem 7.214) that any complex Hilbert space is canonically identified with its (topological) dual space via an antilinear bijection. Thus, the above multiantilinear forms are canonically identified with multilinear forms on $H_1^{\mathrm{td}} \times \cdots \times H_N^{\mathrm{td}}$ and thus $x_1 \otimes \cdots \otimes x_N$, $x_1 \in H_1, \ldots, x_N \in H_N$, are elements of the tensor product $H_1 \otimes \cdots \otimes H_N$ of normed spaces, in the sense of Definition 7.42.

In the sequel, we define (Hilbert-Schmidt) tensor products of Hilbert spaces as completions of the corresponding algebraic tensor products with respect to a crossnorm (Definition 7.44). With this aim, let $\Omega_1, \ldots, \Omega_N$ be arbitrary Hilbert bases of $N \in \mathbb{N}$ separable Hilbert spaces $(H_1, \|\cdot\|^{(1)}), \ldots, (H_N, \|\cdot\|^{(N)})$, respectively. See Definition 7.217 as well as Lemmata 7.218 and 7.219. For any bounded multiantilinear form $\Phi : H_1 \times \cdots \times H_N \to \mathbb{K}$, we then define the positive quantity

$$\|\Phi\|_{H_1 \otimes_2 \cdots \otimes_2 H_N} \doteq \sqrt{\sum_{\mathrm{e}_1 \in \Omega_1} \cdots \sum_{\mathrm{e}_N \in \Omega_N} |\Phi(\mathrm{e}_1, \ldots, \mathrm{e}_N)|^2} \, .$$

We are now in a position to define the (Hilbert-Schmidt) tensor product of a finite sequence of separable Hilbert spaces:

Definition 7.257 (Tensor Product Hilbert Spaces) Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. The space

$$H_1 \otimes_2 \cdots \otimes_2 H_N \doteq \{\Phi : H_1 \times \cdots \times H_N \to \mathbb{K} \; : \; \Phi \text{ b.m.a.f.}, \; \|\Phi\|_{H_1 \otimes_2 \cdots \otimes_2 H_N} < \infty\}$$

of bounded multiantilinear forms (b.m.a.f.) is by definition the "Hilbert-Schmidt[10] tensor product" of $H_1, \ldots, H_N$. In particular,

$$H_1 \otimes \cdots \otimes H_N \subseteq H_1 \otimes_2 \cdots \otimes_2 H_N \, .$$

For simplicity, we sometimes use the shorter term "tensor product" instead of "Hilbert-Schmidt tensor product," but we keep the symbol $\otimes_2$ to avoid any confusion.

We show in the following proposition that the tensor product of a finite number of separable Hilbert spaces, as defined above, is itself a separable Hilbert space:

Proposition 7.258 *The Hilbert-Schmidt tensor product*

$$(H_1 \otimes_2 \cdots \otimes_2 H_N, \|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N})$$

of $N \in \mathbb{N}$ separable Hilbert spaces $H_1, \ldots, H_N$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ is itself a Hilbert space. Additionally, for all $x_1 \in H_1, \ldots, x_N \in H_N$,

$$x_1 \otimes \cdots \otimes x_N \in H_1 \otimes_2 \cdots \otimes_2 H_N$$

and $\{\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N\}_{(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N}$ is a Hilbert basis of $H_1 \otimes_2 \cdots \otimes_2 H_N$.

Proof

1. Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Their tensor product $H \doteq H_1 \otimes_2 \cdots \otimes_2 H_N$ is clearly a vector space and $\|\cdot\|_H$ is a

[10] We use this name, because, for $N = 2$, the corresponding space can be canonically identified with the space of Hilbert-Schmidt operators $H_1 \to H_2$. We omit the details of this identification, for it is not relevant here.

seminorm in H which satisfies the parallelogram identity. Assume that $\|\Phi\|_H = 0$. In particular, $\Phi(\mathrm{e}_1, \ldots, \mathrm{e}_N) = 0$ for all $(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N$. By antilinearity of Φ in each one of its arguments, $\Phi(x_1, \ldots, x_N) = 0$ for all $x_1 \in \mathrm{span}(\Omega_1), \ldots, x_N \in \mathrm{span}(\Omega_N)$. As Φ is a bounded multiantilinear form, for all convergent sequences $x_{1,n} \in H_1, \ldots, x_{N,n} \in H_N$, $n \in \mathbb{N}$, whose limits are, respectively, $x_1 \in H_1, \ldots, x_N \in H_N$, one has

$$\lim_{n\to\infty} \Phi(x_{1,n}, \ldots, x_{N,n}) = \Phi(x_1, \ldots, x_N) .$$

For $\mathrm{span}(\Omega_1) \subseteq H_1, \ldots, \mathrm{span}(\Omega_N) \subseteq H_N$ are dense sets (Lemma 7.218), we arrive at $\Phi = 0$ and, hence, $\|\cdot\|_H$ is a norm. Since this norm satisfies the parallelogram identity, the tensor product H is a pre-Hilbert space; see Definition 7.194.

2. To prove that the tensor product H is complete with respect to the norm $\|\cdot\|_H$, consider the linear transformation $\xi : H \to \ell^2(\Omega_1 \times \cdots \times \Omega_N)$ uniquely defined by

$$\xi(\Phi)(\mathrm{e}_1, \ldots, \mathrm{e}_N) \doteq \Phi(\mathrm{e}_1, \ldots, \mathrm{e}_N) , \qquad (\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N .$$

See Definition 7.228 for the definition of ℓ^2 spaces. Recall that $\ell^2(\Omega_1 \times \cdots \times \Omega_N)$ is a Hilbert space; see, e.g., Exercise 7.229. Directly by the definitions of the corresponding norms, this mapping is isometric (norm preserving).

3. We show now that ξ is surjective: Fix any $f \in \ell^2(\Omega_1 \times \cdots \times \Omega_N)$ and define the function $\Phi_f : H \to \mathbb{K}$ by

$$\Phi_f(x_1, \ldots, x_N) \doteq \sum_{(\mathrm{e}_1,\ldots,\mathrm{e}_N)\in\Omega_1\times\cdots\times\Omega_N} f(\mathrm{e}_1, \ldots, \mathrm{e}_N) \langle x_1, \mathrm{e}_1\rangle^{(1)} \cdots \langle x_N, \mathrm{e}_N\rangle^{(N)}$$

for any $x_1 \in H_1, \ldots, x_N \in H_N$. Clearly,

$$\sum_{(\mathrm{e}_1,\ldots,\mathrm{e}_N)\in\Omega_1\times\cdots\times\Omega_N} |\langle x_1, \mathrm{e}_1\rangle^{(1)} \cdots \langle x_N, \mathrm{e}_N\rangle^{(N)}|^2 \tag{7.7}$$

$$\leq \left(\sum_{\mathrm{e}_1\in\Omega_1} |\langle x_1, \mathrm{e}_1\rangle^{(1)}|^2\right) \cdots \left(\sum_{\mathrm{e}_N\in\Omega_N} |\langle x_N, \mathrm{e}_N\rangle^{(N)}|^2\right) .$$

(In fact, the above inequality is an equality.) Thus, by Parseval's equality (Corollary 7.226),

$$\sum_{(\mathrm{e}_1,\ldots,\mathrm{e}_N)\in\Omega_1\times\cdots\times\Omega_N} |\langle x_1, \mathrm{e}_1\rangle^{(1)} \cdots \langle x_N, \mathrm{e}_N\rangle^{(N)}|^2 \leq \left(\|x_1\|^{(1)} \cdots \|x_N\|^{(N)}\right)^2 .$$

From this estimate and the Cauchy-Schwarz inequality for $\ell^2(\Omega_1 \times \cdots \times \Omega_N)$, it follows that the sum

$$\sum_{(\mathrm{e}_1,\ldots,\mathrm{e}_N)\in\Omega_1\times\cdots\times\Omega_N} f(\mathrm{e}_1,\ldots,\mathrm{e}_N)\, \langle x_1, \mathrm{e}_1\rangle^{(1)} \cdots \langle x_N, \mathrm{e}_N\rangle^{(N)}$$

absolutely converges and its absolute value is bounded from above by

$$\|x_1\|^{(1)} \cdots \|x_N\|^{(N)}\, \|f\|_{\ell^2(\Omega_1\times\cdots\times\Omega_N)} .$$

Thus, we conclude that Φ_f is a well-defined bounded multiantilinear form. Moreover, by construction, $\xi(\Phi_f) = f$. Hence, ξ is a one-to-one correspondence between $H \doteq H_1 \otimes_2 \cdots \otimes_2 H_N$ and $\ell^2(\Omega_1 \times \cdots \times \Omega_N)$ which preserves the norm.

4. From this and the completeness of $\ell^2(\Omega_1 \times \cdots \times \Omega_N)$, it follows that the tensor product H is complete. In particular, H and $\ell^2(\Omega_1 \times \cdots \times \Omega_N)$ are unitary equivalent as Hilbert spaces (Definition 7.209), ξ being a unitary transformation (thanks to Lemma 7.210). Observe that the estimate (7.7) yields

$$\|x_1 \otimes \cdots \otimes x_N\|_H < \infty$$

and, hence,

$$x_1 \otimes \cdots \otimes x_N \in H$$

for every $x_1 \in H_1, \ldots, x_N \in H_N$. Recall that $\{\delta_{(\mathrm{e}_1,\ldots,\mathrm{e}_N),\cdot}\}_{(\mathrm{e}_1,\ldots,\mathrm{e}_N)\in\Omega_1\times\cdots\times\Omega_N}$ is a Hilbert basis of $\ell^2(\Omega_1 \times \cdots \times \Omega_N)$. See Exercise 7.229. Finally, by unitarity of ξ^{-1} and the identity

$$\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N = \xi^{-1}(\delta_{(\mathrm{e}_1,\ldots,\mathrm{e}_N),\cdot})$$

one sees that $\{\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N\}_{(\mathrm{e}_1,\ldots,\mathrm{e}_N)\in\Omega_1\times\cdots\times\Omega_N}$ is a Hilbert basis of the tensor product H.

□

In the next proposition, we show that the Hilbert-Schmidt tensor product of a finite sequence of separable Hilbert spaces does not depend on the choice of Hilbert bases in its definition:

Proposition 7.259 *Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then,*

$$\mathrm{span}\{x_1 \otimes \cdots \otimes x_N \ :\ x_1 \in H_1, \ldots, x_N \in H_N\} ,$$

which is nothing else than the algebraic tensor product $H_1 \otimes \cdots \otimes H_N$, is a dense subset of $H_1 \otimes_2 \cdots \otimes_2 H_N$ and, for all $x_1, x_1' \in H_1, \ldots, x_N, x_N' \in H_N$, one has

$$\langle x_1' \otimes \cdots \otimes x_N', x_1 \otimes \cdots \otimes x_N \rangle_{H_1 \otimes_2 \cdots \otimes_2 H_N} = x_1 \otimes \cdots \otimes x_N (x_1', \ldots, x_N') \ .$$

In particular,

$$\|x_1 \otimes \cdots \otimes x_N\|_{H_1 \otimes_2 \cdots \otimes_2 H_N} = \|x_1\|^{(1)} \cdots \|x_N\|^{(N)} \ , \qquad x_1 \in H_1, \ldots, x_N \in H_N \ .$$

In fact, the space $H_1 \otimes_2 \cdots \otimes_2 H_N$ of bounded multiantilinear forms $\Phi : H_1 \times \cdots \times H_N \to \mathbb{K}$ as well as its norm $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ do not depend on the particular choice of Hilbert bases $\Omega_1, \ldots, \Omega_N$ in their definition.

Proof

1. Clearly,

$$\begin{aligned} \mathrm{span}\{\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N \ &: \ \mathrm{e}_1 \in \Omega_1, \ldots, \mathrm{e}_N \in \Omega_N\} \\ &\subseteq \mathrm{span}\{x_1 \otimes \cdots \otimes x_N \ : \ x_1 \in H_1, \ldots, x_N \in H_N\} \\ &\subseteq H_1 \otimes_2 \cdots \otimes_2 H_N \ . \end{aligned}$$

   Thus, for $\{\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N\}_{(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N}$ is a Hilbert basis of $H_1 \otimes_2 \cdots \otimes_2 H_N$,

$$\mathrm{span}\{x_1 \otimes \cdots \otimes x_N \ : \ x_1 \in H_1, \ldots, x_N \in H_N\}$$

   is a dense subspace of $H_1 \otimes_2 \cdots \otimes_2 H_N$, thanks to Lemma 7.218.
2. The polarization identity (Theorem 7.204) for the norm $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ implies that, for all $x_1, x_1' \in H_1, \ldots, x_N, x_N' \in H_N$,

$$\begin{aligned} &\langle x_1' \otimes \cdots \otimes x_N', x_1 \otimes \cdots \otimes x_N \rangle_{H_1 \otimes_2 \cdots \otimes_2 H_N} \\ &= \sum_{(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N} \langle x_1', \mathrm{e}_1 \rangle^{(1)} \langle \mathrm{e}_1, x_1 \rangle^{(1)} \cdots \langle x_N', \mathrm{e}_N \rangle^{(N)} \langle \mathrm{e}_N, x_N \rangle^{(N)} \ . \end{aligned}$$

   Note that the above sum absolutely converges because

$$\begin{aligned} &\sum_{(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N} |\langle x_1', \mathrm{e}_1 \rangle^{(1)} \langle \mathrm{e}_1, x_1 \rangle^{(1)} \cdots \langle x_N', \mathrm{e}_N \rangle^{(N)} \langle \mathrm{e}_N, x_N \rangle^{(N)} | \\ &\leq \left(\sum_{\mathrm{e}_1 \in \Omega_1} |\langle x_1', \mathrm{e}_1 \rangle^{(1)} \langle \mathrm{e}_1, x_1 \rangle^{(1)} | \right) \cdots \left(\sum_{\mathrm{e}_N \in \Omega_N} |\langle x_N', \mathrm{e}_N \rangle^{(N)} \langle \mathrm{e}_N, x_N \rangle^{(N)} | \right) \\ &\leq \|x_1'\| \, \|x_1\| \cdots \|x_N'\| \, \|x_N\| < \infty \ . \end{aligned}$$

   The last equality follows from Parseval's identity (Corollary 7.226) for H_m and the Cauchy-Schwarz inequality for $\ell^2(\Omega_m)$, where $m \in \{1, \ldots, N\}$.

3. In fact, because of the absolute convergence of all sums in the above bound, one even has that

$$\begin{aligned}
&\left\langle x_1' \otimes \cdots \otimes x_N', x_1 \otimes \cdots \otimes x_N \right\rangle_{H_1 \otimes_2 \cdots \otimes_2 H_N} \\
&= \sum_{e_1,\ldots,e_N \in \Omega_1 \times \cdots \times \Omega_N} \left\langle x_1', e_1 \right\rangle^{(1)} \langle e_1, x_1 \rangle^{(1)} \cdots \left\langle x_N', e_N \right\rangle^{(N)} \langle e_N, x_N \rangle^{(N)} . \\
&= \left(\sum_{e_1 \in \Omega_1} \left\langle x_1', e_1 \right\rangle^{(1)} \langle e_1, x_1 \rangle^{(1)} \right) \cdots \left(\sum_{e_N \in \Omega_N} \left\langle x_N', e_N \right\rangle^{(N)} \langle e_N, x_N \rangle^{(N)} \right)
\end{aligned}$$

for all $x_1, x_1' \in H_1, \ldots, x_N, x_N' \in H_N$.

4. By linearity and continuity (which follows from the Cauchy-Schwarz inequality) of the right argument of any scalar product together with Corollary 7.226, we arrive at

$$\begin{aligned}
&\left\langle x_1' \otimes \cdots \otimes x_N', x_1 \otimes \cdots \otimes x_N \right\rangle_{H_1 \otimes_2 \cdots \otimes_2 H_N} \qquad (7.8) \\
&= \left(\left\langle x_1', \sum_{e_1 \in \Omega_1} \langle e_1, x_1 \rangle^{(1)} e_1 \right\rangle^{(1)} \right) \cdots \left(\left\langle x_N', \sum_{e_1 \in \Omega_1} \langle e_N, x_N \rangle^{(N)} e_N \right\rangle^{(N)} \right) \\
&= \left\langle x_1', x_1 \right\rangle^{(1)} \cdots \left\langle x_N', x_N \right\rangle^{(N)} = x_1 \otimes \cdots \otimes x_N (x_1', \ldots, x_N')
\end{aligned}$$

for all $x_1, x_1' \in H_1, \ldots, x_N, x_N' \in H_N$.

5. Note, in particular, that this special scalar product does not depend on the choice of Hilbert bases $\Omega_1, \ldots, \Omega_N$ in the definition of the norm $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$. Take $\Phi \in H_1 \otimes_2 \cdots \otimes_2 H_N$. By density, there is a sequence

$$\Phi_n \in \operatorname{span}\{x_1 \otimes \cdots \otimes x_N \ : \ x_1 \in H_1, \ldots, x_N \in H_N\} , \qquad n \in \mathbb{N} ,$$

converging to Φ, with respect to the norm $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$. In particular, by the continuity of norms (which follows from their subadditivity),

$$\|\Phi\|_{H_1 \otimes_2 \cdots \otimes_2 H_N} = \lim_{n \to \infty} \|\Phi_n\|_{H_1 \otimes_2 \cdots \otimes_2 H_N} .$$

But, by (7.8), $\|\Phi_n\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$, $n \in \mathbb{N}$, do not depend on the choice of Hilbert bases defining the norm $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$. Hence, for any $\Phi \in H_1 \otimes_2 \cdots \otimes_2 H_N$, $\|\Phi\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ is independent of the choice of Hilbert bases.

6. Finally, a similar argument shows that any tensor product Hilbert space $H_1 \otimes_2 \cdots \otimes_2 H_N$, defined with respect to some choice of Hilbert bases, is a subspace of any tensor product Hilbert space $H_1 \otimes_2 \cdots \otimes_2 H_N$, defined for the same Hilbert spaces, but for any other choices of Hilbert bases. Thus, any choice of Hilbert bases yields the same space, independent of the choice the Hilbert bases in their definition.

□

Corollary 7.260 *Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Then $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ is a reasonable crossnorm (Definition 7.44) in the algebraic tensor product $H_1 \otimes \cdots \otimes H_N$. The Hilbert-Schmidt tensor product $H_1 \otimes_2 \cdots \otimes_2 H_N$ is a completion of the algebraic one with respect to this norm.*

Proof The fact that $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ is a crossnorm in $H_1 \otimes \cdots \otimes H_N$ directly follows from the last proposition. For all $\varphi_1 \in H_1^{\mathrm{td}}, \ldots, \varphi_N \in H_N^{\mathrm{td}}$, which we identify with vectors $x_1 \in H_1, \ldots, x_N \in H_N$ via the Riesz-Fréchet theorem (Theorem 7.214), the element $\varphi_1 \otimes \cdots \otimes \varphi_N$ seen as a linear form on $H_1 \otimes \cdots \otimes H_N$ acts as

$$\varphi_1 \otimes \cdots \otimes \varphi_N(x) = \langle x_1 \otimes \cdots \otimes x_N, x \rangle_{H_1 \otimes_2 \cdots \otimes_2 H_N} , \qquad x \in H_1 \otimes \cdots \otimes H_N ,$$

thanks again to the last proposition. By the Cauchy-Schwarz inequality, it follows that

$$\|\varphi_1 \otimes \cdots \otimes \varphi_N\|_{\mathrm{op}} = \|x_1\|^{(1)} \cdots \|x_N\|^{(N)} = \|\varphi_1\|_{\mathrm{op}} \cdots \|\varphi_N\|_{\mathrm{op}} .$$

As a consequence, $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ is a reasonable crossnorm. See Definition 7.44. Note from Proposition 7.259 that the Hilbert-Schmidt tensor product $H_1 \otimes_2 \cdots \otimes_2 H_N$ is a completion of the algebraic one, with respect to this norm. □

The following corollary says that the Hilbert-Schmidt tensor product is an associative operation:

Corollary 7.261 (Associativity of Hilbert Tensor Products) *Let H_1, H_2, H_3 be three separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. The canonical inclusions*

$$H_1 \otimes H_2 \otimes H_3 \to (H_1 \otimes H_2) \otimes H_3 ,$$
$$H_1 \otimes H_2 \otimes H_3 \to H_1 \otimes (H_2 \otimes H_3)$$

(see Corollary 7.13) uniquely extend to unitary transformations

$$H_1 \otimes_2 H_2 \otimes_2 H_3 \to (H_1 \otimes_2 H_2) \otimes_2 H_3 ,$$
$$H_1 \otimes_2 H_2 \otimes_2 H_3 \to H_1 \otimes_2 (H_2 \otimes_2 H_3) .$$

In other words, the Hilbert spaces $(H_1 \otimes_2 H_2) \otimes_2 H_3$, $H_1 \otimes_2 (H_2 \otimes_2 H_3)$, and $H_1 \otimes_2 H_2 \otimes_2 H_3$ are canonically identified with each other, and the (Hilbert-Schmidt) tensor product operation $\otimes_2$ can be seen as associative.

Proof Exercise. □

Given $N \in \mathbb{N}$ Hilbert spaces $H_1, \ldots, H_N$ over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ and linear transformations $A_1 \in \mathcal{L}(H_1), \ldots, A_N \in \mathcal{L}(H_N)$, recall that there is a unique linear transformation

$$A_1 \otimes \cdots \otimes A_N \in \mathcal{L}(H_1 \otimes \cdots \otimes H_N)$$

on (the algebraic tensor product) $H_1 \otimes \cdots \otimes H_N$, such that

$$A_1 \otimes \cdots \otimes A_N(x_1 \otimes \cdots \otimes x_N) = A_1(x_1) \otimes \cdots \otimes A_N(x_N) , \quad x_1 \in H_1, \ldots, x_N \in H_N .$$

See Corollary 7.11. We show in the sequel that if $A_1 \in \mathcal{B}(H_1), \ldots, A_N \in \mathcal{B}(H_N)$, i.e., these operators are bounded, then $A_1 \otimes \cdots \otimes A_N$ is bounded with respect to the norm $\|\cdot\|_{H_1 \otimes_2 \cdots \otimes_2 H_N}$ and has thus a unique extension as a bounded linear operator on $H_1 \otimes_2 \cdots \otimes_2 H_N$, which is again denoted by $A_1 \otimes \cdots \otimes A_N$.

Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ with Hilbert bases $\Omega_1 \subseteq H_1, \ldots, \Omega_N \subseteq H_N$, respectively. Note that the tensor product $A_1 \otimes \cdots \otimes A_N$ of N bounded operators $A_1 \in \mathcal{B}(H_1), \ldots, A_N \in \mathcal{B}(H_N)$ defines a unique linear mapping from

$$\operatorname{span}\{\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N : (\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N\}$$

to $H_1 \otimes_2 \cdots \otimes_2 H_N$ such that

$$A_1 \otimes \cdots \otimes A_N(\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N) = A_1(\mathrm{e}_1) \otimes \cdots \otimes A_N(\mathrm{e}_N)$$

for all $(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N$. We show below that the operator norm of this linear mapping is bounded.

Lemma 7.262 *Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ with Hilbert bases $\Omega_1 \subseteq H_1, \ldots, \Omega_N \subseteq H_N$, respectively. Fix bounded operators $A_1 \in \mathcal{B}(H_1), \ldots, A_N \in \mathcal{B}(H_N)$. Then, there is a unique $A_1 \otimes \cdots \otimes A_N \in \mathcal{B}(H_1 \otimes_2 \cdots \otimes_2 H_N)$ such that*

$$A_1 \otimes \cdots \otimes A_N(\mathrm{e}_1 \otimes \cdots \otimes \mathrm{e}_N) = A_1(\mathrm{e}_1) \otimes \cdots \otimes A_N(\mathrm{e}_N)$$

for all $(\mathrm{e}_1, \ldots, \mathrm{e}_N) \in \Omega_1 \times \cdots \times \Omega_N$. Additionally,

$$\|A_1 \otimes \cdots \otimes A_N\|_{\mathrm{op}} \leq \|A_1\|_{\mathrm{op}} \cdots \|A_N\|_{\mathrm{op}} .$$

Proof Note that it is enough to consider the special case $\mathrm{id}_{H_1} \otimes \cdots \otimes A_j \otimes \cdots \otimes \mathrm{id}_{H_N}$ and write

$$A_1 \otimes \cdots \otimes A_N = \prod_{j=1}^{N} \mathrm{id}_{H_1} \otimes \cdots \otimes A_j \otimes \cdots \otimes \mathrm{id}_{H_N} .$$

For simplicity and without loss of generality, let $N = 2$. Take any

$$x = \sum_{(\mathrm{e}_1, \mathrm{e}_2) \in I \times J} c_{\mathrm{e}_1, \mathrm{e}_2} \mathrm{e}_1 \otimes \mathrm{e}_2 \in \operatorname{span}\{\mathrm{e}_1 \otimes \mathrm{e}_2 : (\mathrm{e}_1, \mathrm{e}_2) \in \Omega_1 \times \Omega_2\} ,$$

where $I \subseteq \Omega_1$ and $J \subseteq \Omega_2$ are arbitrary finite subsets, while $c_{e_1,e_2} \in \mathbb{K}$ are arbitrary constants for any $(e_1, e_2) \in I \times J$. Then,

$$(A_1 \otimes \mathrm{id}_{H_2})(x) = \sum_{(e_1,e_2)\in I\times J} c_{e_1,e_2} (A_1(e_1) \otimes e_2)$$
$$= \sum_{e_2\in J} A_1 \left(\sum_{e_1\in J} c_{e_1,e_2} e_1 \right) \otimes e_2 \ .$$

Since, for any $e_2, e_2' \in J$ so that $e_2 \neq e_2'$,

$$\left[A_1 \left(\sum_{e_1\in J} c_{e_1,e_2} e_1 \right) \otimes e_2 \right] \perp \left[A_1 \left(\sum_{e_1\in J} c_{e_1,e_2} e_1 \right) \otimes e_2' \right] ,$$

by the Pythagorean theorem and Corollary 7.226,

$$\left\| (A_1 \otimes \mathrm{id}_{H_2})(x) \right\|^2 = \sum_{e_2\in J} \left\| A_1 \left(\sum_{e_1\in J} c_{e_1,e_2} e_1 \right) \otimes e_2 \right\|^2$$
$$= \sum_{e_2\in J} \left\| A_1 \left(\sum_{e_1\in J} c_{e_1,e_2} e_1 \right) \right\|^2$$
$$\leq \|A_1\|_{\mathrm{op}}^2 \sum_{e_2\in J} \sum_{e_1\in J} |c_{e_1,e_2}|^2 = \|A_1\|_{\mathrm{op}}^2 \, \|x\|_{H_1 \otimes_2 H_2}^2 \ .$$

In a similar way, one shows that $\left\| (\mathrm{id}_{H_1} \otimes A_2)(x) \right\| \leq \|A_2\|_{\mathrm{op}} \, \|x\|_{H_1 \otimes_2 H_2}$. From this type of estimate generalized to any $N \geq 2$, we conclude the lemma. □

Exercise 7.263 Under the assumptions of Lemma 7.262, show that, for all $x_1 \in H_1, \ldots, x_N \in H_N$,

$$A_1 \otimes \cdots \otimes A_N (x_1 \otimes \cdots \otimes x_N) = A_1(x_1) \otimes \cdots \otimes A_N(x_N) \ ,$$

i.e., $A_1 \otimes \cdots \otimes A_N \in \mathcal{B}(H_1 \otimes_2 \cdots \otimes_2 H_N)$ is an extension of the algebraic tensor product of the linear operators $A_1 \in \mathcal{B}(H_1), \ldots, A_N \in \mathcal{B}(H_N)$.

From this last result, one finally deduces that the tensor product $A_1 \otimes \cdots \otimes A_N \in \mathcal{B}(H_1 \otimes_2 \cdots \otimes_2 H_N)$ of bounded operators of the last lemma does not depend on the choice of Hilbert bases:

Corollary 7.264 *Let $H_1, \ldots, H_N$ be $N \in \mathbb{N}$ separable Hilbert spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Fix bounded operators $A_1 \in \mathcal{B}(H_1), \ldots, A_N \in \mathcal{B}(H_N)$. Then, $A_1 \otimes \cdots \otimes A_N \in \mathcal{B}(H_1 \otimes_2 \cdots \otimes_2 H_N)$ is the unique bounded operator on $H_1 \otimes_2 \cdots \otimes_2 H_N$ such that*

$$A_1 \otimes \cdots \otimes A_N(x_1 \otimes \cdots \otimes x_N) = A_1(x_1) \otimes \cdots \otimes A_N(x_N)$$

for all $x_1 \in H_1, \ldots, x_N \in H_N$.

In fact, as discussed above, $A_1 \otimes \cdots \otimes A_N \in \mathcal{B}(H_1 \otimes_2 \cdots \otimes_2 H_N)$ is nothing else than the unique continuous extension of the corresponding algebraic tensor product of linear operators, from the algebraic tensor product $H_1 \otimes \cdots \otimes H_N$ to the Hilbert-Schmidt tensor product $H_1 \otimes_2 \cdots \otimes_2 H_N \supseteq H_1 \otimes \cdots \otimes H_N$.

7.4 Riesz Spaces

7.4.1 Basic Definitions and Properties

Recall that infima and suprema in an ordered vector space are unique, when they exist. The partially ordered set $(P, \succeq)$ is called a "lattice" if any two elements $p, p' \in P$ have an infimum and a supremum, denoted, respectively, by $p \wedge p'$ and $p \vee p'$.

Exercise 7.265 Let $(P, \succeq)$ be any partially ordered set which is a lattice. Show that any *finite* subset of P has an infimum and a supremum. Prove, moreover, that for any finite sequence $p_1, p_2, \ldots, p_N \in P$, $N \in \mathbb{N}$, one has

$$p_1 \wedge (p_2 \wedge (\cdots \wedge p_N) \cdots) = \inf\{p_1, \ldots, p_N\} ,$$

$$p_1 \vee (p_2 \vee (\cdots \vee p_N) \cdots) = \sup\{p_1, \ldots, p_N\} .$$

In particular, the lattice operations $\wedge, \vee$ of $(P, \succeq)$ are commutative and associative.

Observe also that, in lattices, a minimal (maximal) upper (lower) bound of a subset is automatically the supremum (infimum), i.e., the least (largest) upper (lower) bound of this subset:

Lemma 7.266 *Let $(P, \succeq)$ be any lattice and $\Omega \subseteq P$ any upper (lower) bounded subset. $p \in P$ is the supremum (infimum) of Ω iff, for any upper (lower) bound $p' \in P$ of Ω, $p \succeq p'$ ($p' \succeq p$) implies that $p' = p$.*

Proof We prove only the upper bound case, the proof of the lower bound one being very similar. Let $\Omega \subseteq P$ be a upper bounded subset. If $p \in P$ is the supremum, i.e., the least upper bound, of Ω , then, trivially, for any upper bound $p' \in P$ of Ω, $p \succeq p'$ implies that $p' = p$, i.e., p is a minimal upper bound of Ω. Assume now, conversely, that p is a minimal upper bound of Ω. Suppose, by contradiction, that it is not the least upper bound, i.e., there is some upper bound $p' \in P$ for Ω, with $p' \not\succeq p$. Then $p' \wedge p$ would be an upper bound for Ω satisfying

$$p \succeq p' \wedge p \neq p \ .$$

This last equation shows that p would not be minimal if p is not the least upper bound. □

Definition 7.267 (Riesz Space)

(i) A "Riesz space" or "vector lattice" V is, by definition, an ordered vector space over $\mathbb{R}$ which is a lattice, in the usual sense of partially ordered sets. In particular, Riesz spaces are directed vector spaces (see Definition 1.13).
(ii) A "Riesz subspace" $W \subseteq V$ of a Riesz space V is a vector subspace which is a lattice, whose lattice operations ($\wedge$,$\vee$) coincide with the restriction of the lattice operations of V to W. In particular, order-closed subspaces of Riesz spaces are Riesz subspaces.
(iii) Riesz subspaces of V that are order-full (Definition 1.4) are called "Riesz ideals." In particular, Riesz ideals are order ideals of V.

$(\mathbb{R}, \geq)$ is a simple example of a Riesz space: For any $x_1, x_2 \in \mathbb{R}$,

$$x_1 \wedge x_2 = \min\{x_1, x_2\} \ , \qquad x_1 \vee x_2 = \max\{x_1, x_2\} \ .$$

Another example is the space $\mathcal{F}(\Omega; \mathbb{R})$ of real-valued functions, where Ω is any nonempty set. More generally, if V is a Riesz space, then $\mathcal{F}(\Omega; V)$ is again a Riesz space. In this case, for all $f, f' \in \mathcal{F}(\Omega; V)$, one has

$$[f \wedge f'](p) = f(p) \wedge f'(p) \ , \qquad [f \vee f'](p) = f(p) \vee f'(p) \ , \qquad p \in \Omega \ .$$

Recall that this (Riesz) space is order-complete iff V is order-complete (Definition 1.20) and that the same is true for the Archimedean property. See Example 1.3. In particular, $\mathcal{F}(\Omega; \mathbb{R})$ is an Archimedean order-complete Riesz space. Remind that a preordered vector space V is "Archimedean" if $0 \succeq \tilde{v}$ whenever $v \succeq \alpha\tilde{v}$ for all $\alpha \in \mathbb{R}_0^+$ and some $v \in V$; see Definition 1.2.

It is very important to notice that Riesz spaces may have vector subspaces that are themselves new Riesz spaces, but the lattice operations of the subspace do not correspond to the restriction to the subspace of the operations of the big space. In other words, subspaces of Riesz spaces are not necessarily Riesz subspaces, even if these are lattices:

Example 7.268 Consider the Riesz space $V \doteq \mathcal{F}([0, 1]; \mathbb{R})$. Let (the functions) $f, f' \in V$ be defined by

$$f(s) = s \ , \ \ f'(s) = 1 \ , \qquad s \in [0, 1] \ .$$

$W \doteq \text{span}(\{f, f'\}) \subseteq V$ is a clearly a vector subspace of V, which is a lattice. However, it is not a Riesz subspace: In fact, note that the supremum of $\{f, f' - f\}$ in W is f', but

$$\max\{f(1/2), f'(1/2) - f(1/2)\} = 1/2 < f'(1/2) = 1 .$$

The following important example is prototypical for Riesz spaces:

Exercise 7.269 Let M be any metric space. Prove that $C(M; \mathbb{R})$ is a Riesz subspace of $\mathcal{F}(M; \mathbb{R})$.

Hint: Prove first that, for any $f, g \in C(M; \mathbb{R})$, $\min\{f(\cdot), g(\cdot)\}$ and $\max\{f(\cdot), g(\cdot)\}$ are continuous mappings $M \to \mathbb{R}$.

Riesz spaces of this kind are Archimedean (as subspaces of Archimedean spaces), but generally *not* σ-order-complete.

Definition 7.270 (Function Spaces) We say that a real vector space V is a "function space" whenever it is a Riesz subspace of $\mathcal{F}(\Omega; \mathbb{R})$ for some nonempty set Ω.

In particular, function spaces are Archimedean. By the last exercise, observe that $C(M; \mathbb{R})$, where M is any metric space, is a function space in the sense of the above definition. Note that the space V of Example 7.268 is indeed a subspace of some $\mathcal{F}(\Omega; \mathbb{R})$, i.e., its elements are real-valued functions, but is not a function space in the sense of the definition.

Theorem 7.271 *For any Archimedean Riesz space V, there are a nonempty set Ω and an injective bipositive*[11] *mapping $V \to \mathcal{F}(\Omega; \mathbb{R})$. In particular, any Archimedean Riesz space is equivalent to a function space.*

This abstract result on representation of Archimedean Riesz spaces is well-known and will not be proven here. It is only given here for completeness. For more details, see, for instance, [5, Chapter 7].

For the next exercise on elementary properties of lattice operations ($\wedge$,$\vee$), recall that if $(v_i)_{i \in I}$ is a monotonically increasing net in a preordered vector space V and $v \in V$ is a supremum for (the upward directed subset) $\{v_i : i \in I\} \subseteq V$, then we write $v_i \uparrow v$ and v is called a "(monotone) order limit" of the monotone net $(v_i)_{i \in I}$. Mutatis mutandis for monotonically decreasing nets; see Definition 1.17.

Exercise 7.272 Let V be any Riesz space. Prove the following assertions:

(i) For all $v, v' \in V$,

$$-(v \vee v') = (-v) \wedge (-v') \text{ and } -(v \wedge v') = (-v) \vee (-v') .$$

[11] That is, it reflects positivity; see Definition 1.7.

(ii) For all $v, v', v'' \in V$,

$$v'' + (v \vee v') = (v'' + v) \vee (v'' + v') \,,$$
$$v'' + (v \wedge v') = (v'' + v) \wedge (v'' + v') \,.$$

(iii) For all $\alpha \geq 0$ and $v, v' \in V$,

$$\alpha(v \vee v') = (\alpha v) \vee (\alpha v') \text{ and } \alpha(v \wedge v') = (\alpha v) \wedge (\alpha v') \,.$$

(iv) For all $v, v' \in V$,

$$v + v' = v \vee v' + v \wedge v'.$$

(v) For all $v, v', v'' \in V^+$,

$$v \wedge v' + v \wedge v'' \succeq v \wedge (v' + v'') \,.$$

Hint: Prove first that

$$v' \succeq v \wedge (v' + v'') - v \wedge v'' \,.$$

(vi) For any decreasing net $(v_i)_{i \in I}$, $v_i \downarrow v$, and any $v' \in V$,

$$(v_i \wedge v') \downarrow (v \wedge v') \text{ and } (v_i \vee v') \downarrow (v \vee v') \,.$$

(vii) For any increasing net $(v_i)_{i \in I}$, $v_i \uparrow v$, and any $v' \in V$,

$$(v_i \wedge v') \uparrow (v \wedge v') \text{ and } (v_i \vee v') \uparrow (v \vee v') \,.$$

The following characterization of ordered vector spaces that are Riesz spaces is frequently useful:

Exercise 7.273 Let V be any real ordered vector space. Show that V is a Riesz space iff, for all $v \in V$, the subset $\{0, v\} \subseteq V$ has a supremum.

One simple though important property of Riesz spaces is the following assertion:

Lemma 7.274 (Riesz Decomposition Property) *Let V be any Riesz space. If*

$$v, v_1, \ldots, v_n \in V^+ \,, \qquad n \in \mathbb{N} \,,$$

are positive vectors for which

$$v_1 + \cdots + v_n \succeq v,$$

then there are positive vectors

$$v'_1, \ldots, v'_n \in V^+$$

such that $v_k \succeq v'_k$, $k = 1, \ldots, n$, *and*

$$v'_1 + \cdots + v'_n = v \ .$$

Proof Note that the property is trivial for $n = 1$. We prove the general case by induction in $n \in \mathbb{N}$. Suppose that the property is true for some $n \in \mathbb{N}$ and take arbitrary $v, v_1, \ldots, v_{n+1} \in V^+$ such that

$$v_1 + \cdots + v_{n+1} \succeq v \ .$$

Define $v'_{n+1} \doteq v \wedge v_{n+1}$. Clearly, $v_{n+1} \succeq v'_{n+1}$ and $(v - v'_{n+1}) \in V^+$. Observe from Exercise 7.272 that

$$\begin{aligned} v - v'_{n+1} &= v - v \wedge v_{n+1} \\ &= v + (-v) \vee (-v_{n+1}) \\ &= 0 \vee (v - v_{n+1}) \ . \end{aligned}$$

Thus, by positivity of $v_1, \ldots, v_n \in V^+$,

$$v_1 + \cdots + v_n \succeq v - v'_{n+1} \ .$$

Hence, by the induction hypothesis, there are $v'_1, \ldots, v'_n \in V^+$ such that $v_k \succeq v'_k$ for every $k \in \{1, \ldots, n\}$ and

$$v'_1 + \cdots + v'_n = v - v'_{n+1} \ .$$

□

In Riesz spaces, the existence of suprema of subsets that are order-bounded from above can always be seen as the existence of order limits of bounded monotone nets (Definition 1.17):

Lemma 7.275 *Let V be any Riesz space. If there is a nonempty subset $\Omega \subseteq V$ which is bounded from above but has no supremum, then there is a monotonically increasing net in V, which is also bounded from above but has no order limit.*

Proof Take an arbitrary nonempty subset $\Omega \subseteq V$ that is bounded from above. Let $\mathcal{P}_f(\Omega)$ be the collection of all *finite* subsets of Ω. It is partially ordered by the inclusion relation for sets. It is also a directed set with respect to this order relation (Definition 1.13). In particular, it may be used as the index set of a net (Definition 1.15). For all $\tilde{\Omega} \in \mathcal{P}_f(\Omega)$, define

$$v_{\tilde{\Omega}} \doteq \sup \tilde{\Omega} \in V \ .$$

Note that the net $(v_{\tilde{\Omega}})_{\tilde{\Omega}\in\mathcal{P}_f(\Omega)}$ is monotonically increasing and that, for any upper bound v of Ω, $v \succeq v_{\tilde{\Omega}}$ for every $\tilde{\Omega} \in \mathcal{P}_f(\Omega)$. In particular, the net $(v_{\tilde{\Omega}})_{\tilde{\Omega}\in\mathcal{P}_f(\Omega)}$ is bounded from above. Assume now that Ω has no supremum. Suppose, by contradiction, that $\{v_{\tilde{\Omega}} : \tilde{\Omega} \in \mathcal{P}_f(\Omega)\}$ has a supremum, that is,

$$v_{\tilde{\Omega}} \uparrow \sup\{v_{\tilde{\Omega}} : \tilde{\Omega} \in \mathcal{P}_f(\Omega)\} .$$

For any upper bound $v \in V$ of Ω, one has $v \succeq v_{\tilde{\Omega}}$ for every $\tilde{\Omega} \in \mathcal{P}_f(\Omega)$ (see Lemma 7.266) and, hence,

$$v \succeq \sup\{v_{\tilde{\Omega}} : \tilde{\Omega} \in \mathcal{P}_f(\Omega)\} .$$

Therefore, as $\sup\{v_{\tilde{\Omega}} : \tilde{\Omega} \in \mathcal{P}_f(\Omega)\}$ is an upper bound for Ω,

$$\sup \Omega = \sup\{v_{\tilde{\Omega}} : \tilde{\Omega} \in \mathcal{P}_f(\Omega)\} .$$

This contradicts the non-existence of a supremum for Ω. Therefore, $(v_{\tilde{\Omega}})_{\tilde{\Omega}\in\mathcal{P}_f(\Omega)}$ has no order limit. □

From the last lemma, we arrive at the following equivalent characterization of order completeness in Riesz spaces, via monotone nets:

Corollary 7.276 *Let V be any Riesz space. It is order-complete (σ-order-complete) iff all its monotonically increasing bounded nets (sequences) of positive elements have order limits.*

Proof Exercise. □

A similar statement can be proven for the order closedness of Riesz subspaces of Riesz spaces. Recall that a vector subspace $W \subseteq V$ of a preordered vector space V is "order-closed" if the suprema in V of subsets of W are again elements of W. See Definition 1.20.

Lemma 7.277 *Let V be any Riesz space. A Riesz subspace $W \subseteq V$ is order-closed (σ-order-closed) iff, for any increasing monotone net $(w_i)_{i\in I}$ (sequence $(w_n)_{n\in\mathbb{N}}$) of positive elements of W, $w_i \uparrow v$ ($w_n \uparrow v$) (i.e., the net (sequence) has an order limit in V), one has $v \in W$.*

Proof We only consider the order-closedness, the proof for the σ-order-closedness being an obvious adaptation of this case. Assume that W is an order-closed vector subspace. In particular, it is a Riesz subspace. Then, clearly, for any increasing monotone net $(w_i)_{i\in I}$ in W, $w_i \uparrow v$, one has $v \in W$. Assume now that this property of nets is valid. Let $\Omega \subseteq W$ be any nonempty subset which is bounded from above in V. Consider the increasing monotone net $(w_{\tilde{\Omega}})_{\tilde{\Omega}\in\mathcal{P}_f(\Omega)}$ in V defined in the same way as in the proof of the last lemma. Note that $(w_{\tilde{\Omega}})_{\tilde{\Omega}\in\mathcal{P}_f(\Omega)}$ is a net in W, as W is a Riesz subspace. If Ω has a supremum in V, then $w_{\tilde{\Omega}} \uparrow \sup \Omega$. By the assumption on monotone nets, $\sup \Omega \in W$ and W is thus order-closed. To finish the proof, note that the assumption on positive nets in the lemma implies the same property for general nets, because W is a directed vector space. □

7.4.2 The Absolute Value in General Vector Lattices

In Riesz spaces, a notion of "absolute value," and related notions, can naturally be introduced:

Definition 7.278 (Absolute Value in a Riesz Space) Let V be any Riesz space.

(i) For all $v \in V$, we define its "absolute value" by

$$|v| \doteq v \vee (-v) \ .$$

In particular, $|-v| = |v|$ for all $v \in V$ and $|v| = v$ for every positive $v \in V^+$.

(ii) For any $v \in V$,

$$v^+ \doteq \frac{1}{2}(|v| + v) \ , \qquad v^- \doteq \frac{1}{2}(|v| - v) \ ,$$

are, respectively, called the "positive part" and "negative part" of v. In particular,

$$|v| = v^+ + v^- \text{ and } v = v^+ - v^- .$$

(iii) We say that the vectors $v, v' \in V$ are "disjoint," whenever $|v| \wedge |v'| = 0$. In this case, we use the notation $v \perp v'$. Finally, for any nonempty subset $\Omega \subseteq V$, we define its "order-disjoint complement" by

$$\Omega^{\mathrm{d}} \doteq \{v' \in V \ : \ v \perp v' \text{ for all } v \in \Omega\} \subseteq V \ .$$

Observe that, as is usual for other notions of disjoint complement, one has (a) $\Omega \subseteq \Omega^{\mathrm{dd}}$, (b) $\Omega \cap \Omega^{\mathrm{d}} \subseteq \{0\}$, and (c) $\Omega^{\mathrm{d}} \subseteq \Omega'^{\mathrm{d}}$, whenever $\Omega' \subseteq \Omega$.

Lemma 7.279 *Let V be any Riesz space. For any nonempty $\Omega \subseteq V$, one has $(\Omega^{\mathrm{d}})^{\mathrm{dd}} = \Omega^{\mathrm{d}}$.*

Proof On the one hand, by the first basic property of the disjoint complement given above, one has $\Omega^{\mathrm{d}} \subseteq (\Omega^{\mathrm{d}})^{\mathrm{dd}}$. On the other hand, by the first and third properties,

$$(\Omega^{\mathrm{d}})^{\mathrm{dd}} = (\Omega^{\mathrm{dd}})^{\mathrm{d}} \subseteq \Omega^{\mathrm{d}} \ .$$

□

Exercise 7.280 Let V be a Riesz space. Show (directly by the axioms of Riesz spaces) that:

(i) For $v \in V$, $|v| = 0$ iff $v = 0$, and $v \in V^+$ iff $|v| = v$.
(ii) For all $v \in V$, $v^+, v^- \succeq 0$ and $v^+ \perp v^-$.
(iii) For all $v \in V$ and $v', v'' \succeq 0$, $v' \perp v''$, such that $v = v' - v''$, one has $v^+ = v'$ and $v^- = v''$.
(iv) For all $v \in V$,

$$v^+ = v \vee 0 \qquad \text{and} \qquad v^- = (-v) \vee 0.$$

(v) For all $v, v' \in V$,

$$v \wedge v' = \frac{1}{2}(v + v' - |v - v'|) \quad \text{and} \quad v \vee v' = \frac{1}{2}(v + v' + |v - v'|)\ .$$

(vi) For all $v, v' \in V$,

$$|v + v'| \in [||v| - |v'||, |v| + |v'|]\ .$$

Hint: Prove that

$$|v| + |v'| \succeq |v + v'|$$

and deduce from this, by a convenient choice of v and v', that

$$|v + v'| \succeq |v| - |v'|\ .$$

(vii) For all $v, v' \in V$,

$$|v'| \succeq ||v + v'| - |v||\ .$$

(viii) For all $v, v' \in V$,

$$|v - v'| \succeq |v^+ - v'^+| \text{ and } |v - v'| \succeq |v^- - v'^-|\ .$$

Notice that the above properties are trivial for $(\mathbb{R}, \geq)$ and any function space (Definition 7.270). From Theorem 7.271, they must hold true in any Archimedean Riesz space.

From the properties of the absolute value, one gets the following characterization of the Archimedean property of Riesz spaces:

Exercise 7.281 Let V be any Riesz space. Show that V is Archimedean iff, for any $v \in V$ and $\tilde{v} \in V^+$, $v \succeq \alpha\tilde{v}$ for all $\alpha \geq 0$ only if $\tilde{v} = 0$.

In other words, in a Riesz space, the condition defining the Archimedean property (see Definition 1.2) has only to be checked for *positive* (instead of all non-negative) $\tilde{v} \neq 0$.

Definition 7.282 (Solid Sets) Let V be any Riesz space. We say that the subset $\Omega \subseteq V$ is "solid" in V if, for all $v' \in V$ and $v \in \Omega$, one has $|v| \succeq |v'|$ only if $v' \in \Omega$. In particular, in a solid subset $\Omega \subseteq V$, $v \in \Omega$ implies $|v| \in \Omega$.

The following lemma gives an alternative definition for Riesz ideals (Definition 7.267 (iii)), via the solidness:

Lemma 7.283 *Let V be any Riesz space. The solid vector subspaces of V are exactly the Riesz ideals of V.*

Proof

1. Let $W \subseteq V$ be any solid vector subspace of V. Then, by Exercise 7.280 (v), for any $w \in W$, $|w| \in W$. Thus, for all $w, w' \in W$,

$$w \wedge w' = \frac{1}{2}(w + w' - |w - w'|) \in W ,$$

$$w \vee w' = \frac{1}{2}(w + w' + |w - w'|) \in W .$$

This shows that W is a Riesz subspace; see Definition 7.267 (i).

2. As W is solid, for all $w \in W^+$, one has that

$$\{v \in V \ : \ w \succeq v \succeq 0\} \subseteq W .$$

But, as W is a vector space, this implies that, for all $w, w' \in W$,

$$\{v \in V \ : \ w \succeq v \succeq w'\} \subseteq W .$$

In other words, W is an order ideal. See Definition 1.4.

3. Let now $W \subseteq V$ be any Riesz ideal of V. As W is (by definition) a Riesz subspace, for all $w \in W$, one has $|w| \in W$. For all $w \in W$ and $v \in V$ so that $|w| \succeq |v|$, one has $|w| \succeq v^+ \succeq 0$ and $|w| \succeq v^- \succeq 0$. As W is (by definition) an order ideal, it follows that $v^+, v^- \in W$ and, hence, $v \in W$. In other words, W is a solid vector subspace of V. See Definition 7.282.

□

Among other things, the absolute value allows us to extend the notion of order convergence to non-monotone nets:

Definition 7.284 (Order Convergence) Let V be any Riesz space. We say that the net $(v_i)_{i \in I}$ is "order-convergent" to $v \in V$ if there is a second net $(v'_i)_{i \in I}$ with the same index set, such that $v'_i \downarrow 0$ and $v'_i \succeq |v - v_i|$ for every $i \in I$. In this case, v is called the "order limit" of the net $(v_i)_{i \in I}$, and we use the notation $v_i \xrightarrow{o} v$.

This notion of convergence can be understood as an abstract version of the pointwise convergence for functions and is reminiscent of "Lebesgue's dominated convergence." See Exercise 7.286 and Proposition 7.313 below.

It is important to notice that if $v_i \stackrel{o}{\to} v$ in a Riesz subspace V' of the Riesz space V, it does not follow in general that $v_i \stackrel{o}{\to} v$ in V as well.

Exercise 7.285 Prove that order limits in Riesz spaces are unique.

Exercise 7.286 Let $V \subseteq \mathcal{F}(\Omega; \mathcal{R})$ be a function space in the sense of Definition 7.270. i.e., a Riesz subspace of the Riesz space $\mathcal{F}(\Omega; \mathbb{R})$. Show that a *bounded* net $(v_i)_{i\in I}$ in V converges pointwise to $v \in V$ (i.e., $\lim_{i\in I} v_i(p) = v(p)$) only if it is order-convergent to the same v. Conversely, show that if V is order-closed (σ-order-closed), then any net $(v_i)_{i\in I}$ (sequence $(v_n)_{n\in\mathbb{N}}$) in V converges pointwise to $v \in V$ if it is order-convergent to the same v.

If the Riesz space is order-complete (σ-order-complete), then any order limit of nets can be seen as a monotone limit via the "limit inferior" or "limit superior," exactly as is in the case of real numbers:

Exercise 7.287 Let V be any order-complete (σ-order-complete) Riesz space and let $(v_i)_{i\in I}$ $((v_n)_{n\in\mathbb{N}})$ be a bounded net (sequence) in V that is order-convergent to some $v \in V$. Show that

$$\begin{aligned} v &= \liminf_{i\in I} v_i \doteq \sup\left\{\inf\{v_j \ : \ j \succeq i\} \ : \ i \in I\right\} \\ &= \limsup_{i\in I} v_i \doteq \inf\left\{\sup\{v_j \ : \ j \succeq i\} \ : \ i \in I\right\} . \end{aligned}$$

(Mutatis mutandis for sequences $(v_n)_{n\in\mathbb{N}}$.) Observe that the net (sequence) is bounded, being order-convergent. Thus, as V is order-complete (σ-order-complete), $\liminf_{i\in I} v_i$ and $\limsup_{i\in I} v_i$ are well-defined.

From the last exercise, we obtain that order-closed (σ-order-closed) Riesz subspaces of order-complete (σ-order-complete) Riesz spaces are closed with respect to order limits of bounded nets:

Lemma 7.288 *Let V be any order-complete (σ-order-complete) Riesz space and $V' \subseteq V$ any order-closed (σ-order-closed) Riesz subspace. Let $(v_i)_{i\in I}$ $((v_n)_{n\in\mathbb{N}})$ be a net (sequence) in V' that is order-convergent to some $v \in V$. Then, $v \in V'$.*

By combining the last lemma with Exercise 7.286, we arrive at the following result:

Corollary 7.289 *Let $V \subseteq \mathcal{F}(\Omega; \mathbb{R})$ be a function space (see Definition 7.270) that is order-closed (σ-order-closed). For any bounded net $(v_i)_{i\in I}$ (sequence $(v_n)_{n\in\mathbb{N}}$) that converges pointwise to some $f \in \mathcal{F}(\Omega; \mathbb{R})$, one has $f \in V$.*

The above definition of convergence in Riesz spaces leads to the following notion of continuity for linear transformations between Riesz spaces:

Definition 7.290 (Order and σ-Order Continuity) Let V_1 and V_2 be two Riesz spaces. A linear transformation $\Theta \in \mathcal{L}(V_1; V_2)$ is "order-continuous" if, for every net $(v_{1,i})_{i\in I}$, $v_{1,i} \overset{o}{\to} 0$, in V_1, one has that $\Theta(v_{1,i}) \overset{o}{\to} 0$ in V_2. Similarly, Θ is said to be "σ-order-continuous" if, for every *sequence* $(v_{1,n})_{n\in\mathbb{N}}$, $v_{1,n} \overset{o}{\to} 0$, in V_1, one has that $\Theta(v_{1,n}) \overset{o}{\to} 0$ in V_2. The set of all order-continuous (σ-order-continuous) linear transformation $V_1 \to V_2$ is denoted by $\mathcal{L}_n(V_1; V_2)$ ($\mathcal{L}_c(V_1; V_2)$).

It is important to notice that $\Theta \in \mathcal{L}(V_1; V_2)$ is order-continuous (σ-order-continuous) and V_1' is a Riesz subspace of the Riesz space V_1; it does not follow in general that the restriction $\Theta|_{V_1'} \in \mathcal{L}(V_1'; V_2)$ is order-continuous (σ-order-continuous) as well. This is so, because there may be nets (sequences) that are order-convergent in V_1', but not in V_1. See remarks after Definition 7.284.

Obviously, any order-continuous linear transformation $V_1 \to V_2$ is σ-order-continuous, i.e.,

$$\mathcal{L}_n(V_1; V_2) \subseteq \mathcal{L}_c(V_1; V_2) .$$

The opposite inclusion is generally wrong. Note that $\mathcal{L}_n(V_1; V_2)$ and $\mathcal{L}_c(V_1; V_2)$ are vector subspaces of $\mathcal{L}(V_1; V_2)$, by basic properties of the absolute value and order-convergent monotone nets.

In the case of *positive* linear transformations, order (σ-order) continuity can be rephrased in terms of *monotone* nets (*monotone* sequences) only, as follows:

Exercise 7.291 Let V_1 and V_2 be two Riesz spaces and Θ any fixed *positive* linear transformation $V_1 \to V_2$, i.e., $\Theta \in \mathcal{L}(V_1; V_2)^+$. Prove that the following three properties of Θ are equivalent:

(i) $\Theta \in \mathcal{L}_c(V_1; V_2)$ ($\Theta \in \mathcal{L}_n(V_1; V_2)$).
(ii) For any decreasing monotone sequence $(v_{1,n})_{n\in\mathbb{N}}$ (monotone net $(v_{1,i})_{i\in I}$), $v_{1,n} \downarrow 0$ ($v_{1,i} \downarrow 0$), of positive vectors of V_1, one has $\Theta(v_{1,n}) \downarrow 0$ ($\Theta(v_{1,i}) \downarrow 0$).
(iii) For any positive $v_1 \in V_1^+$ and any increasing monotone sequence $(v_{1,n})_{n\in\mathbb{N}}$ (increasing monotone net $(v_{1,i})_{i\in I}$), $v_{1,n} \uparrow v_1$ ($v_{1,i} \uparrow v_1$), of positive vectors of V_1, one has $\Theta(v_{1,n}) \uparrow \Theta(v_1)$ ($\Theta(v_{1,i}) \uparrow \Theta(v_1)$).

The following lemma directly follows from Exercise 7.291 (iii):

Lemma 7.292 *Let Ω be any countable set and let $\ell_1(\Omega; \mathbb{R})$ be the Riesz space of absolutely summable functions $\Omega \to \mathbb{R}$. In other words, the elements of $\ell_1(\Omega; \mathbb{R})$ are all functions $f : \Omega \to \mathbb{R}$ satisfying*

$$\sum_{p\in\Omega} |f(p)| < \infty .$$

The linear functional sum : $\ell_1(\Omega; \mathbb{R}) \to \mathbb{R}$ *defined by*

$$\mathrm{sum}\,(f) \doteq \sum_{p\in\Omega} f\,(p) \in \mathbb{R}\ , \qquad f \in \ell_1(\Omega;\mathbb{R})\ ,$$

is order-continuous. (See, in this context, Definition 7.223 (ii) and Proposition 7.224 (i).)

Proof Exercise. □

In Archimedean Riesz spaces, σ-order-continuous linear transformations $V_1 \to V_2$ are order-bounded (Definition 1.23):

Exercise 7.293 Let V_1 and V_2 be two Riesz spaces. If V_1 is Archimedean and V_2 has an order unit (Definition 1.2), then

$$\mathcal{L}_{\mathrm{n}}(V_1;V_2) \subseteq \mathcal{L}_{\mathrm{c}}(V_1;V_2) \subseteq \mathcal{L}_{\mathrm{ob}}(V_1;V_2)\ .$$

Much more will be said below about the relation between order continuity and order boundedness of linear functionals, i.e., for the case $V_2 = \mathbb{R}$.

7.4.3 Bands

We now introduce so-called bands, which are pivotal objects in the theory of Riesz spaces:

Definition 7.294 (Band) Order-closed Riesz ideals of a Riesz space are called "bands" of this space. If $W \subseteq V$ is a band of the Riesz space V for which

$$V = W \oplus W^{\mathrm{d}}$$

(see Definition 7.2), then W is said to be a "projection band."

By Exercise 1.22, for any metric space M and nonempty *open* subset $\Omega \subseteq M$,

$$C_\Omega(M;\mathbb{R}) \doteq \{f \in C(M;\mathbb{R})\ :\ f(\Omega) = \{0\}\}$$

is a band in the Riesz space $C(M;\mathbb{R})$ (Exercise 7.269), but if Ω was not open, $C_\Omega(M;\mathbb{R})$ would be, in general, only a Riesz ideal.

Exercise 7.295 Let V be any Riesz space and $W \subseteq V$ a projection band. Show that, for all $v \in V^+$, the unique decomposition $v = w + v'$, $w \in W$, $v' \in W^{\mathrm{d}}$, preserves positivity, i.e., $w \in W^+$ and $v' \in W^{\mathrm{d}+}$.

Hint: Use Exercise 7.280 (iii) and the solidity of W and W^{d}.

The following exercise gives a further important example of a band:

Exercise 7.296 Let V be any Riesz space. Prove that, for any nonempty subset $\Omega \subseteq V$, the order-disjoint complement $\Omega^{\mathrm{d}} \subseteq V$ is a band (i.e., an order-closed solid subspace) of V.

From this exercise, we conclude, in particular, that Riesz ideals decomposing the underling Riesz spaces as direct sums (Definition 7.2) are automatically projection bands:

Lemma 7.297 *Let V be any Riesz space and let $W_1, W_2 \subseteq V$ be Riesz ideals for which*

$$V = W_1 \oplus W_2 .$$

Then $W_1 = W_2^{\mathrm{d}}$ and $W_2 = W_1^{\mathrm{d}}$. In particular, W_1 and W_2 are both (projection) bands.

Proof

1. Take any $w_1 \in W_1$ and $w_2 \in W_2$. As W_1 and W_2 are Riesz subspaces, $|w_1| \in W_1$ and $|w_2| \in W_2$. As these subspaces are solid (being Riesz ideals; see Lemma 7.283), one has

$$|w_1| \wedge |w_2| \in W_1 \cap W_2 .$$

   But $W_1 \cap W_2 = \{0\}$, because $V = W_1 \oplus W_2$. (See Definition 7.2.) Thus, $|w_1| \wedge |w_2| = 0$, i.e., $w_1 \perp w_2$ for every $w_1 \in W_1$ and $w_2 \in W_2$. Hence, $W_1 \subseteq W_2^{\mathrm{d}}$ and $W_2 \subseteq W_1^{\mathrm{d}}$.
2. Take any $v \in W_2^{\mathrm{d}}$ and (uniquely) decompose it as $v = w_1 + w_2$, for $w_1 \in W_1$, $w_2 \in W_2$. Note that $v - w_1 \in W_2^{\mathrm{d}}$ and, thus,

$$w_2 \in W_2 \cap W_2^{\mathrm{d}} = \{0\} .$$

   This shows that $W_2^{\mathrm{d}} \subseteq W_1$, i.e., $W_2^{\mathrm{d}} = W_1$. We prove that $W_1^{\mathrm{d}} = W_2$ in a similar manner.

□

In fact, the last lemma shows that the notion of band is mandatory in considering disjoint decompositions of Riesz spaces.

In Archimedean Riesz spaces, all bands are disjoint complements:

Proposition 7.298 *Let V be an Archimedean Riesz space. For any band $W \subseteq V$, one has that*

$$W^{\mathrm{dd}} \doteq (W^{\mathrm{d}})^{\mathrm{d}} = W .$$

Proof

1. Let $W \subseteq V$ be any band of V. Clearly, $W \subseteq W^{\mathrm{dd}}$.
2. Take now an arbitrary $v \in W^{\mathrm{dd}} \cap V^{+}$, $v \neq 0$. In order to prove that $v \in W$, we will consider the subset

$$\Omega_v \doteq \{w \in W \ : \ v \succeq w \succeq 0\} \subseteq W \ .$$

3. By construction, v is an upper bound for Ω_v. We will prove that it is even the supremum of Ω_v in V. Recall the suprema in lattices are the same as minimal upper bound, thanks to Lemma 7.266. Thus, suppose, by contradiction, that there is another $\tilde{v} \in V^{+}$, $\tilde{v} \neq v$, $v \succeq \tilde{v}$, which is an upper bound for Ω_v. Note meanwhile that W^{dd} is (by definition) a disjoint complement of a band W and is thus a band, by Exercise 7.296. In particular, W^{dd} is solid; see Definition 7.294 and Lemma 7.283. Therefore, $\tilde{v} \in W^{\mathrm{dd}}$. In particular, $v - \tilde{v} \neq 0$, $v - \tilde{v} \succeq 0$, and $v - \tilde{v} \in W^{\mathrm{dd}}$.
4. Observing that $W^{\mathrm{d}} \cap W^{\mathrm{dd}} = \{0\}$, it follows that $v - \tilde{v} \notin W^{\mathrm{d}}$ with $|v - \tilde{v}| = v - \tilde{v} \succeq 0$, thanks to Exercise 7.280 (i). Hence, there is $w \in W$ such that $|w| \wedge (v - \tilde{v}) \neq 0$. As W is solid (being a band), it follows that, for some $w \in W^{+}$,

$$\tilde{w} \doteq w \wedge (v - \tilde{v}) \neq 0 \ .$$

Note that $w \succeq \tilde{w}$ and, thus, $\tilde{w} \in W^{+}$, as W is solid (being a band). Moreover, $v \succeq \tilde{w} \succeq 0$, that is, $\tilde{w} \in \Omega_v$.
5. Observe that

$$v = \tilde{v} + (v - \tilde{v}) \succeq \tilde{w} + \tilde{w} = 2\tilde{w} \ ,$$

i.e., $2\tilde{w} \in \Omega_v$. By iterating this estimate, we see that $n\tilde{w} \in \Omega_v$ for every $n \in \mathbb{N}$. By the Archimedean property, it follows that $\tilde{w} = 0$, which leads to a contradiction. Thus, $v = \sup \Omega_v$. As W is order-closed (being a band), $v \in W$. Therefore, $W^{\mathrm{dd}} \cap V^{+} \subseteq W \subseteq W^{\mathrm{dd}}$.
6. Let $v \in W^{\mathrm{dd}}$ be arbitrary (i.e., not necessarily positive). As $v \in W^{\mathrm{dd}}$ is a Riesz subspace (being a band), $v^{+}, v^{-} \in W^{\mathrm{dd}}$. By the previous part of the proof, $v^{+}, v^{-} \in W$, for these two vectors are positive. Hence, $v = v^{+} - v^{-} \in W$. It thus follows that $W^{\mathrm{dd}} = W$.

□

One important property of bands of Riesz spaces is that they are stable with respect to order limits:

Proposition 7.299 *Let V be a Riesz space and $W \subseteq V$ any band. For any order-convergent net $(v_i)_{i \in I}$ of elements of W, $v_i \overset{\mathrm{o}}{\to} v$, one has that $v \in W$.*

Proof Take any order-convergent net $(v_i)_{i \in I}$ of elements of W with $v_i \overset{\mathrm{o}}{\to} v$. Let $(v'_i)_{i \in I}$ be a net in V such that $v'_i \downarrow 0$ and $v'_i \succeq |v - v_i|$ for every $i \in I$, which must exist, by the definition of order convergence (Definition 7.284). By Exercise 7.280

(vi), $|v_i| \succeq |v| - v_i'$. Thus, as W is solid (being a band), $(|v| - v_i')^+ \in W$. Observe that $(|v| - v_i') \uparrow |v|$. Hence,

$$(|v| - v_i')^+ = (|v| - v_i') \vee 0 \uparrow |v| \vee 0 = |v| \ ,$$

thanks to Exercises 7.272 (vii) and 7.280 (iv). As W is order-closed (being a band), we deduce that $|v| \in W$. Hence, again by solidness of W, $v \in W$. □

From the last two propositions, we see that bands play in the theory of Riesz spaces an analogous role as the one of closed subspaces in the theory of Hilbert spaces. The next corollary strengthens this analogy:

Proposition 7.300 (Riesz) *Let V be an order-complete Riesz space and $W \subseteq V$ any band. Then*

$$V = W \oplus W^{\mathrm{d}} \ .$$

That is, any band in an arbitrary order-complete Riesz space is a projection band.

Proof

1. Observe, for a given vector $v \in V$, that if $v = w + v'$ with $w \in W$ and $v' \in W^{\mathrm{d}}$, then such a decomposition is unique: Assume that $v = \tilde{w} + \tilde{v}'$ for $\tilde{w} \in W$ and $\tilde{v}' \in W^{\mathrm{d}}$. Then

$$\tilde{w} - w = v' - \tilde{v}' \ .$$

In particular,

$$(\tilde{w} - w), (v' - \tilde{v}') \in W \cap W^{\mathrm{d}} = \{0\} \ .$$

Hence, $\tilde{w} = w$ and $\tilde{v}' = v$.

2. Now we prove that, under the assumptions of the proposition, such a decomposition exists for any $v \in V$. Being a Riesz space, V is directed; see Definitions 1.13 and 7.267 (i). Thus, by Exercise 1.14, V^+ generates V, and, as W and W^{d} are (Riesz) subspaces, it suffices to prove the existence of the decomposition for arbitrary $v \in V^+$. Thus, fix any $v \in V^+$ and (similarly as in the proof of Proposition 7.298) define the upper bounded subset

$$\Omega_v \doteq \{w' \in W \ : \ v \succeq w'\} \ .$$

Let $w \doteq \sup \Omega_v \in W^+$ and $v' \doteq v - w \in V^+$. Notice that w exists, because V is order-complete, and belongs to W, for this subspace is order-closed, being a band.

3. We show now that $v - w \in W^{\mathrm{d}}$. Suppose, by contradiction, that $v - w \notin W^{\mathrm{d}}$. In this case, by the solidness of W, there is $w' \in W^+$ such that $w' \wedge (v - w) \succeq 0$ is not zero. Observe from Exercise 7.272 (i–ii) that

$$
\begin{aligned}
v - (w + w' \wedge (v - w)) &= v - w - (w' \wedge (v - w)) \\
&= v - w + (-w') \vee (w - v) \\
&= (v - w - w') \vee 0 \succeq 0 \,.
\end{aligned}
$$

But, by the above inequality, if $w' \wedge (v - w)$ was not zero, then w would not be the supremum of Ω_v, because $w' \wedge (v - w) \in W$, as W is solid.

□

Exercise 7.301 Let V be any Riesz space. For all $v \in V$, define

$$
B_v \doteq \{v' \in V \;:\; |v'| = \sup\{|v'| \wedge \alpha |v| \;:\; \alpha \geq 0\}\} \subseteq V \,.
$$

Show that B_v is the smallest band of V that contains v.

Hints: Note that V itself is a band that contains v. Prove first that B_v is a subset of any band that contains v. Prove next that B_v is a vector subspace of V. Observe that B_v is preserved by the absolute value and conclude from this fact that B_v is a Riesz subspace. Show, finally, that B_v is solid and order-closed.

The band B_v, $v \in V$, of the last exercise is called the "band generated by v." As an example of application of this construction, in measure theory, it yields, by Corollary 7.276, the famous Lebesgue's decomposition theorem.

7.4.4 *Bands as Ranges of Positive Projectors*

Recall that, from the definition of projection bands (Definition 7.294), for any Riesz space V and projection band $W \subseteq V$, one has that

$$
V = W \oplus W^{\mathrm{d}}.
$$

In particular, we can define a mapping $P_W : V \to V$ by the condition $P_W(w+v') = w$ for every $w \in W$ and $v' \in W^{\mathrm{d}}$. Note that this well-defines a mapping, because any $v \in V$ has a unique decomposition as $v = w + v'$, $w \in W$, and $v' \in W^{\mathrm{d}}$. See Definition 7.2. This mapping has the following properties:

Exercise 7.302 Let V be a Riesz space and $W \subseteq V$ any projection band. Show that P_W is a projector, i.e., $P_W \in \mathcal{L}(V)$ (linearity) and $P_W \circ P_W = P_W$, satisfying

$$
\mathrm{id}_V \succeq P_W \succeq 0 \,.
$$

Here, $\mathrm{id}_V \in \mathcal{L}(V)$ is the identity mapping.

Observe that, clearly,

$$
\mathrm{ran}(P_W) \doteq P_W(V) = W \,.
$$

In particular, the range $\mathrm{ran}(P_W)$ is a projection band. We will prove below that any such a band in V is the range of a projector $P \in \mathcal{L}(V)^+$ with $\mathrm{id}_V \succeq P$. Therefore, the set of projection bands of V can be identified with the set of all positive projectors on V that are bounded from above by the identity operator id_V. Recall that, in an order-complete Riesz space, any band is a projection band, thanks to Proposition 7.300. Thus, we have, in this case, a one-to-one correspondence between bands and positive projectors bounded by id_V. Interestingly, closed subspaces of Hilbert spaces have a quite analogous property. See Proposition 7.238.

Lemma 7.303 *Let V be an arbitrary Riesz space. For an arbitrary $\Theta \in \mathcal{L}(V)$, the following three properties are equivalent:*

(i) $\Theta = P_W$ *for some projection band* $W \subseteq V$.
(ii) Θ *is a projector (i.e.,* $\Theta \circ \Theta = \Theta$*) and* $\mathrm{id}_V \succeq \Theta \succeq 0$.
(iii) *For all* $v, v' \in V$, $\Theta(v) \perp (\mathrm{id}_V - \Theta)(v')$.

Proof The implication (i)⇒(ii) refers to the last exercise. We prove next (ii)⇒(iii): Assume that Θ is a projector satisfying $\mathrm{id}_V \succeq \Theta \succeq 0$. Take any positive $v, v' \in V^+$ and define

$$w \doteq \Theta(v) \wedge (\mathrm{id}_V - \Theta)(v') \in V^+ .$$

Then, from $w \in V^+$ and $(\mathrm{id}_V - \Theta)(v') \succeq w$, we arrive at

$$0 = (\Theta - \Theta \circ \Theta)(v') = \Theta((\mathrm{id}_V - \Theta)(v')) \succeq \Theta(w) \succeq 0 .$$

In other words, $\Theta(w) = 0$. Similarly, from $w \in V^+$ and $\Theta(v) \succeq w$, we arrive at

$$0 = (\Theta - \Theta \circ \Theta)(v) = (\mathrm{id}_V - \Theta)(\Theta(v)) \succeq (\mathrm{id}_V - \Theta)(w) \succeq 0 .$$

That is, $(\mathrm{id}_V - \Theta)(w) = 0$. Hence,

$$w = (\mathrm{id}_V - \Theta)(w) + \Theta(w) = 0 .$$

This implies that

$$\Theta(v) \wedge (\mathrm{id}_V - \Theta)(v') = 0$$

for all $v, v' \in V$ (and not only for $v, v' \in V^+$), because both Θ and $(\mathrm{id}_V - \Theta)$ are positive, whence, for all $v \in V$,

$$\Theta(|v|) \succeq |\Theta(v)| , \qquad (\mathrm{id}_V - \Theta)(|v|) \succeq |(\mathrm{id}_V - \Theta)(v)| .$$

Keeping in mind Definition 7.278 (iii), we conclude (ii)⇒(iii). It remains to prove (iii)⇒(i): Assume that, for all $v, v' \in V$, $\Theta(v) \perp (\mathrm{id}_V - \Theta)(v')$. Let

$$W \doteq \mathrm{ran}(\Theta)^{\mathrm{dd}} \text{ and } W' \doteq \mathrm{ran}(\mathrm{id}_V - \Theta)^{\mathrm{dd}}.$$

By Lemma 7.279, one has

$$W^{\mathrm{d}} = (\mathrm{ran}(\Theta)^{\mathrm{dd}})^{\mathrm{d}} = \mathrm{ran}(\Theta)^{\mathrm{d}} \supseteq \mathrm{ran}(\mathrm{id}_V - \Theta) ,$$
$$W'^{\mathrm{d}} = (\mathrm{ran}(\mathrm{id}_V - \Theta)^{\mathrm{dd}})^{\mathrm{d}} = \mathrm{ran}(\mathrm{id}_V - \Theta)^{\mathrm{d}} \supseteq \mathrm{ran}(\Theta) .$$

By applying once again Lemma 7.279, we arrive at

$$W^{\mathrm{d}} \supseteq W' , \qquad W'^{\mathrm{d}} \supseteq W .$$

In other words, $w \perp w'$ for all $w \in W$ and $w' \in W'$. In particular, it follows that $W \cap W' = \{0\}$. This implies that if some $v \in V$ can be decomposed as $v = w + w'$ for $w \in W$ and $w' \in W'$, then the decomposition is unique. Clearly, for all $v \in V$, one has

$$v = \Theta(v) + (\mathrm{id}_V - \Theta)(v) .$$

As

$$\Theta(v) \in \mathrm{ran}(\Theta) \subseteq W , \qquad (\mathrm{id}_V - \Theta)(v) \in \mathrm{ran}(\mathrm{id}_V - \Theta) \subseteq W' ,$$

we conclude that $V = W \oplus W'$. From Lemma 7.297, W and W' are projection bands and

$$W^{\mathrm{d}} = W' , \qquad W'^{\mathrm{d}} = W .$$

Finally, observing that $v - \Theta(v) \in W^{\mathrm{d}}$, for all $v \in V$,

$$P_W(v) - \Theta(v) = P_W(v - \Theta(v)) = 0 .$$

It follows that $\Theta = P_W$ and, thus, (iii)⇒(i). □

7.4.5 The Order Dual of a Riesz Space

Thiss subsection is devoted to the order duals of Riesz spaces. We will show, in particular, that they are themselves Riesz spaces. We start with a well-known result, which states that linear functionals on Riesz spaces can be identified, by restriction, with the additive positive-valued mappings on the positive cone of the space.

Proposition 7.304 (Kantorovich) *Let V_1 be any Riesz space and V_2 an Archimedean ordered vector space. Let $\Phi : V_1^+ \to V_2^+$ be an arbitrary* additive *mapping, i.e., for all $v_1, v_1' \in V_1^+$,*

$$\Phi(v_1 + v_1') = \Phi(v_1) + \Phi(v_1') .$$

Φ is the restriction to V_1^+ of a unique $\Theta_\Phi \in \mathcal{L}(V_1; V_2)^+$.

Proof

1. Observe first that if such a $\Theta_\Phi \in \mathcal{L}(V_1; V_2)^+$ exists, then, by linearity, for all $v_1 \in V_1$, it has to satisfy

$$\begin{aligned}\Theta_\Phi(v_1) &= \Theta_\Phi(v_1^+ - v_1^-) \\ &= \Theta_\Phi(v_1^+) - \Theta_\Phi(v_1^-) = \Phi(v_1^+) - \Phi(v_1^-) .\end{aligned}$$

In particular, Θ_Φ is unique. So, define the mapping $\Theta_\Phi : V_1 \to V_2$ by the above expression, i.e.,

$$\Theta_\Phi(v_1) \doteq \Phi(v_1^+) - \Phi(v_1^-) , \qquad v_1 \in V_1 .$$

2. We show first that Θ_Φ is additive (on the whole V_1). Take any $v_1, v_1' \in V_1$ and note that

$$(v_1 + v_1')^+ - (v_1 + v_1')^- = v_1 + v_1' = v_1^+ - v_1^- + v_1'^+ - v_1'^- .$$

In other words,

$$(v_1 + v_1')^+ + v_1^- + v_1'^- = (v_1 + v_1')^- + v_1^+ + v_1'^+ .$$

By the additivity of Θ_Φ (i.e., of Φ) on V_1^+, we conclude that

$$\Theta_\Phi((v_1+v_1')^+)+\Theta_\Phi(v_1^-)+\Theta_\Phi(v_1'^-) = \Theta_\Phi((v_1+v_1')^-)+\Theta_\Phi(v_1^+)+\Theta_\Phi(v_1'^+) ,$$

which in turn implies that

$$\begin{aligned}\Theta_\Phi(v_1 + v_1') &= \Theta_\Phi((v_1 + v_1')^+) - \Theta_\Phi((v_1 + v_1')^-) \\ &= (\Theta_\Phi(v_1^+) - \Theta_\Phi(v_1^-)) + (\Theta_\Phi(v_1'^+) - \Theta_\Phi(v_1'^-)) \\ &= \Theta_\Phi(v_1) + \Theta_\Phi(v_1') .\end{aligned}$$

3. Similarly, for all $v_1 \in V_1$, one has

$$\begin{aligned}\Theta_\Phi(-v_1) &= \Theta_\Phi((-v_1)^+) - \Theta_\Phi((-v_1)^-) \\ &= \Theta_\Phi(v_1^-) - \Theta_\Phi(v_1^+) = -\Theta_\Phi(v_1) .\end{aligned}$$

4. Observe also that, from its additivity and positivity of Φ, Θ_Φ is order-preserving: For all $v_1, v_1' \in V_1$ satisfying $v_1 \succeq v_1'$,

$$\begin{aligned}\Theta_\Phi(v_1) &= \Theta_\Phi(v_1' + (v_1 - v_1')) \\ &= \Theta_\Phi(v_1') + \Theta_\Phi(v_1 - v_1') \\ &= \Theta_\Phi(v_1') + \Phi(v_1 - v_1') \succeq \Theta_\Phi(v_1') .\end{aligned}$$

5. To finish the proof, we have to show that, for all $\alpha > 0$ and all $v_1 \in V_1$, $\Theta_\Phi(\alpha v_1) = \alpha\Theta_\Phi(v_1)$. Note that, for $\alpha \in \mathbb{N}$, this is a direct consequence of additivity. Consider now any positive rational number $\alpha = m/n$ with $m, n \in \mathbb{N}$. Then, for all $v_1 \in V_1$,

$$\Theta_\Phi(\alpha v_1) = m\Theta_\Phi(n^{-1}v_1) = m\frac{n}{n}\Theta_\Phi(n^{-1}v_1) = \frac{m}{n}\Theta_\Phi\left(\frac{n}{n}v_1\right) = \alpha\Theta_\Phi(v_1) \ .$$

6. For an arbitrary positive real number $\alpha > 0$, take two sequences of positive rational numbers $\alpha_k, \alpha'_k \in \mathbb{Q}^+$, $k \in \mathbb{N}$, such that $\alpha_k \leq \alpha \leq \alpha'_k$ for all $k \in \mathbb{N}$, and

$$\lim_{k\to\infty}(\alpha'_k - \alpha_k) = 0 \ .$$

Then, for any positive $v_1 \in V_1^+$, as Θ_Φ preserves order, we obtain that, for all $k \in \mathbb{N}$,

$$(\alpha'_k - \alpha)\Theta_\Phi(v_1) \succeq \Theta_\Phi(\alpha v_1) - \alpha\Theta_\Phi(v_1) \succeq -(\alpha - \alpha_k)\Theta_\Phi(v_1) \ .$$

By the Archimedean property (Definition 1.2), from the first inequality, we conclude that

$$0 \succeq \Theta_\Phi(\alpha v_1) - \alpha\Theta_\Phi(v_1) \ ,$$

and, from the second,

$$0 \succeq \alpha\Theta_\Phi(v_1) - \Theta_\Phi(\alpha v_1) \ .$$

Hence, $\alpha\Theta_\Phi(v_1) = \Theta_\Phi(\alpha v_1)$. By additivity and step 3, for all $\alpha \in \mathbb{R}$ and all $v_1 \in V_1$ (both not necessarily positive anymore), we have again $\alpha\Theta_\Phi(v_1) = \Theta_\Phi(\alpha v_1)$.

□

By using the above proposition, we prove next that, given any Riesz space V_1 and an *order-complete* Archimedean ordered vector space V_2, the space $\mathcal{L}_{\text{ob}}(V_1; V_2)$ of all order-bounded linear transformations $V_1 \to V_2$ (Definition 1.23) is again a Riesz space. To this end, recall the definition of intervals: For any v, v' in a preordered vector space V,

$$[v, v'] \doteq (v + V^+) \cap (v' - V^+) = \{v'' \in V \ : \ v' \succeq v'' \succeq v\} \subseteq V \ .$$

See Definition 1.4.

Lemma 7.305 *Let V_1 be any Riesz space and V_2 an order-complete Archimedean ordered vector space. For all $\Theta \in \mathcal{L}_{\text{ob}}(V_1; V_2)$, the subset $\{0, \Theta\} \subseteq \mathcal{L}_{\text{ob}}(V_1; V_2)$ has a supremum in $\mathcal{L}_{\text{ob}}(V_1; V_2)$, which is uniquely defined by the identity*

$$\sup\{0, \Theta\}(v_1) = \sup \Theta([0, v_1]) \in V_2^+ , \qquad v_1 \in V_1^+ .$$

Observe that the right-hand side of the above expression is well-defined, for V_2 is order-complete (Definition 1.20) and Θ order-bounded (Definition 1.23).

Proof

1. Take any $\Theta \in \mathcal{L}_{\text{ob}}(V_1; V_2)$. Define the mapping $\Phi : V_1^+ \to V_2^+$ by

$$\Phi(v_1) \doteq \sup \Theta([0, v_1]) \in V_2^+ , \qquad v_1 \in V_1^+ .$$

Note first that, for any upper bound $\Theta' \in \mathcal{L}(V_1; V_2)^+$ of the subset $\{0, \Theta\} \subseteq \mathcal{L}_{\text{ob}}(V_1; V_2)^+$, one has that, for all $v_1 \in V_1^+$,

$$\Theta'(v_1) = \sup \Theta'([0, v_1]) \succeq \sup \Theta([0, v_1]) = \Phi(v_1) .$$

On the other hand, for all $v_1 \in V_1^+$,

$$\Phi(v_1) = \sup \Theta([0, v_1]) \succeq (0 \vee \Theta(v_1)) \ .$$

Hence, if Φ is the restriction to V_1^+ of some

$$\Theta^+ \in \mathcal{L}(V_1; V_2)^+ \subseteq \mathcal{L}_{\text{ob}}(V_1; V_2)$$

then Θ^+ is the supremum of $\{0, \Theta\} \subseteq \mathcal{L}(V_1; V_2)^+$ (see Lemma 7.266). This positive linear transformation (trivially) satisfies the identity given in the lemma, which determines Θ^+ uniquely, since $\Theta^+(v_1) = \Theta^+(v_1^+) - \Theta^+(v_1^-)$ for any $v_1 \in V_1$.

2. To show the existence of Θ^+, we prove that Φ is additive and use Proposition 7.304 (Kantorovich's theorem). Thus, take two arbitrary $v_1, v_1' \in V_1^+$. From the straightforward equality

$$\Phi(v_1 + v_1') = \sup\{\Theta(v_1'') \ : \ v_1'' \in V_1^+ \ , \ v_1 + v_1' \succeq v_1''\}$$

we conclude that, for all $w_1 \in [0, v_1]$ and $w_1' \in [0, v_1']$,

$$\Phi(v_1 + v_1') \succeq \Theta(w_1 + w_1') = \Theta(w_1) + \Theta(w_1') ,$$

i.e.,

$$\Phi(v_1 + v_1') - \Theta(w_1') \succeq \Theta(w_1) .$$

Taking the supremum with respect to $w_1 \in [0, v_1]$ for any fixed $w_1' \in [0, v_1']$, we arrive at

$$\Phi(v_1 + v_1') - \Theta(w_1') \succeq \Phi(w_1) \text{ , i.e., } \Phi(v_1 + v_1') - \Phi(v_1) \succeq \Theta(w_1') .$$

Now, doing the same for $w'_1 \in [0, v'_1]$, we deduce that

$$\Phi(v_1 + v'_1) - \Phi(v_1) \succeq \Phi(v'_1) \text{ , i.e., } \Phi(v_1 + v'_1) \succeq \Phi(v_1) + \Phi(v'_1) \text{ .}$$

3. Now, by Lemma 7.274, for any $v_1, v'_1 \in V_1^+$ and $v''_1 \in V_1^+$ satisfying $v_1 + v'_1 \succeq v''_1$, there are $w_1 \in [0, v_1]$ and $w'_1 \in [0, v'_1]$, such that $v''_1 = w_1 + w'_1$. Hence,

$$\Phi(v_1) + \Phi(v'_1) \succeq \Theta(w_1) + \Theta(w'_1) = \Theta(v''_1) \text{ .}$$

Taking the supremum with respect to $v''_1 \in [0, v_1 + v'_1]$,

$$\Phi(v_1) + \Phi(v'_1) \succeq \Phi(v_1 + v'_1) \text{ .}$$

As V_2 is an ordered vector space (Definition 1.27), it follows that, for any $v_1, v'_1 \in V_1^+$,

$$\Phi(v_1 + v'_1) = \Phi(v_1) + \Phi(v'_1) \text{ .}$$

In other words, Φ is additive, and Proposition 7.304 ensures the existence of a (unique) positive linear transformation Θ^+ extending Φ to all V_1.

□

Using Exercise 7.273, observe that the last lemma implies that $\mathcal{L}_{\text{ob}}(V_1; V_2)$ is a Riesz space. In particular, $\mathcal{L}_{\text{ob}}(V_1; V_2)$ is directed; see Definitions 1.13 and 7.267 (i). Thus, by Exercise 1.14, $\mathcal{L}(V_1; V_2)^+$ generates $\mathcal{L}_{\text{ob}}(V_1; V_2)$, i.e., any order-bounded linear transformation $V_1 \to V_2$ is the difference of two positive ones. Taking $V_2 = \mathbb{R}$, this means, for instance, that, for any Riesz space V, $\mathcal{L}_{\text{ob}}(V; \mathbb{R})$ is the "order dual" of V, i.e.,

$$\mathcal{L}_{\text{ob}}(V; \mathbb{R}) = V^{\text{od}} \doteq \text{span}(\mathcal{L}(V; \mathbb{R})^+) \text{ .}$$

See Definition 1.25.

Exercise 7.306 Let V_1 be any Riesz space and V_2 an order-complete Archimedean ordered vector space. Show that in the Riesz space $\mathcal{L}_{\text{ob}}(V_1; V_2)$, for all $\Theta, \Theta' \in \mathcal{L}_{\text{ob}}(V_1; V_2)$ and all $v_1 \in V_1^+$, the following equalities are satisfied:

(i) $(\Theta \wedge \Theta')(v_1) = \inf\{\Theta(v'_1) + \Theta'(v_1 - v'_1) : v'_1 \in [0, v_1]\}$.
(ii) $(\Theta \vee \Theta')(v_1) = \sup\{\Theta(v'_1) + \Theta'(v_1 - v'_1) : v'_1 \in [0, v_1]\}$.
(iii) $|\Theta|(v_1) = \sup\{\Theta(v'_1) : |v'_1| \in [0, v_1]\}$.

In the following proposition, we show that, under the conditions of Lemma 7.305, the Riesz space $\mathcal{L}_{\text{ob}}(V_1; V_2)$ is additionally order-complete, i.e., every subset of $\mathcal{L}_{\text{ob}}(V_1; V_2)$ which is bounded from above has a supremum (Definition 1.20):

Proposition 7.307 *Let V_1 be any Riesz space and V_2 an order-complete Archimedean ordered vector space. $\mathcal{L}_{\text{ob}}(V_1; V_2)$ is an order-complete Riesz space.*

Proof The arguments will be similar to those used in the proof of Lemma 7.305.

1. Recall that a Riesz space is order-complete iff any increasing monotone net of positive elements of this space, which is bounded, has an order limit. See Corollary 7.276. Thus, take any such a net $(\Theta_i)_{i\in I}$ in $\mathcal{L}_{\text{ob}}(V_1; V_2)^+$. As $(\Theta_i)_{i\in I}$ is bounded, for any $v_1 \in V_1^+$, $(\Theta_i(v_1))_{i\in I}$ is an increasing monotone net of elements of V_2^+, which is again bounded. In particular, as V_2 is order-complete (Definition 1.20), it has an order limit (Definition 1.17). So, define the mapping $\Phi : V_1^+ \to V_2^+$ by
$$\Phi(v_1) \doteq \sup\{\Theta_i(v_1) \ : \ i \in I\} \ .$$
2. We will show now that Φ is additive. Take two positive elements $v_1, v_1' \in V_1^+$. For all $i \in I$,
$$\Phi(v_1 + v_1') \succeq \Theta_i(v_1) + \Theta_i(v_1') \ .$$
As I is a directed set and the nets $(\Theta_i(v_1))_{i\in I}$ and $(\Theta_i(v_1'))_{i\in I}$ are increasing, it follows that, for all $i, i' \in I$,
$$\Phi(v_1 + v_1') \succeq \Theta_i(v_1) + \Theta_{i'}(v_1') \ .$$
Taking suprema with respect to i and i' separately, we arrive at
$$\Phi(v_1 + v_1') \succeq \Phi(v_1) + \Phi(v_1') \ .$$
On the other hand, clearly, for all $i \in I$,
$$\Phi(v_1) + \Phi(v_1') \succeq \Theta_i(v_1) + \Theta_i(v_1') = \Theta_i(v_1 + v_1')$$
and, hence,
$$\Phi(v_1) + \Phi(v_1') \succeq \Phi(v_1 + v_1') \ .$$
As V_2 is an ordered vector space (Definition 1.27), it follows that
$$\Phi(v_1 + v_1') = \Theta_i(v_1) + \Theta_{i'}(v_1') \ , \qquad v_1, v_1' \in V_1^+ \ .$$
3. From Proposition 7.304 (Kantorovich's theorem), Φ is the restriction to V_1^+ of a unique $\overline{\Theta} \in \mathcal{L}_{\text{ob}}(V_1; V_2)^+$. By construction, $\overline{\Theta} \succeq \Theta_i$ for every $i \in I$. Let $\Theta' \in \mathcal{L}_{\text{ob}}(V_1; V_2)^+$ be any other upper bound for the net $(\Theta_i)_{i\in I}$. Then, for all $v_1 \in V_1^+$ and $i \in I$, $\Theta'(v_1) \succeq \Theta_i(v_1)$, that is,
$$\Theta'(v_1) \succeq \overline{\Theta}(v_1) \ , \qquad v_1 \in V_1^+ \ .$$

Hence, by Lemma 7.266,

$$\overline{\Theta} = \sup\{\Theta_i \;:\; i \in I\}$$

and $\Theta_i \uparrow \overline{\Theta}$; see Definition 1.17.

□

Observe that the last proposition is the analogue for Riesz spaces of a well-known property of normed spaces, stated in Exercise 7.86 (iv).

Noting that $\mathbb{R}$ is an order-complete Archimedean ordered vector space and recalling that $V^{\mathrm{od}} = \mathcal{L}_{\mathrm{ob}}(V;\mathbb{R})$, the last proposition has the following important consequence:

Corollary 7.308 *The order dual of any Riesz space is an order-complete Riesz space. In particular, by Proposition 7.300, any band of the order dual is a projection band.*

By Exercise 7.293, if V is an Archimedean Riesz space, then recall that

$$\mathcal{L}_{\mathrm{n}}(V;\mathbb{R}) \subseteq \mathcal{L}_{\mathrm{c}}(V;\mathbb{R}) \subseteq \mathcal{L}_{\mathrm{ob}}(V;\mathbb{R}) = V^{\mathrm{od}} .$$

Here, $\mathcal{L}_{\mathrm{n}}(V;\mathbb{R})$ ($\mathcal{L}_{\mathrm{c}}(V;\mathbb{R})$) is the space of all order-continuous (σ-order-continuous) linear transformation $V \to \mathbb{R}$; see Definition 7.290. These spaces and their elements are naturally renamed, in this special case, as follows:

Definition 7.309 (Integrals and Singular Functionals) We call $\mathcal{L}_{\mathrm{n}}(V;\mathbb{R}) \subseteq V^{\mathrm{od}}$ the "normal order dual" of the Archimedean Riesz space V and $\mathcal{L}_{\mathrm{c}}(V;\mathbb{R}) \subseteq V^{\mathrm{od}}$ its "continuous order dual." These vector subspaces are, respectively, denoted here by V^{n} and V^{c}. The elements of V^{c} (V^{n}) are called "integrals" ("normal integrals") on V. Linear functionals $\varphi \in V^{\mathrm{od}}$ that are elements of the disjoint complement $(V^{\mathrm{n}})^{\mathrm{d}}$ are called "singular functionals" on V.

It turns out that V^{n} and V^{c} are bands in the order dual V^{od}:

Theorem 7.310 (Ogasawara) *For any Archimedean Riesz space V, V^{n} and V^{c} are bands in the order dual V^{od}.*

A proof of this result can be found in [4, Theorem 4.4].

Using this theorem together with Proposition 7.300 and Corollary 7.308, V^{n} and V^{c} are projection bands, and the following important corollary holds true:

Corollary 7.311 *Let V be an arbitrary Archimedean Riesz space. Every $\varphi \in V^{\mathrm{od}}$ can be uniquely decomposed as a sum of a normal integral and a singular functional on V, i.e.,*

$$V = V^{\mathrm{n}} \oplus (V^{\mathrm{n}})^{\mathrm{d}} .$$

Notice that usual Lebesgue integrals are σ-order continuous linear functionals in the Riesz spaces of the corresponding integrable functions. Thus, they are integrals in the general abstract sense, as defined above. In this special case, observe additionally that Exercise 7.291 (iii) refers to the celebrated (Beppo Levi) monotone convergence theorem. As an important example of normal integrals, we mention the normal positive functionals on commutative von Neumann algebras. See Sect. 2.2 for more details on this type of functional. Recall that any commutative C^*-algebra is a Riesz space, by the results presented in Sect. 4.7.

Similar to the case of (usual) Lebesgue integrals for positive measures, order continuity of general abstract positive integrals, or, equivalently, Exercise 7.291 (iii), which can be seen as a general abstract version of monotone convergence theorem of integration theory, has as important consequences a (general abstract) version of Fatou's lemma, as well as one of Lebesgue's dominated convergence theorem, for integrals and normal integrals:

Proposition 7.312 (Fatou's Lemma) *Let V be any order-complete (σ-order-complete) Riesz space. For any* positive *normal integral $\varphi \in V^{\mathrm{n}}$ (integral $\varphi \in V^{\mathrm{c}}$) and any net $(v_i)_{i\in I}$ (sequence $(v_n)_{n\in\mathbb{N}}$) in V^+, i.e., any net (sequence) of positive vectors of V, one has that*

$$\varphi\left(\liminf_{i\in I} v_i\right) \leq \liminf_{i\in I} \varphi\left(v_i\right) \ ,$$

where, as is usual,

$$\liminf_{i\in I} v_i \doteq \sup\left\{\inf\{v_j \ : \ j \succeq i\} \ : \ i \in I\right\} \in V^+ \ ,$$

$$\liminf_{i\in I} \varphi\left(v_i\right) \doteq \sup\left\{\inf\{\varphi\left(v_j\right) \ : \ j \succeq i\} \ : \ i \in I\right\} \in \mathbb{R}_0^+ \ .$$

(Mutatis mutandis for integrals $\varphi \in V^{\mathrm{c}}$ and sequences $(v_n)_{n\in\mathbb{N}}$.)

Proof Take any net $(v_i)_{i\in I}$ in V^+ and a (positive) normal integral $\varphi \in V^{\mathrm{n}}$. For all $i \in I$, define

$$w_i \doteq \inf\{v_j \ : \ j \succeq i\} \in V \ .$$

Observe that $(w_i)_{i\in I}$ is a monotonically increasing net satisfying

$$\liminf_{i\in I} v_i = \sup_{i\in I} w_i \ ,$$

that is, $w_i \uparrow \liminf_{i\in I} v_i$. Clearly, $v_i \succeq w_i \succeq 0$ for all $i \in I$ and $\liminf_{i\in I} v_i \succeq 0$. Thus, as φ is order-continuous and positive, we arrive at

$$\varphi\left(\liminf_{i\in I} v_i\right) = \lim_{i\in I} \varphi\left(w_i\right) = \liminf_{i\in I} \varphi\left(w_i\right) \leq \liminf_{i\in I} \varphi\left(v_i\right) \ .$$

See also Exercise 7.291 (iii). (Mutatis mutandis for integrals $\varphi \in V^{\mathrm{c}}$ and sequences $(v_n)_{n\in\mathbb{N}}$.) □

We combine Exercise 7.287 (well-known for real numbers) with the last proposition in order to derive an abstract version of Lebesgue's dominated convergence theorem for normal integrals:

Proposition 7.313 (Dominated Convergence) *Let V be any order-complete (σ-order-complete) Riesz space and $\varphi \in V^{\mathrm{n}}$ ($\varphi \in V^{\mathrm{c}}$) any* positive *normal integral (integral). Take a net $(v_i)_{i\in I}$ (sequence $(v_n)_{n\in\mathbb{N}}$) in V that is order-convergent to some $v \in V$. If there is $w \in V^+$ such that $w \succeq |v_i|$ for all $i \in I$, then*

$$\lim_{i\in I} \varphi(v_i) = \varphi(v) \ .$$

Proof Take any net $(v_i)_{i\in I}$ in V satisfying the conditions of the proposition. Define the nets $(v_i^{\pm})_{i\in I}$ in V^+ (i.e., new nets of positive vectors) by $v_i^{\pm} \doteq w \pm v_i$. By the last proposition (i.e., Fatou's lemma for abstract normal integrals), one has that

$$\varphi\left(\liminf_{i\in I} v_i^{\pm}\right) \leq \liminf_{i\in I} \varphi\left(v_i^{\pm}\right) \ .$$

Observing that

$$\liminf_{i\in I} v_i^{\pm} = w + \liminf_{i\in I}(\pm v_i) \ ,$$

$$\liminf_{i\in I} \varphi\left(v_i^{\pm}\right) = \varphi(w) + \liminf_{i\in I}(\pm\varphi(v_i)) \ ,$$

we conclude that

$$\varphi\left(\liminf_{i\in I}(\pm v_i)\right) \leq \liminf_{i\in I}(\pm\varphi(v_i)) \ .$$

Note that $(-v_i)_{i\in I}$ is order-convergent to $-v$, as $(v_i)_{i\in I}$ is order-convergent to v, by assumption. From Exercise 7.287, we then obtain that

$$\pm\,\varphi(v) \leq \liminf_{i\in I}(\pm\varphi(v_i)) \ .$$

Thus,

$$\limsup_{i\in I} \varphi(v_i) = -\liminf_{i\in I}(-\varphi(v_i)) \leq \varphi(v) \leq \liminf_{i\in I} \varphi(v_i) \leq \limsup_{i\in I} \varphi(v_i)$$

and the assertion follows. (Mutatis mutandis for integrals $\varphi \in V^{\mathrm{c}}$ and sequences $(v_n)_{n\in\mathbb{N}}$.) □

The following simple corollary of the last proposition combined with Lemma 7.292 is a well-known property of series that is very useful in analysis:

Corollary 7.314 (Dominated Convergence for Series) *Let Ω be any countable set and let $(f_i)_{i\in I}$ be a net of absolutely summable functions $\Omega \to \mathbb{C}$. Assume the existence of a (positive valued) absolutely summable function $g : \Omega \to \mathbb{C}$ such that $|f_i(p)| \leq g(p)$ for all $i \in I$ and $p \in \Omega$. If the net converges pointwise, i.e., the limit $\lim_{i\in I} f_i(p)$ exists for all $p \in \Omega$, then the function $p \mapsto \lim_{i\in I} f_i(p)$ is absolutely summable in Ω, and one has the equality*

$$\lim_{i\in I} \sum_{p\in\Omega} f_i(p) = \sum_{p\in\Omega} \lim_{i\in I} f_i(p) \ .$$

Proof Exercise. *Hint*: Prove that the Riesz space $\ell_1(\Omega; \mathbb{R})$ of absolutely summable functions $\Omega \to \mathbb{C}$ is order-complete. See also Exercise 7.286. □

7.4.6 Normed Lattices

If a given Riesz space is to be endowed with a norm, one usually requires some compatibility between the norm and the lattice structure. This refers to normed lattices defined as follows:

Definition 7.315 (Normed Lattice) Let V be a Riesz space. We say that a norm $\|\cdot\|$ in V is a "lattice norm" if, for all $v, v' \in V$, $|v| \succeq |v'|$ implies that $\|v\| \geq \left\|v'\right\|$. A Riesz space endowed with a lattice norm is called a "normed lattice." If the normed lattice is a Banach space (Definition 7.82), then it is called a "Banach lattice."

The space of real numbers $\mathbb{R}$ with the absolute value norm and $C(K; \mathbb{R})$ with the supremum norm, where K is some compact metric space, are simple examples of Banach lattices. Observe also that, up to a norm equivalence, for a given Riesz space V, a lattice norm in V with respect to which V is complete (i.e., a Banach lattice) is unique:

Theorem 7.316 *Let V be any Riesz space and $\|\cdot\|$, $\|\cdot\|'$ any two* lattice norms *in V. If V is complete with respect to both $\|\cdot\|$ and $\|\cdot\|'$, then these two norms are equivalent.*

The above result is only given for completeness and will not be proven here. For more details, see, for instance, [3, 9.10 Corollary].

Exercise 7.317 Let V be a Riesz space and $\|\cdot\|$ a norm in V.

(i) Show that the norm $\|\cdot\|$ is a lattice norm iff it is a *monotone* norm such that $\||v|\| = \|v\|$ for every $v \in V$.
(ii) Prove that, for any normed lattice V, the mapping $|\cdot| : V \to V$ is continuous.
(iii) Show that if $\|\cdot\|$ is a lattice norm, then $(V, \|\cdot\|)$ is an ordered normed space, i.e., the positive cone $V^+ \subseteq V$ is closed with respect to the norm $\|\cdot\|$(Definition 1.32). In particular, V is Archimedean. See Exercise 1.33.

(iv) Prove that if V is Archimedean (and has order units), then, for any order unit $u \in V$, the norm $\|\cdot\|_u$ (Definition 1.41) is a lattice norm.

Combining the last exercise with Lemma 1.34 and Proposition 1.43, we arrive at the following statement:

Corollary 7.318 *Let V_1 be any Banach lattice and V_2 any Riesz space having an order unit. Then $\mathcal{B}(V_1; V_2)^+ = \mathcal{L}(V_1; V_2)^+$. In particular, for any Banach lattice V, one has*

$$V^{\mathrm{n}} \subseteq V^{\mathrm{c}} \subseteq V^{\mathrm{od}} \subseteq V^{\mathrm{td}} \doteq \mathcal{B}(V; \mathbb{R}) .$$

Proof Exercise. □

Recall that, for any Archimedean Riesz space V (thus, in particular, for normed lattices), $V^{\mathrm{c}} \subseteq V^{\mathrm{od}}$.

7.4.7 Some Remarks on the Relations Between Riesz Spaces and C-Algebras*

Note that C^*-algebras are studied in details in Chap. 4. If a given (unital) C^*-algebra $\mathcal{A}$ (see, e.g., Definition 7.85) is *commutative*, then $\mathrm{Re}\{\mathcal{A}\}$ is equivalent, as a preordered vector space, to a function space. This fact is discussed and proved in detail in Sect. 4.7. In particular, $\mathrm{Re}\{\mathcal{A}\}$ is a Riesz space, in this case. It turns out that $\mathrm{Re}\{\mathcal{A}\}$ being a Riesz space is equivalent to the commutativity of $\mathcal{A}$:

Theorem 7.319 (Sherman) *Let $\mathcal{A}$ be any (not necessarily unital) C^*-algebra. $\mathcal{A}$ is commutative iff* $\mathrm{Re}\{\mathcal{A}\}$ *is a Riesz space.*

This is a deep result of C^*-algebra theory, and a proof of this result can be found in [6]. Thus, from a physical point of view, a system is classical iff the space of its observables is a Riesz space.

Another interesting fact about Riesz structures in C^*-algebras is the following observations: For any two commutative unital C^*- *subalgebras* $\mathcal{A}'$ and $\mathcal{A}''$ of some (not necessarily commutative) C^*-algebra $\mathcal{A}$, if $\mathcal{A}'' \subseteq \mathcal{A}'$, then $\mathrm{Re}\{\mathcal{A}''\}$ is a Riesz subspace of $\mathrm{Re}\{\mathcal{A}'\}$ (and not only a vector subspace which is itself a Riesz space). This fact is a direct consequence of the following property of commutative C^*-algebras:

Lemma 7.320 *Let $\mathcal{A}$ be a commutative C^*-algebra. For any pair of real elements $A, A' \in \mathrm{Re}\{\mathcal{A}\}$, the supremum $A \vee A' \in \mathcal{A}$ is an element of the smallest C^*-subalgebra of $\mathcal{A}$ containing both A and A'.*

Proof Since $\mathcal{A}$ is a commutative C^*-algebra, recall that it is equivalent to a function space (in the sense of Definition 7.270), by Corollary 4.126. Therefore, the supremum $A \vee A' \in \mathcal{A}$ exists for any pair of real elements $A, A' \in \text{Re}\{\mathcal{A}\}$. By simple properties of ordered vector spaces,

$$A \vee A' = (A - A') \vee 0 + A' , \qquad A, A' \in \text{Re}\{\mathcal{A}\} .$$

Thus, it suffices to prove that $(A-A')\vee 0$ is an element of the smallest C^*-subalgebra of $\mathcal{A}$ containing A and A'. By Corollary 4.126 together with Lemma 4.107 (ii),

$$(A - A') \vee 0 = f(A - A') , \qquad A, A' \in \text{Re}\{\mathcal{A}\} ,$$

where f is the continuous function $f(s) = (|s| + s)/2$, $s \in \sigma(A - A')$, on the spectrum $\sigma(A - A') \subseteq \mathbb{R}$ of $A - A'$. As $f(0) = 0$, the assertion then follows from Lemma 4.106. □

Recall that commutative unital C^*-subalgebras of a unital C^*-algebra $\mathcal{A}$ associated with a given physical system are called the "classical contexts" of this system. This fact allows one, for instance, to see the classical contexts of a (generally non-classical) physical system as a (poset) category of Riesz spaces. This kind of construction is possibly relevant in the study of foundations of quantum theory. Observe that the view on quantum systems as (poset) categories of commutative C^*-algebras is being exploited in research on mathematical foundations of quantum theory since a couple of decades. See, for instance, [7] for a comprehensive review on this subject.

By Sherman's theorem, for a given C^*-algebra $\mathcal{A}$, $\text{Re}\{\mathcal{A}\}$ is strictly speaking a Riesz space only if $\mathcal{A}$ is commutative. However, even in the non-commutative case, $\text{Re}\{\mathcal{A}\}$ is not that far, in a sense, from being a Riesz space. One first strong relation between $\text{Re}\{\mathcal{A}\}$, where $\mathcal{A}$ is an arbitrary (i.e., not necessarily commutative) C^*-algebra, and a Riesz space is established by the following well-known result:

Theorem 7.321 *Every Archimedean real ordered vector space that is directed is equivalent (as a preordered vector space) to an* order-dense *subspace of an Archimedean Riesz space. This Riesz space is unique up to an isomorphism of preordered vector spaces. See Definitions 1.10 and 1.19.*

See [8, Theorem 4.3] and related references therein.

For any C^*-algebra $\mathcal{A}$, $\text{Re}\{\mathcal{A}\}$ satisfies the assumptions of the last theorem. In particular, $\text{Re}\{\mathcal{A}\}$ can always be seen as an order-dense vector subspace of a (up to isomorphisms of preordered vector spaces) unique Archimedean Riesz space.

A second strong relation between the order structure of any (unital) C^*-algebra and Riesz spaces is as follows: In an arbitrary C^*-algebra $\mathcal{A}$, any positive element $A \in \mathcal{A}^+$ has a unique "positive square root," that is, there is a unique $\sqrt{A} \in \mathcal{A}^+$ such that $A = (\sqrt{A})^2$. See, e.g., Proposition 4.102 and [51, Theorem 2.2.10]. Moreover, it turns out that, for any $A \in \text{Re}\{\mathcal{A}\}$, $A^2 \in \mathcal{A}^+$. See Corollary 4.103. These two facts can be used to introduce (a version of) the absolute value for elements of $\text{Re}\{\mathcal{A}\}$ and then define from it binary operations in $\text{Re}\{\mathcal{A}\}$ capturing as much as possible from the properties of usual lattice operations. The absolute value of any $A \in \text{Re}\{\mathcal{A}\}$ is, by definition, the positive element $|A| \doteq \sqrt{A^2}$. In fact, for a given $A \in \mathrm{Re}\{\mathcal{A}\}$, $|A|$ is the unique positive element of $\mathcal{A}$ for which $A^2 = |A|^2$. With this definition, the absolute value in a C^*-algebra has various properties in common with the absolute value of a vector lattice. For instance, for any $A \in \mathrm{Re}\{\mathcal{A}\}$ and $\alpha \in \mathbb{R}$, one has that $|\alpha A| = |\alpha||A|$ and

$$A^{\pm} \doteq \frac{1}{2}(|A| \pm A) \in \mathcal{A}^+ .$$

Observe that, as in the Riesz space case (Definition 7.278), one has that

$$A = A^+ - A^-, \qquad |A| = A^+ + A^- .$$

See, e.g., Proposition 4.102.

Thus, inspired by Exercise 7.280 (v), we define two binary operations $\curlywedge, \curlyvee$: $\mathrm{Re}\{\mathcal{A}\} \times \mathrm{Re}\{\mathcal{A}\} \to \mathrm{Re}\{\mathcal{A}\}$ as follows: For all $A, A' \in \mathrm{Re}\{\mathcal{A}\}$,

$$A \curlywedge A' \doteq \frac{1}{2}(A + A' - |A - A'|) ,$$
$$A \curlyvee A' \doteq \frac{1}{2}(A + A' + |A - A'|) .$$

These operations (trivially) satisfy, for all $A, A', A'' \in \mathrm{Re}\{\mathcal{A}\}$ and $\alpha \geq 0$, the following identities (compare them with those in Exercise 7.272):

1. $A \curlywedge A' = A' \curlywedge A$,
2. $A \curlyvee A' = A' \curlyvee A$,
3. $A + A' = A \curlywedge A' + A \curlyvee A'$,
4. $A^+ = 0 \curlyvee A$,
5. $A^- = 0 \curlywedge (-A)$,
6. $-A \curlyvee A' = (-A') \curlywedge (-A)$,
7. $-A \curlywedge A' = (-A') \curlyvee (-A)$,
8. $\alpha(A \curlyvee A') = (\alpha A') \curlyvee (\alpha A)$,
9. $\alpha(A \curlywedge A') = (\alpha A') \curlywedge (\alpha A)$,
10. $A'' + A \curlywedge A' = (A'' + A) \curlywedge (A'' + A')$,
11. $A'' + A \curlyvee A' = (A'' + A) \curlyvee (A'' + A')$.

Additionally, the operation $\curlywedge$ ($\curlyvee$) refers to minimal (maximal), not necessarily least (largest), upper (lower) bounds of two-point sets:

Theorem 7.322 *Let $\mathcal{A}$ be any (unital) C^*-algebra. For every $A, A', A'' \in \mathrm{Re}\{\mathcal{A}\}$ such that $A, A' \geq A''$ ($A'' \geq A, A'$), one has that $A'' \geq A \curlywedge A'$ ($A \curlyvee A' \geq A''$) only if $A'' = A \curlywedge A'$ ($A'' = A \curlyvee A'$).*

The proof of this theorem is quite involved and will not be discussed here. The theorem is only stated here for the sake of completeness of the present discussion. For a complete proof,[12] see [9, Corollary 6.10].

If $\mathcal{A}$ is commutative, then $\curlywedge$ ($\curlyvee$) coincide with the lattice operation $\wedge$ ($\vee$) in the Riesz space $\mathrm{Re}\{\mathcal{A}\}$. In fact, for any classical context $\mathcal{A}'$ of $\mathcal{A}$, $\curlywedge$ and $\curlyvee$ preserve $\mathrm{Re}\{\mathcal{A}'\}$ and coincide with the lattice operation $\wedge$ and $\vee$ of $\mathrm{Re}\{\mathcal{A}'\}$:

Lemma 7.323 *Let $\mathcal{A}$ be any C^*-algebra. For any pair of real elements $A, A' \in \mathrm{Re}\{\mathcal{A}\}$, $A \curlyvee A'$, $A \curlywedge A' \in \mathrm{Re}\{\mathcal{A}\}$ are elements of the smallest C^*-subalgebra of $\mathcal{A}$ containing both A and A'. If A, A' belong to some classical context $\mathcal{A}'$ of $\mathcal{A}$, i.e., $\mathcal{A}'$ is a commutative subalgebra of $\mathcal{A}$, then*

$$A \curlyvee A' = A \vee A' \qquad \textit{and} \qquad A \curlywedge A' = A \wedge A' \,,$$

where $\vee$ and $\wedge$ are the lattice operations of the Riesz space $\mathrm{Re}\{\mathcal{A}'\}$.

Proof The fact that $A \curlyvee A'$ and $A \curlywedge A'$ are real elements of $\mathcal{A}$ directly follows from the definition of the operations $\curlyvee$ and $\curlywedge$. From Lemma 4.106, we deduce that $A \curlyvee A'$ and $A \curlywedge A'$ are elements of the smallest C^*-subalgebra of $\mathcal{A}$ containing A and A'. Assume now that A and A' are in some classical context $\mathcal{A}'$ of $\mathcal{A}$. By Properties 10 and 11 of the operations $\curlyvee$ and $\curlywedge$, given above, one has that

$$A \curlyvee A' = A' + (A - A') \curlyvee 0 \qquad \text{and} \qquad A \curlywedge A' = A' + (A - A') \curlywedge 0 \,.$$

By combining the above identities with the corresponding ones for the lattice operations $\vee$ and $\wedge$, in order to prove that $A \curlyvee A' = A \vee A'$ and $A \curlywedge A' = A \wedge A'$, it suffices to show that

$$(A - A') \curlyvee 0 = (A - A') \vee 0 \qquad \text{and} \qquad (A - A') \curlywedge 0 = (A - A') \wedge 0 \,.$$

By the proof of Lemma 7.320, the first identity follows from the observation that

$$(A - A') \curlyvee 0 = f(A' - A) \,,$$

where f is the continuous function $f(s) = (|s| + s)/2$, $s \in \sigma(A - A')$, on the spectrum $\sigma(A - A') \subseteq \mathbb{R}$ of $A - A'$. The second identity then follows from Property 7 of the operations $\curlyvee$ and $\curlywedge$, together with the corresponding ones for the lattice operations $\vee$ and $\wedge$. □

To close this subsection we state an important result of Kadison's [100] in this context, which says that the (C^*-)algebras $\mathcal{B}(H)$ of bounded operators on complex

[12] This proof refers to concrete C^*-algebras, but, of course, easily extends to general ones, recalling that any C^*-algebra is $*$-isomorphic to a concrete one, by the Gelfand-Naimark theorem (Theorem 4.89).

Hilbert spaces H are, in a sense, the ones that are the poorest, in which regards the (quasi-)lattice structure:

Proposition 7.324 (Kadison's Anti-Lattice Theorem) *Let H be any complex Hilbert space and take any to self-adjoint operators $A, A' \in \mathcal{B}(H)^{\mathbb{R}}$. Then, $\{A, A'\}$ has a supremum in $\mathcal{B}(H)^{\mathbb{R}}$ iff these elements are comparable, i.e., $A \geq A'$ or $A' \geq A$.*

Proof See [100]. □

7.5 Convexity

In this section, we present various results related to convexity that are essential in our study of equilibrium states of quantum lattices. For most of them, we provide complete proofs. However, in some cases, we only give a proof for a weaker version of the full result and refer to the literature for the general proof. In fact, the presentation of complete proofs would, in some cases, largely exceed the scope of the book. In fact, the proofs of simple versions of these cases serve as a pedagogical illustration, which can be hopefully instructive for the non-expert. For instance, this is the case of Choquet's theorem (Theorem 7.339), representing a stronger version of Corollary 7.338 for which a proof is given here. Standard materials like the Hahn-Banach (separation) theorem (Theorem 7.331) are also stated without proofs.

7.5.1 Basic Notions

We start with the standard notion of convexity for sets:

Definition 7.325 (Convex Sets) Let V be any real vector space. The set $M \subseteq V$ is "convex" when for all $v, v' \in M$ and $\alpha \in [0, 1]$,

$$\alpha v + (1 - \alpha) v' \in M \ .$$

Definition 7.326 (Convex Hull) Let V be any real vector space and $M \subseteq V$ any nonempty set. $\mathrm{co}(M) \subseteq V$ denotes the smallest convex subset of V that contains M. Similarly, if V is a metric vector space, $\overline{\mathrm{co}}(M) \subseteq V$ denotes the smallest *closed* convex subset of V that contains M. $\mathrm{co}(M)$ and $\overline{\mathrm{co}}(M)$ are, respectively, called the "convex hull" and the "closed convex hull" of M.

Observe that V is (trivially) a convex subset of V that contains $M \subseteq V$ and that the intersection of any family of subsets with these properties (convexity and the inclusion of M) is again a subset of V with the same properties. Therefore, $\mathrm{co}(M) \subseteq V$ is well-defined. Any arbitrary intersection of closed subsets is closed and so, the closed convex hull $\overline{\text{co}}(M) \subseteq V$ is also well-defined. The following exercise provides a more constructive definition for $\text{co}(M)$ and $\overline{\text{co}}(M)$:

Exercise 7.327 Let V be any real vector space and $M \subseteq V$ any nonempty set. Show that

$$\text{co}(M) = \{\alpha_1 v_1 + \cdots + \alpha_n v_n \ : \ \alpha_1, \ldots, \alpha_n \in \mathbb{R}_0^+,\ \alpha_1 + \cdots + \alpha_n = 1,\ n \in \mathbb{N}\} .$$

If V is a metric vector space, show additionally that $\overline{\text{co}}(M)$ is nothing else than the closure of $\text{co}(M)$.

We have the following notion of convexity of real-valued functions, which is directly related to the notion of convex sets:

Definition 7.328 (Convex Functions) Let V be any real vector space and $M \subseteq V$ a convex set. A function $f : M \to \mathbb{R}$ is "convex" if, for all $v, v' \in M$ and $\alpha \in [0, 1]$,

$$f(\alpha v + (1-\alpha)v') \leq \alpha f(v) + (1-\alpha) f(v') .$$

We say that f is "concave" if $-f$ is convex. Finally, f is said to be "affine" whenever it is simultaneously convex and concave.

By iterating the inequality in the above definition, note that, for any convex function $f : M \to \mathbb{R}$, one has that

$$f(\alpha_1 v_1 + \cdots + \alpha_n v_n) \leq \alpha_1 f(v_1) + \cdots + \alpha_n f(v_n)$$

for all $n \in \mathbb{N}$, $v_1, \ldots, v_n \in \mathbb{R}$ and all $\alpha_1, \ldots, \alpha_n \in \mathbb{R}_0^+$ satisfying $\alpha_1 + \cdots + \alpha_n = 1$. Additionally, the convexity of a function is nothing else than the convexity of its epigraph (Definition 7.147):

Exercise 7.329 Let V be any real vector space and $M \subseteq V$ a convex set. Show that the function $f : M \to \mathbb{R}$ is convex iff its epigraph

$$\text{epi}(f) \doteq \{(v, s) : \ v \in M\ ,\ s \geq f(v)\} \subseteq V \times \mathbb{R}$$

is a convex set.

The following bound is an important, albeit simple, instance of the convexity of functions $f : \mathbb{R} \to \mathbb{R}$:

Lemma 7.330 (Jensen's Inequality) *Let M be any nonempty set and V a subspace of the preordered (real) vector space $\mathcal{F}(M; \mathbb{R})$ that contains the constant functions. Take any normalized (i.e., $\varphi(1) = 1$) positive linear functional $\varphi \in V'^{+}$. For any convex function $f : \mathbb{R} \to \mathbb{R}$ and any $g \in V \subseteq \mathcal{F}(M; \mathbb{R})$ for which $f \circ g \in V$, one has that*

$$f(\varphi(g)) \leq \varphi(f \circ g) .$$

Proof Recall that any convex function $f : \mathbb{R} \to \mathbb{R}$ is continuous and, thus, for all $x_0 \in \mathbb{R}$, there is a constant $\alpha_0 \in \mathbb{R}$ such that

$$f(x) - f(x_0) \geq \alpha_0(x - x_0) \,, \qquad x \in \mathbb{R} \,.$$

In other words, α_0 defines a tangent of f at x_0. Note that the existence of α_0 is a simple application of Proposition 3.21 (i). In fact, in this simple case, one can directly check that, by convexity of f, the following monotone limit exists and has the required property:

$$\alpha_0 = \lim_{x \downarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \,.$$

In other words, one may take α_0 as being the "right derivative" of f at x_0, which is well-defined for any convex function $f : \mathbb{R} \to \mathbb{R}$. Choose $\varphi \in V'^{+}$ with $\varphi(1) = 1$ and $g \in V$ such that $f \circ g \in V$. Let $x_0 = \varphi(g)$ and take the corresponding $\alpha_0 \in \mathbb{R}$ as above. Then, for all $p \in M$,

$$f \circ g(p) - f(\varphi(g)) \geq \alpha_0 g(p) - \alpha_0 \varphi(g) \,,$$

that is,

$$f \circ g - f(\varphi(g)) \geq \alpha_0 g - \alpha_0 \varphi(g)$$

in the preordered vector space V (whose order relation is inherited from $\mathcal{F}(M; \mathbb{R})$). As $\varphi(1) = 1$, by positivity of φ, we arrive at

$$\varphi(f \circ g) - f(\varphi(g)) \geq \alpha_0 \varphi(g) - \alpha_0 \varphi(g) = 0 \,.$$

□

Observe that if we take $n \in \mathbb{N}$, $M = \{1, \ldots, n\}$ and $V = \mathcal{F}(M; \mathbb{R})$, then the above lemma is just saying that

$$f(\alpha_1 x_1 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n)$$

for all $x_1, \ldots, x_n \in \mathbb{R}$ and all $\alpha_1, \ldots, \alpha_n \in \mathbb{R}_0^+$ with $\alpha_1 + \cdots + \alpha_n = 1$. Hence, Jensen's inequality can be seen as a kind of infinitary version of the usual convexity property of functions $\mathbb{R} \to \mathbb{R}$.

In metric vector spaces, it turns out that disjoint convex sets can be separated by *continuous* linear functionals:

Theorem 7.331 (Hahn-Banach Separation Theorem) *Let* (V, d) *be any metric vector space over* $\mathbb{R}$ *and let* $M, M' \subseteq V$ *be two nonempty* disjoint convex *sets.*

(i) If M is open *, then there are a continuous linear functional $\varphi \in V^{\mathrm{td}}$ and a constant $\alpha \in \mathcal{R}$, such that*

$$\varphi(v) < \alpha \leq \varphi(v') , \qquad v \in M , \ v' \in M' .$$

(ii) If the metric vector space (V, d) is locally convex (Definition 7.117), M is compact, and M' is closed, then there are a continuous linear functional $\varphi \in V^{\mathrm{td}}$ and two constants $\alpha^{\flat}, \alpha^{\sharp} \in \mathbb{R}$, such that

$$\varphi(v) \leq \alpha^{\flat} < \alpha^{\sharp} \leq \varphi(v') , \qquad v \in M , \ v' \in M' .$$

Proof For a proof of the theorem, see, for instance, [18, 3.4 Theorem]. □

As one-point sets are (trivially) convex and compact in any metric vector space, we have the following simple, though important, corollary of the last theorem:

Corollary 7.332 (Continuous Linear Functions Separate Points) *Let V be any metric vector space over $\mathbb{R}$. If V is locally convex, then, for any $v, v' \in V$, $v \neq v'$, there is a continuous linear functional $\varphi \in V^{\mathrm{td}}$ such that $\varphi(v) < \varphi(v')$.*

7.5.2 The Extreme Boundary of Convex Sets

In the present subsection, we introduce the notion of extremality in connection with the convexity property and prove an important (well-known) result related to extreme points of convex sets: the Krein-Milman theorem (Proposition 7.334).

Definition 7.333 (Extreme Subsets and Points) Let V be any real vector space and $M_1 \subseteq V$ a nonempty set. $M_2 \subseteq M_1$ is an "extreme subset" of M_1 whenever, for all $v_1, v_1' \in M_1$ and $\alpha \in (0, 1)$, such that

$$\alpha v_1 + (1 - \alpha) v_1' \in M_2 ,$$

one has that $v_1, v_1' \in M_2$. $v_1 \in M_1$ is an "extreme point" of M_1 if $\{v_1\} \subseteq M_1$ is an extreme subset of M_1. $\mathcal{E}(M_1) \subseteq M_1$ denotes the (possibly empty) set of all extreme points of M_1, whereas $\mathcal{ES}(M_1) \subseteq \mathcal{P}(M_1)$ denotes the collection of all extreme subsets of M_1. Observe that $\mathcal{E}(M_1)$ is naturally seen as a subset of $\mathcal{ES}(M_1)$, by identifying $v_1 \in \mathcal{E}(M_1)$ with $\{v_1\} \in \mathcal{ES}(M_1)$. The subset $\mathcal{E}(M_1) \subseteq M_1$ is called the "extreme boundary" of M_1. The elements of $\mathcal{ES}(M_1)$ are also called "faces" of M_1.

In turns out that any nonempty compact set in a locally convex space (see Definition 7.117 for the metric case) always have extreme points:

Proposition 7.334 (Krein-Milman Theorem for Metric Locally Convex Spaces) *Let (V, d) be any metric vector space over $\mathbb{R}$ and $K \subseteq V$ a nonempty compact set.*

If (V, d) is a locally convex space, then $\mathcal{E}(K)$ is nonempty, and $K \subseteq \overline{\mathrm{co}}(\mathcal{E}(K))$. In particular, if K is convex, then one has the equality $K = \overline{\mathrm{co}}(\mathcal{E}(K))$.

Proof

1. Let $\mathcal{CES}(K)$ denote the family of all nonempty closed extreme subsets (or faces) of K. Remark that $\mathcal{CES}(K) \neq \emptyset$, because $K \in \mathcal{CES}(K)$. We consider the partial order relation in $\mathcal{CES}(K)$ given by the reverse inclusion, that is, $M \leq M'$ if $M' \subseteq M$. Let $\mathcal{C} \subseteq \mathcal{CES}(K)$ be any "chain" (or totally ordered family), i.e., for any pair $M, M' \in \mathcal{C}$, either $M \leq M'$ or $M' \leq M$. Then one has that
$$\cap\, \mathcal{C} \in \mathcal{CES}(K)\ .$$
In fact, $\cap\mathcal{C}$ is closed, being the intersection of a family of closed sets. Additionally, by Lemma 7.160 combined with Proposition 7.181, $\cap\mathcal{C}$ is nonempty, K being compact, and the intersection operation preserves the property of subsets being extreme. Thus, by Zorn's lemma, the partially ordered set $\mathcal{CES}(K)$ has at least one maximal element $K_\infty \in \mathcal{CES}(K)$, that is, for any $K' \in \mathcal{CES}(K)$, one has that $K' \subseteq K_\infty$ only if $K' = K_\infty$.
2. We now show that such a maximal element must be of the form $K_\infty = \{v_\infty\}$ for some $v_\infty \in K$. In particular, $v_\infty \in \mathcal{E}(K)$ and $\mathcal{E}(K)$ is thus nonempty. Assume, by contradiction, that there are two vectors $v_\infty, v'_\infty \in K_\infty$, $v_\infty \neq v'_\infty$. Then, as V is locally convex, by Corollary 7.332, there is a continuous linear functional $\varphi \in V^{\mathrm{td}}$ such that $\varphi(v'_\infty) < \varphi(v_\infty)$. Define
$$K_\infty(\varphi) \doteq \{v \in K_\infty \ :\ \varphi(v) = \sup_{v' \in K_\infty} \varphi(v')\} \subseteq K_\infty\ .$$
As φ is continuous and K_∞ is compact, by Proposition 7.172, $K_\infty(\varphi)$ is nonempty and compact (thus, closed). By linearity of φ, $K_\infty(\varphi) \in \mathcal{CES}(K)$. Now, note that, by construction, $v'_\infty \notin K_\infty(\varphi)$ and, hence, $K_\infty(\varphi) \subsetneq K_\infty$. This would contradict the maximality of K_∞. Thus, K_∞ must contain exactly one unique point $v_\infty \in K$ and $\mathcal{E}(K)$ is thus nonempty.
3. Assume, again by contradiction, that $K \nsubseteq \overline{\mathrm{co}}(\mathcal{E}(K))$, that is, there is $\tilde{v} \in K$ with $\tilde{v} \notin \overline{\mathrm{co}}(\mathcal{E}(K))$. Then, noting again that one-point sets are trivially convex and compact, by Theorem 7.331, there is $\varphi \in V^{\mathrm{td}}$ such that, for all $v \in \overline{\mathrm{co}}(\mathcal{E}(K))$,
$$\varphi(v) < \varphi(\tilde{v})\ .$$
Exactly as in Point 2,
$$K(\varphi) \doteq \{v \in K \ :\ \varphi(v) = \sup_{v' \in K} \varphi(v')\} \in \mathcal{CES}(K)\ ,$$
but $K(\varphi) \cap \overline{\mathrm{co}}(\mathcal{E}(K)) = \emptyset$. Again by Zorn's lemma, there is a maximal element $\tilde{K} \in \mathcal{CES}(K)$ with $\tilde{K} \subseteq K(\varphi)$. But this would imply the existence of an

extreme point of K not contained in $\overline{\text{co}}(\mathcal{E}(K))$, which is clearly wrong. Thus, $K \subseteq \overline{\text{co}}(\mathcal{E}(K))$. Finally, if K is not only compact (thus closed) but also convex, then

$$\overline{\text{co}}(\mathcal{E}(K)) \subseteq \overline{\text{co}}(K) = K \subseteq \overline{\text{co}}(\mathcal{E}(K)) \ .$$

□

Observe that the above proposition holds true also for topological vector spaces that are not necessarily metric or locally convex. See, for instance, [18, 3.23 Theorem]. The above version of the Krein-Milman theorem is sufficient for our purposes.

7.5.3 Barycenters of States of Unital C-Algebras*

In the present subsection, we introduce the notion of "barycenter" of a probability measure in a convex set. This notion can be thought of as being an infinitary version of convex combinations of points in a convex set. To keep the discussions as simple as possible, we only consider barycenters within some weak*-closed convex set of states of an arbitrary unital separable C^*-algebra, which is the case of interest here. However, with rather simple adaptations, the constructions given below can also be used for more general convex sets of linear functionals.

Definition 7.335 (Barycenters of States of Unital C^*-Algebras) Let $\mathcal{A}$ be a separable unital C^*-algebra. Take any nonempty weak*-closed subset $E \subseteq E(\mathcal{A})$ of states on $\mathcal{A}$ and a state $\mu \in E(C(E;\mathbb{C}))$[13] on the C^*-algebra of continuous functions on the weak*-compact[14] set E. We then define a state $\rho_\mu \in E(\mathcal{A})$ on the original C^*-algebra $\mathcal{A}$ by

$$\rho_\mu \doteq \mu \circ \Xi \ ,$$

where Ξ is the Gelfand transform[15] of Definition 4.79. The state ρ_μ is called the "barycenter" of (the probability measure) μ. If $\rho \in E(\mathcal{A})$ is the barycenter of some $\mu \in E(C(E;\mathbb{C}))$, then we say that (the probability measure) μ represents the state ρ.

[13] Recall that states of the C^*-algebra of continuous functions on a compact metric space are nothing else than probability measures. See Definition 4.10 and Theorem 4.68, as well as the related discussions.

[14] See Proposition 4.84.

[15] By definition, Ξ is a mapping $\mathcal{A} \to C(E(\mathcal{A});\mathbb{C})$. It is naturally seen here as a mapping $\mathcal{A} \to C(E;\mathbb{C})$, by restriction of continuous functions on $E(\mathcal{A})$ to E.

Note here that, by Proposition 4.78 (i), the Gelfand transform Ξ is a positive linear transformation and, thus, $\rho_\mu \doteq \mu \circ \Xi$ is a positive linear functional on $\mathcal{A}$. As it is (trivially) normalized, i.e., $\rho_\mu(\mathfrak{1}) = 1$, it is a state on $\mathcal{A}$. See Definition 4.62.

Observe from Lemma 4.6 that the C^*-algebra of continuous functions on a compact metric space is always separable. In particular, $C(E;\mathbb{C})$ in the last definition is a separable C^*-algebra. Thus, we can provide the space of states $E(C(E;\mathbb{C}))$ with a metric, whose associated topology is the weak * topology for linear functions. See Sect. 4.5.1. The barycenter mapping $\mu \mapsto \rho_\mu$ is a weak*-continuous transformation $E(C(E;\mathbb{C})) \to E(\mathcal{A})$:

Exercise 7.336 Let $\mathcal{A}$ be a separable unital C^*-algebra. Show that for any sequence (of probability measures) $\mu_n \in E(C(E;\mathbb{C}))$, $n \in \mathbb{N}$, converging in the weak* topology to some $\mu \in E(C(E;\mathbb{C}))$, one has that the corresponding sequence of states $\rho_{\mu_n} \in E(\mathcal{A})$ on $\mathcal{A}$ converges in the weak* topology to $\rho_\mu \in E(\mathcal{A})$.

In the next proposition, we establish a relation between the barycenter of a probability measure and the support (see Definition 4.11) of this measure:

Proposition 7.337 *Let $\mathcal{A}$ be a separable unital C^*-algebra and $\tilde{E}, E \subseteq E(\mathcal{A})$ nonempty closed subsets with $\tilde{E} \subseteq E$. If $\rho \in \overline{\mathrm{co}}(\tilde{E})$, then ρ is the barycenter of some $\mu \in E(C(E;\mathbb{C}))$, whose unique extension to a probability measure (i.e., a state on the C^*-algebra $\mathfrak{M}_{\mathrm{b}}(E;\mathbb{C})$ that is an integral; see Definition 4.10 and Theorem 4.68) is supported in the set $\tilde{E}$.*

Proof

1. Take any state $\rho \in \overline{\mathrm{co}}(\tilde{E})$. Then there is a sequence $(\rho_n)_{n\in\mathbb{N}} \subseteq \mathrm{co}(\tilde{E})$ such that

$$\rho = \lim_{n\to\infty} \rho_n$$

in the weak* topology of $E(\mathcal{A})$. Note that $\rho_n \in \mathrm{co}(\tilde{E})$ for $n \in \mathbb{N}$ means that

$$\rho_n = \alpha_1^{(n)} \rho_1^{(n)} + \cdots + \alpha_{k_n}^{(n)} \rho_{k_n}^{(n)} , \qquad k_n \in \mathbb{N} ,$$

for some states $\rho_1^{(n)}, \ldots, \rho_{k_n}^{(n)} \in \tilde{E}$ and positive constants $\alpha_1^{(n)}, \ldots, \alpha_{k_n}^{(n)} \geq 0$ satisfying $\alpha_1^{(n)} + \cdots + \alpha_{k_n}^{(n)} = 1$. For any $n \in \mathbb{N}$ and $k \in \{1, \ldots, k_n\}$, define the probability measure $\mu_k^{(n)} \in E(\mathfrak{M}_{\mathrm{b}}(E;\mathbb{C}))$ by

$$\mu_k^{(n)}(f) \doteq f(\rho_k^{(n)}) , \qquad f \in \mathfrak{M}_{\mathrm{b}}(E) .$$

See Exercise 4.66. By construction, for every $n \in \mathbb{N}$, ρ_n is the barycenter of (the restriction to $C(E;\mathbb{C}) \subseteq \mathfrak{M}_{\mathrm{b}}(E;\mathbb{C})$)

$$\mu_n \doteq \alpha_1^{(n)} \mu_1^{(n)} + \cdots + \alpha_{k_n}^{(n)} \mu_{k_n}^{(n)} \in E(\mathfrak{M}_{\mathrm{b}}(E)) .$$

Clearly, all the probability measures $\mu_n \in E(\mathfrak{M}_b(E;\mathbb{C}))$, $n \in \mathbb{N}$, have support in $\tilde{E}$. As $E(C(E;\mathbb{C}))$ is weak*-compact, we can assume that the restriction of $\mu_n \in E(\mathfrak{M}_b(E))$ to $C(E;\mathbb{C})$ converges in the weak* topology to some $\tilde{\mu}_\infty \in E(C(E;\mathbb{C}))$.

2. By construction, ρ is the barycenter of $\tilde{\mu}_\infty$. Let $\mu_\infty \in E(\mathfrak{M}_b(E;\mathbb{C}))$ be the unique extension of $\tilde{\mu}_\infty$ to a probability measure (see Theorem 4.68). We show now that also μ_∞ is supported in $\tilde{E}$. With this aim, let d be any metric in $E(C(E;\mathbb{C}))$ whose associated topology is the weak* one (see Definition 4.80). Then, for all $n \in \mathbb{N}$, define the continuous function $f_n \in C(E;\mathbb{C})$ by

$$f_n(\rho) = \frac{nd(\rho, \tilde{E})}{1 + nd(\rho, \tilde{E})} \in [0, 1] , \qquad \rho \in E ,$$

where

$$d(\rho, \tilde{E}) \doteq \inf\{d(\rho, \rho') \ : \rho' \in \tilde{E}\} .$$

Note that this sequence is monotonically increasing and the pointwise limit f_∞ of the sequence $(f_n)_{n\in\mathbb{N}}$, as defined by

$$f_\infty(\rho) = \lim_{n\to\infty} f_n(\rho) , \qquad \rho \in E ,$$

is nothing else than $\chi_{E\backslash\tilde{E}}$, the characteristic function of the subset $E\backslash\tilde{E} \subseteq E$. In particular, as $\tilde{E}$ is by assumption closed, $\chi_{E\backslash\tilde{E}} \in \mathfrak{M}_b(E;\mathbb{C})$ (see Definition 4.8), and, as probability measures are by definition σ-order-continuous, one has that

$$\mu_\infty(\chi_{E\backslash\tilde{E}}) = \lim_{n\to\infty} \mu_\infty(f_n) = \lim_{n\to\infty} \lim_{m\to\infty} \mu_m(f_n) .$$

But $\mu_m(f_n) = 0$, because $f_n(\tilde{E}) = \{0\}$ and μ_m is supported in $\tilde{E}$. Thus, $\mu_\infty(\chi_{E\backslash\tilde{E}}) = 0$. Take any bounded measurable function $g \in \mathfrak{M}_b(E;\mathbb{C})$ satisfying $g|_{\tilde{E}} = 0$. Then, $(1 - \chi_{E\backslash\tilde{E}})g = 0$ and, consequently,

$$\mu_\infty(g) = \mu_\infty(g\chi_{E\backslash\tilde{E}}) .$$

From the Cauchy-Schwarz for states, we then arrive at

$$|\mu_\infty(g)|^2 \leq \mu_\infty(|g|^2)\mu_\infty(\chi_{E\backslash\tilde{E}}) = 0 .$$

Thus, μ_∞ is supported in $\tilde{E}$.

□

By combining Propositions 7.334 (Krein-Milman theorem) and 7.337, we arrive at the following important corollary, which can be seen as a weaker version of Choquet's theorem (Theorem 7.339) stated later on:

Corollary 7.338 (Krein-Milman Theorem—Barycenter Version) *Let $\mathcal{A}$ be a separable unital C^*-algebra and $E \subseteq E(\mathcal{A})$ a nonempty closed* convex *set of states on $\mathcal{A}$. Every $\rho \in E$ is the barycenter of some $\mu \in E(C(E;\mathbb{C}))$, whose unique extension to a probability measure is supported in $\overline{\mathcal{E}(E)}$, the closure of the extreme boundary of the convex set E.*

The above corollary is indeed useful in many situations, but in statistical mechanics, it frequently happens that $\overline{\mathcal{E}(E)} = E$, in which case the corollary is trivial. In fact, note that the sets E_1 of invariant states on the algebras of quantum spins (fermions) $\mathrm{Spin}(N,\Gamma)$ $(\mathrm{CAR}(\Omega,\Gamma))$ of Sect. 5.1 are so-called Poulsen simplices. In particular, they have a (weak*-)dense extreme boundary. See Theorem 6.1 and related discussions. Therefore, we state below a stronger version of the last corollary saying that probability measures representing states can be chosen as being supported in $\mathcal{E}(E)$ and not only in the closure $\overline{\mathcal{E}(E)}$. This refers to the celebrated Choquet's theorem, in the case of convex sets of states on C^*-algebras:

Theorem 7.339 (Choquet's Theorem—Version for States) *Let $\mathcal{A}$ be a separable unital C^*-algebra and $E \subseteq E(\mathcal{A})$ a nonempty closed* convex *set of states on $\mathcal{A}$. Every $\rho \in E$ is the barycenter of some $\mu \in E(C(E;\mathbb{C}))$, whose unique extension to a probability measure is supported in $\mathcal{E}(E)$, the extreme boundary of the convex set E.*

For a complete proof of Choquet's theorem for convex sets of continuous linear functionals on general separable Banach spaces, we refer to [13].

It is important to notice at this point that Choquet's theorem may fail to hold if the C^*-algebra $\mathcal{A}$ is not separable. Thus, the separability of the considered C^*-algebra is essential in our study of equilibrium states of infinite quantum systems. By contrast, Corollary 7.338 does have an extension for the non-separable case.

7.5.4 The γ-Regularization

In the study of convex sets and functions, it is very natural to use continuous affine functionals, like straight lines in the plane. For instance, lower semicontinuous convex functions $\mathbb{R} \to \mathbb{R}$ are always given by lower envelopes of continuous affine functionals. This motivates the introduction of the so-called γ-regularization of real-valued functions on metric[16] vector spaces, defined as follows:

Definition 7.340 (γ-Regularization of Real-Valued Functions) Let (V,d) be any metric vector space over $\mathbb{R}$ and $M \subseteq V$ a nonempty convex set. For any real-valued function $f : M \to \mathbb{R}$, its "γ-regularization" $\gamma(f)$ on M, also known as the "lower envelop of f," is the function defined as being the supremum over all affine and continuous minorants $m : V \to \mathbb{R}$ of f, i.e., for all $x \in M$,

[16] Or, more generally, on topological vector spaces.

$$\gamma\,(f)\,(x) \doteq \sup\,\{m(x) : m \in \mathrm{A}(M) \text{ and } m \le f\}\ ,$$

where $\mathrm{A}(M)$ is the space of all affine (i.e., both convex and concave) continuous real-valued functions on M.

Since the γ-regularization $\gamma\,(f)$ of a real-valued function f is a supremum over continuous functions, $\gamma\,(f)$ is lower semicontinuous, thanks to Lemma 7.144. It is additionally convex, being the supremum of convex functions (in this case, even affine). In locally convex spaces, every convex and lower semicontinuous function on a closed convex domain M equals its own γ-regularization on K.

Proposition 7.341 (γ-Regularization of Convex Lower Semicontinuous Functions) *Let (V, d) be any locally convex metric vector space over $\mathbb{R}$ and $M \subseteq V$ a nonempty closed convex set. Take any real-valued function $f : K \to \mathbb{R}$. Then $\gamma\,(f) = f$ iff f is convex and lower semicontinuous.*

Proof As already explained, if $\gamma\,(f) = f$, then f is convex and lower semicontinuous. We prove now the converse assertion:

1. Take a convex and lower semicontinuous function f on K. Since K is, by assumption, convex and closed, by Lemma 7.148 and Exercise 7.329, $\mathrm{epi}(f) \subseteq V \times \mathbb{R}$ (see Definition 7.147) is convex and closed with respect to the Cartesian product metric for $V \times \mathbb{R}$ (see Definition 7.93). Recall that $V \times \mathbb{R}$ is a locally convex metric space with respect to the Cartesian product metric, thanks to Exercise 7.119.
2. Fix $x \in K$. Then, in order to prove that $\gamma\,(f)\,(x) = f\,(x)$, we show that, for any arbitrary parameter $s < f(x)$, there is an affine continuous real-valued function m on K such that $m \le f$ and $m\,(x) > s$. In fact, the one-point set $\{(x, s)\} \subseteq V \times \mathbb{R}$ is trivially compact and convex, and $\mathrm{epi}(f)$ is convex and closed, as explained in Point 1. Therefore, from the Hahn-Banach separation theorem for locally convex spaces (Theorem 7.331 (ii)), there are a continuous linear functional $\Theta : V \times \mathbb{R} \to \mathbb{R}$ and a constant $\zeta \in \mathbb{R}$ such that, for all $(x', s') \in \mathrm{epi}(f)$,

$$\Theta(x, s) > \zeta \ge \Theta(x', s')\ . \tag{7.9}$$

   Any continuous linear functional like Θ is of the form

$$\Theta(x', s') = \theta(x') - cs'\ , \qquad (x', s') \in V \times \mathbb{R}\ ,$$

   where $c \in \mathbb{R}$ is some constant and $\theta : V \to \mathbb{R}$ is a linear continuous functional. By applying (7.11) to $x' = x$, note that $c \neq 0$, while $c \ge 0$ because

$$\sup\{s' \in \mathbb{R}\ :\ (x', s') \in \mathrm{epi}(f)\} = \infty\ , \qquad x' \in K\ .$$

3. If one takes $s' = f(x)$ and $x' = x$, then we conclude that

$$s < m(x) \doteq c^{-1}\theta(x) - c^{-1}\zeta \leq f(x) .$$

Additionally, by taking $s' = f(x')$, $x' \in M$, we arrive at

$$m\left(x'\right) \doteq c^{-1}\theta(x') - c^{-1}\zeta \leq f(x'), \qquad x' \in K .$$

□

This last proposition is a standard result. It implies that $\gamma(f)$ is the largest lower semicontinuous and convex minorant of f:

Corollary 7.342 (Largest Lower Semicontinuous Convex Minorants) *Let (V, d) be any locally convex metric vector space over $\mathbb{R}$ and $M \subseteq V$ a nonempty closed convex set. Take any real-valued function $f : K \to \mathbb{R}$. Then its γ-regularization $\gamma(f)$ is its largest lower semicontinuous convex minorant on K.*

Proof For any lower semicontinuous convex real-valued function $g : K \to \mathbb{R}$ satisfying $g \leq f$, we infer from Definition 7.340 and Proposition 7.341 that

$$g(x) = \gamma(g)(x) \doteq \sup\{m(x) : m \in \mathrm{A}(K) \text{ and } \theta \leq g \leq f\} \leq \gamma(f)(x)$$

for any $x \in K$. □

Proposition 7.341 has another interesting consequence: An extension of the Bauer maximum principle which, in the case of convex functions, refers to the following lemma:

Lemma 7.343 (Bauer Maximum Principle) *Let (V, d) be any locally convex metric vector space over $\mathbb{R}$ and $K \subseteq V$ a nonempty compact convex set. An upper semicontinuous convex real-valued function $f : K \to \mathbb{R}$ attains its maximum at an extreme point of K, i.e.,*

$$\sup f(K) = \max f(\mathcal{E}(K)) .$$

Here, $\mathcal{E}(K)$ is the (nonempty) set of extreme points of K (see Proposition 7.334).

Proof

1. Let

$$\Omega \doteq \{x \in K : f(v) = \sup f(K)\} .$$

Since $f : K \to \mathbb{R}$ is upper semicontinuous and convex, Ω is a nonempty closed (and thus compact) subset of K, thanks to Proposition 7.172. Moreover, by convexity of f, for all $x, x' \in K$ and $\alpha \in (0, 1)$ such that

$$\alpha x + (1 - \alpha)x' \in \Omega \ ,$$

one has that $x, x' \in \Omega$. In other words, Ω is an extreme subset of K. See Definition 7.333. Let $\mathcal{CES}(\Omega, K)$ be the family of all nonempty closed subsets of Ω which are also extreme subsets (or faces) of K. We consider the partial order relation in $\mathcal{CES}(\Omega, K)$ given by the reverse inclusion, that is, $M \leq M'$ if $M' \subseteq M$. By using Zorn's lemma, like in the proof of Proposition 7.334 with $\mathcal{CES}(K)$, the partially ordered set $\mathcal{CES}(\Omega, K)$ has at least one maximal element $\Omega_\infty \in \mathcal{CES}(\Omega, K)$, that is, for any $\Omega' \in \mathcal{CES}(\Omega, K)$, one has $\Omega' \subseteq \Omega_\infty$ only if $\Omega' = \Omega_\infty$.

2. We now show that such a maximal element must be of the form $\Omega_\infty = \{x_\infty\}$ for some $x_\infty \in K$. In particular, $x_\infty \in \mathcal{E}(K) \cap \Omega$. This is done exactly like in the proof of Proposition 7.334: Assume by contradiction that $\Omega_\infty = \{x_\infty, x'_\infty\}$ with $x_\infty \neq x'_\infty$. The Hahn-Banach separation theorem (Theorem 7.331 (ii)) implies the existence of a continuous linear functional $\theta : V \to \mathbb{R}$ such that $\theta\left(x'_\infty\right) > \theta\,(x_\infty)$. Now, define the closed set

$$\Omega' \doteq \{x \in \Omega : \theta\,(x) = \sup \theta\,(\Omega)\} \ .$$

Observe that Ω' is an extreme subset of Ω and therefore an extreme subset of K. See Definition 7.333. Since $x'_\infty \notin \Omega'$, it follows that $\Omega' \subsetneq \Omega_\infty$, which is absurd because $\Omega' \in \mathcal{CES}(\Omega, K)$ and Ω_∞ is a maximal element of $\mathcal{CES}(\Omega, K)$.

□

By combining Proposition 7.341 with Lemma 7.343, one can extend the Bauer maximum principle to study convex functions that are not necessarily upper semicontinuous:

Lemma 7.344 (Extension of the Bauer Maximum Principle) *Let (V, d) be any locally convex metric vector space over $\mathbb{R}$ and $K \subseteq V$ a nonempty compact convex set. Let $f_\pm : K \to \mathbb{R}$ be two convex real-valued functions such that f_- and f_+ are, respectively, lower and upper semicontinuous. Then, it suffices to take the supremum of the sum $f \doteq f_- + f_+$ in the (nonempty) set $\mathcal{E}(K)$ of extreme points of K, i.e.,*

$$\sup\, f\,(K) = \sup f\,(\mathcal{E}(K)) \ .$$

Proof We first use Proposition 7.341 in order to write $f_- = \gamma\,(f_-)$ as a supremum over affine and continuous functions. Then we commute this supremum with the one over K and apply Lemma 7.343 to obtain that

$$\sup f\,(K) = \sup \{\sup\, [m + f_+]\,(\mathcal{E}(K)) : m \in \mathrm{A}(K) \text{ and } m \leq f_-\} \ .$$

The lemma then follows by commuting again both suprema and by using $f_- = \gamma\,(f_-)$, thanks again to Proposition 7.341. □

Observe, however, that under the conditions of the lemma above, the supremum of $f = f_- + f_+$ is generally not attained on $\mathcal{E}(K)$, in contrast with the usual Bauer maximum principle (Lemma 7.343). Note that this extension has been recently observed, in 2012 [25].

We now conclude this section by an important, relatively recent, result which is pivotal for the study of equilibrium states of mean-field models done in Chap. 6.

Theorem 7.345 (Minimization of Real-Valued Functions) *Let (V, d) be any locally convex metric vector space over $\mathbb{R}$ and $K \subseteq V$ a nonempty compact convex set. Let $f : K \to [\mathrm{k}, \infty)$ be any real-valued functional bounded from below by some $\mathrm{k} \in \mathbb{R}$. Then, the following assertions hold true:*

(i)

$$\inf f(K) = \inf \gamma(f)(K) \ .$$

(ii) The set M of minimizers of $\gamma(f)$ over K equals the closed convex hull of the closed set

$$\Omega(f, K) \doteq \left\{ x \in K : \exists \{x_i\}_{i \in I} \subset K \text{ with } x_i \to x \text{ and } \lim_I f(x_i) = \inf f(K) \right\}$$

of generalized minimizers of f over K, i.e.,

$$M = \overline{\mathrm{co}}\,(\Omega(f, K)) \ .$$

Proof

1. By Definition 7.340, $\gamma(f) \leq f$ on K and thus

$$\inf \gamma(f)(K) \leq \inf f(K) \ .$$

The converse inequality is derived by restricting the supremum in Definition 7.340 to constant functions $m : K \to \mathbb{R}$ with $\mathrm{k} \leq m \leq f$.

2. Observe from Propositions 7.172 and 7.341 that the variational problem $\inf \gamma(f)(K)$ has minimizers and the set $M \doteq \Omega(\gamma(f), K)$ of all minimizers of $\gamma(f)$ is convex and compact. For any $y \in \Omega(f, K)$, by definition, there is a net $\{x_i\}_{i \in I} \subseteq K$ of approximating minimizers of f on K converging to y. In particular, since the function $\gamma(f)$ is lower semicontinuous and $\gamma(f) \leq f$ on K, we have that

$$\gamma(f)(y) \leq \liminf_I \gamma(f)(x_i) \leq \lim_I f(x_i) = \inf f(K) = \inf \gamma(f)(K) \ ,$$

i.e., $y \in M$. Since M is convex and compact, we obtain that

$$M \supseteq \overline{\mathrm{co}}\,(\Omega(f, K)) \ .$$

3. The proof of the converse inclusion is much more involved and we omit it here. It uses, in particular, the Urysohn lemma and a crucial property concerning the γ-regularization of real-valued functions in relation with the concept of barycenters [20, Corollary I.3.6.]. For more details, we recommend either [1, Theorem 10.37] or [25, Theorem 1.4].

□

This general fact related to the minimization of non-convex and non-lower semi-continuous real-valued functions on compact convex sets has only been observed[17] in 2012 [25]. Notice that, previously, related results were obtained in [98], but only for $V = \mathbb{R}^n$, $n \in \mathbb{R}$. It is an essential argument in the proof of Theorem 6.25 and, therefore, for our current study of equilibrium states of mean-field models.

By using the theory of compact convex subsets of locally convex metric vector spaces V (e.g., Theorem 7.339), Theorem 7.345 yields a very useful characterization of the set $\Omega(f, K)$ of all generalized minimizers of f over K. In fact, recall that any compact convex subset K is the closure of the convex hull of the (nonempty) set $\mathcal{E}(K)$ of its extreme points, i.e., of the points which cannot be expressed as (nontrivial) convex combinations of other elements in K. This is the Krein-Milman theorem (Proposition 7.334). Moreover, among all subsets $Z \subseteq K$ generating the convex set K, $\mathcal{E}(K)$ is—in a sense—the smallest one. This is the Milman theorem. See, e.g., [18, Theorem 3.25]. It follows from Theorem 7.345 together with [18, Theorems 3.4 (b), 3.23, 3.25] that extreme points of the compact convex set M of minimizers of $\gamma(f)$ over K are generalized minimizers of f. This refers to the following result:

Theorem 7.346 (Minimization of Real-Valued Functions—II) *Let* (V, d) *be any locally convex metric vector space over* $\mathbb{R}$ *and* $K \subseteq V$ *a nonempty compact convex set. Let* $f : K \to [\mathrm{k}, \infty)$ *be any real-valued functional bounded from below by some* $\mathrm{k} \in \mathbb{R}$. *Then extreme points of the compact convex set* $M \doteq \Omega(\gamma(f), K)$ *belong to the set of generalized minimizers of* f, *i.e.,* $\mathcal{E}(M) \subseteq \Omega(f, K)$.

This last result makes possible a *full characterization* of the closure of the set $\Omega(f, K)$ in the following sense: Since M is compact and convex, we can study the minimization problem $\inf f(K_M)$ for any closed (and hence compact) convex subset $K_M \subseteq M$. Applying Theorem 7.345, we get

$$\inf f(K_M) = \inf \gamma\left(f|_{K_M}\right)(K_M) \ . \tag{7.10}$$

If

$$\inf f(K_M) = \inf f(K)$$

[17] Assertion (i) is, however, trivial.

then, by Theorem 7.346,

$$\mathcal{E}\left(M_{K_M}\right) \subseteq \Omega\left(f|_{K_M}, K_M\right) \subseteq \Omega\left(f, K\right) \ ,$$

where M_{K_M} is the compact convex set of minimizers of $\gamma\left(f|_{K_M}\right)$ over $K_M \subseteq M$. In general, $\mathcal{E}\left(M_{K_M}\right) \backslash \mathcal{E}\left(M\right) \neq \emptyset$ because M_{K_M} is not necessarily a face of M. Thus, we discover in this manner new points of $\Omega\left(f, K\right)$ not contained in $\mathcal{E}\left(M\right)$. Choosing a sufficiently large family $\{K_M\}$ of closed convex subsets of M, we can exhaust the set $\Omega\left(f, K\right)$ through the union $\cup\{\mathcal{E}\left(M_{K_M}\right)\}$. Note that this construction can be performed in an inductive way: For each set M_{K_M} of minimizers, consider further closed convex subsets $K_M' \subset M_{K_M}$. The art consists in choosing the family $\{K_M\}$ appropriately, i.e., it should be as small as possible, and the extreme points of M_{K_M} should possess some reasonable characterization. Of course, the latter heavily depends on the function f and on particular properties of the compact convex set K (e.g., density of $\mathcal{E}(K)$, metrizability, etc.). This general strategy is exactly what is done in [1] to prove Theorem 6.36.

Since the γ-regularization involves affine functions, we conclude this section by making explicit the affine functions over states on unital C^*-algebras $\mathcal{A}$, as is used in Chap. 6. For some fixed dense sequence $\mathcal{S} = (x_n)_{n\in\mathbb{R}}$ in $\mathcal{A}$, $d \equiv d_{\mathcal{S}}$ denotes the metric in (the topological dual space) $\mathcal{A}^{\mathrm{td}}$, which is introduced in Definition 4.80. See also Exercise 4.82. Observe additionally that $(\mathcal{A}^{\mathrm{td}}, d)$ is a locally convex metric vector space. See Exercise 7.123.

Proposition 7.347 (Affine Weak*-Continuous Real-Valued Functions over States) *Let $\mathcal{A}$ be any separable unital C^*-algebra and $(\mathcal{A}^{\mathrm{td}}, d)$ be the locally convex metric vector space of Definition 4.80 for some dense sequence $\mathcal{S}$ in $\mathcal{A}$. See Exercises 4.82 and 7.123. For any convex weak*-compact subset $E \subseteq E\left(\mathcal{A}\right) \subseteq \mathrm{Re}\{\mathcal{A}^{\mathrm{td}}\}$,*

$$\mathrm{A}(E) = \{\rho \mapsto \rho\left(A\right) \in \mathbb{R} : A \in \mathrm{Re}\{\mathcal{A}\}\} \ ,$$

where $\mathrm{A}\left(E\right)$ is the space of all affine (i.e., both convex and concave) weak*-continuous real-valued functions on E.

Proof Clearly,

$$\{\rho \mapsto \rho\left(A\right) \in \mathbb{R} : A \in \mathrm{Re}\{\mathcal{A}\}\} \subseteq \mathrm{A}(E) \ .$$

Conversely, fix $f \in \mathrm{A}\left(E\right)$. Since E is a convex weak *-compact subset of $E\left(\mathcal{A}\right)$, we deduce from [60, Corollary 6.3] the existence of an increasing sequence $\{f_n\}_{n\in\mathbb{N}}$ of affine continuous real-valued functions on $(\mathrm{Re}\{\mathcal{A}^{\mathrm{td}}\}, d)$ that uniformly converges to f, as $n \to \infty$. Meanwhile, observe from Proposition 7.126 that any affine weak*-continuous real-valued functions g on $(\mathrm{Re}\{\mathcal{A}^{\mathrm{td}}\}, d)$ is of the form

$$g\left(\sigma\right) = \sigma\left(A\right) + g\left(0\right) \ , \qquad \sigma \in \mathrm{Re}\{\mathcal{A}^{\mathrm{td}}\} \ ,$$

for some self-adjoint element $A \in \mathrm{Re}\{\mathcal{A}\} \cap \mathrm{span}\mathcal{S}$, because the continuous real-valued function $g - g(0)$ is linear. We thus deduce the existence of a sequence $\{A_n\}_{n\in\mathcal{N}} \subseteq \mathrm{Re}\{\mathcal{A}\} \cap \mathrm{span}\mathcal{S}$ such that

$$f_n(\sigma) = \sigma(A_n) + f_n(0) \ , \qquad \sigma \in \mathrm{Re}\{\mathcal{A}^{\mathrm{td}}\} \ .$$

Since $\rho(\mathfrak{1}) = 1$ for $\rho \in E$ (Definition 4.62), by Proposition 4.78 (ii), the uniform convergence of $\{f_n\}_{n\in\mathbb{N}}$ to f on E yields that $\{A_n + f_n(0)\mathfrak{1}\}_{n\in\mathbb{N}} \subseteq \mathrm{Re}\{\mathcal{A}\}$ is a Cauchy sequence, which thus converges to some $A \in \mathrm{Re}\{\mathcal{A}\}$, as $n \to \infty$. It follows that $f(\rho) = \rho(A)$ for any state $\rho \in E$, where $A \in \mathrm{Re}\{\mathcal{A}\}$. □

7.5.5 The Legendre-Fenchel Transform

In this subsection, we define and prove some important, albeit simple, properties of the Legendre-Fenchel transform. In view of the applications in Chap. 5, we will indeed consider the infinite dimensional but not in full generally. For simplicity and clearness of the exposition, we only consider Legendre-Fenchel transforms of convex lower semicontinuous functions on some weak*-compact convex set of continuous linear functionals on some separable[18] normed space. This case is general enough for the applications to quantum statistical mechanics presented in this book. For a more thorough exposition on Legendre-Fenchel transforms, we recommend [20]. Throughout the present subsection, X thus denotes an arbitrary *separable* real normed space. For some fixed dense sequence $\mathcal{S} = (x_n)_{n\in\mathbb{R}}$ in X, $d \equiv d_{\mathcal{S}}$ denotes the metric in (the topological dual space) X^{td}, which is introduced in Definition 4.80. See also Exercise 4.82. Observe additionally that (X^{td}, d) is a locally convex metric vector space. See Exercise 7.123.

Definition 7.348 (Legendre-Fenchel Transform) Let K be any nonempty norm-bounded[19] (weak*)-compact set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous function. Then, the "Legendre-Fenchel" transform of f is the function $f^* : X \to \mathbb{R}$ defined by

$$f^*(x) \doteq \sup_{\varphi\in K}\{\varphi(x) - f(\varphi)\} \ , \qquad x \in X \ .$$

Notice that, for any fixed $x \in X$, the mapping $\varphi \mapsto \varphi(x) - f(\varphi)$ from K to $\mathbb{R}$ is upper semicontinuous: In fact, as K is norm-bounded, one can easily check that, for any fixed $x \in X$, the mapping $\varphi \mapsto \varphi(x)$ is continuous with respect

[18] Recall that considering separable normed spaces has the technical advantage that weak*-compact sets of continuous linear functionals on this spaces are metrizable, as explained in Sect. 4.5.1 for sets of states on unital C^*-algebras.

[19] That is, $\sup_{\varphi\in K} \|\varphi\|_{\mathrm{op}} < \infty$.

to the metric $d \equiv d_{\mathcal{S}}$ in K. Thus, by Proposition 7.146, $\varphi \mapsto \varphi(x) - f(\varphi)$ is upper semicontinuous on K, as f is lower semicontinuous. With this remark, by Proposition 7.172, for any fixed $x \in X$, the mapping $\varphi \mapsto \varphi(x) - f(\varphi)$ takes a maximum in K. In particular, the set of real numbers

$$\{\varphi(x) - f(\varphi) : \varphi \in K\} \subseteq \mathbb{R}$$

is bounded from above and, thus, the Legendre-Fenchel transform f^* of f is well-defined on the whole space X.

The following inequality is important in the theory of Legendre-Fenchel transforms, albeit elementary:

Lemma 7.349 (Young's Inequality) *Let K be any nonempty compact set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous function. Then, for all $\varphi \in K$ and $x \in X$,*

$$f(\varphi) + f^*(x) \geq \varphi(x) .$$

Proof The lemma directly follows from the definition of the Legendre-Fenchel transform f^*. □

As a simple consequence of its definition, the Legendre-Fenchel transform f^* is always a convex function on X. Additionally, in the special case of our interest, it is also continuous:

Proposition 7.350 *Let K be any nonempty norm-bounded compact set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous function. Then the Legendre-Fenchel transform $f^* : X \to \mathbb{R}$ is convex and Lipschitz continuous, i.e., for some constant $L \in \mathbb{R}_0^+$, one has that*

$$|f^*(x) - f^*(x')| \leq L \left\| x - x' \right\| , \qquad x, x' \in X .$$

Proof Take any $x, x' \in X$. Then, for all $\alpha \in [0, 1]$,

$$\begin{aligned}
&\sup_{\varphi \in K} \{\varphi(\alpha x + (1-\alpha)x') - f(\varphi)\} \\
&= \sup_{\varphi \in K} \{\alpha\varphi(x) - \alpha f(\varphi) + (1-\alpha)\varphi(x') - f(\varphi)(1-\alpha)\} \\
&\leq \alpha \sup_{\varphi \in K} \{\varphi(x) - f(\varphi)\} + (1-\alpha) \sup_{\varphi \in K} \{\varphi(x') - f(\varphi)\} .
\end{aligned}$$

Thus, for any $x, x' \in X$ and all $\alpha \in [0, 1]$,

$$f^*(\alpha x + (1-\alpha)x') \leq \alpha f^*(x) + (1-\alpha) f^*(x') ,$$

that is, $f^* : X \to \mathbb{R}$ is convex. As K is norm-bounded, there is $L \in \mathbb{R}_0^+$ such that $\sup_{\varphi \in K} \|\varphi\|_{\mathrm{op}} \leq L$. Then, clearly, for all $x, x' \in X$,

$$|(\varphi(x) - f(\varphi)) - (\varphi(x') - f(\varphi))| = |\varphi(x - x')| \leq L \left\|x - x'\right\|$$

and, since the mapping $\varphi \mapsto \varphi(x) - f(\varphi)$ takes a maximum at some $\varphi_x \in K$, we arrive at

$$|f^*(x) - f^*(x')| \leq (\varphi_x(x) - f(\varphi_x)) - (\varphi_x(x') - f(\varphi_x)) \leq L \left\|x - x'\right\|$$

for all $x, x' \in X$. □

For every $\varphi \in K$, define the affine mapping $h_\varphi : X \to \mathbb{R}$ by

$$h_\varphi(x) \doteq \varphi(x) - f(\varphi) , \qquad x \in X .$$

With this definition, $f^* = \sup_{\varphi \in K} h_\varphi$. If the topology of X^{td} is chosen such that h_φ is continuous for every $\varphi \in K$ (for instance, this is so for the (global) weak* topology of X^{td}), then f^* is the supremum of a family of continuous mapping. Hence, by an adequate version of Lemma 7.144, the Legendre-Fenchel transform f^* is lower semicontinuous. In the usual setting for the theory of Legendre-Fenchel transforms, one rather finds this situation, which is more general than the one that we present here with compact sets K on (X^{td}, d). Nevertheless, note that the pressure function of Definition 5.28 is a Legendre-Fenchel transform (the one of the entropy density functional times the constant β^{-1}) in the more restrictive sense of Definition 7.348. In fact, the pressure function is such a transform of a function defined on a set of states, which are, by definition, norm-one linear functionals. Note at this point that, because of this fact, the Lipschitz constant (L in the last proposition) for the pressure function is 1. This is nothing else than a version of Bogoliubov's inequality (Corollary 3.25) for infinite systems. See Proposition 5.30 (i).

We now derive a kind of "inversion formula" for the above-defined Legendre-Fenchel. In fact, as is well-known and demonstrated below in our special case, such a formula for convex functions is provided by the "double Legendre-Fenchel" transform, which is defined as follows:

Definition 7.351 (Double Legendre-Fenchel Transform) Let K be any nonempty norm-bounded compact set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous function. Then, the "double Legendre-Fenchel transform" of f is the function $f^{**} : \mathrm{dom}(f^{**}) \to \mathbb{R}$ defined by

$$f^{**}(\varphi) \doteq \sup_{x \in X} \{\varphi(x) - f^*(x)\} , \qquad \varphi \in \mathrm{dom}(f^{**}) \subseteq X^{\mathrm{td}} ,$$

where

$$\mathrm{dom}(f^{**}) \doteq \left\{ \varphi \in X^{\mathrm{td}} \ : \ \sup_{x \in X} \{\varphi(x) - f^*(x)\} < \infty \right\} .$$

Young's inequality (Lemma 7.349) already yields important properties of the double Legendre-Fenchel transform:

Corollary 7.352 *Let K be any nonempty norm-bounded compact set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous function. Then $K \subseteq \mathrm{dom}(f^{**})$ and $f^{**}(\varphi) \leq f(\varphi)$ for every $\varphi \in K$.*

Proof By Young's inequality, for all $\varphi \in K$ and all $x \in X$, one has that

$$f(\varphi) \geq \varphi(x) - f^*(x) .$$

Thus, by taking the supremum with respect to $x \in X$, the corollary follows. □

In the next theorem, we show that if the set K is convex and $f : K \to \mathbb{R}$ is a convex function, then the double Legendre-Fenchel transform f^{**} is nothing else than the original function f:

Theorem 7.353 (Inversion Formula for the Legendre-Fenchel Transform) *Let K be any nonempty norm-bounded, compact, and* convex *set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous* convex *function. Then $\mathrm{dom}(f^{**}) = K$ and $f^{**}(\varphi) = f(\varphi)$ for all $\varphi \in K$.*

Proof

1. As $K \subseteq X^{\mathrm{td}}$ is convex and closed (being compact) while f is convex and lower semicontinuous, by Lemma 7.148 and Exercise 7.329, $\mathrm{epi}(f) \subseteq X^{\mathrm{td}} \times \mathbb{R}$ (see Definition 7.147) is convex and closed with respect to the Cartesian product metric for $X^{\mathrm{td}} \times \mathbb{R}$ (see Definition 7.93). Recall that $X^{\mathrm{td}} \times \mathbb{R}$ is a locally convex metric space with respect to the Cartesian product metric, by Exercises 7.119 and 7.123. Thus, from the Hahn-Banach separation theorem for locally convex spaces (Theorem 7.331 (ii)), for all $\varphi \in K$ and all $s \in \mathbb{R}$ with $s < f(\varphi)$, there are a continuous linear functional $\Theta : X^{\mathrm{td}} \times \mathbb{R} \to \mathbb{R}$ and a constant $\gamma \in \mathbb{R}$, such that, for all $(\varphi', s') \in \mathrm{epi}(f)$, one has

$$\Theta(\varphi, s) > \gamma \geq \Theta(\varphi', s') . \tag{7.11}$$

Note here that the one-point set $\{(\varphi, s)\} \subseteq X^{\mathrm{td}} \times \mathbb{R}$ is trivially compact and convex. Remark additionally that any continuous linear functional $\Theta : X^{\mathrm{td}} \times \mathbb{R} \to \mathbb{R}$ is of the form

$$\Theta(\varphi', s') = \theta(\varphi') - cs' , \qquad (\varphi', s') \in X^{\mathrm{td}} \times \mathbb{R} ,$$

where $c \in \mathbb{R}$ is some constant and $\theta : X^{\mathrm{td}} \to \mathbb{R}$ is a linear continuous functional. Since

$$\sup\{s' \in \mathbb{R} \ : \ (\varphi', s') \in \mathrm{epi}(f)\} = \infty \ , \qquad \varphi' \in K \ ,$$

one infers from (7.11) that $c \geq 0$, but also that $c \neq 0$, because, otherwise, one would arrive at the contradiction $\theta(\varphi) > \theta(\varphi)$. Hence, $c > 0$, and we have

$$c^{-1}\theta(\varphi) - s > c^{-1}\gamma \geq c^{-1}\theta(\varphi') - s' \ , \qquad (\varphi', s') \in \mathrm{epi}(f) \ .$$

Now, by Proposition 7.126, there is some $x \in X$ such that

$$c^{-1}\theta(\varphi') = \varphi'(x) \ , \qquad \varphi' \in X^{\mathrm{td}} \ .$$

In particular, by choosing $s' = f(\varphi')$,

$$\varphi(x) - s > c^{-1}\gamma \geq \varphi'(x) - f(\varphi') \ , \qquad \varphi' \in K \ .$$

Taking the supremum with respect to $\varphi' \in K$, we then conclude that

$$s < \varphi(x) - f^*(x) \ .$$

Hence, $f^{**}(\varphi) > s$ for all $\varphi \in K$ and all $s < f(\varphi)$. In particular, for any $\varepsilon > 0$, $f^{**}(\varphi) > f(\varphi) - \varepsilon$, i.e., $f^{**}(\varphi) \geq f(\varphi)$. By Corollary 7.352, $K \subseteq \mathrm{dom}(f^{**})$ and $f^{**}(\varphi) \leq f(\varphi)$ for all $\varphi \in K$. From these last two observations, we then arrive at $f^{**}(\varphi) = f(\varphi)$ for all $\varphi \in K$.

2. Take now some $\varphi \notin K$ and some constant $s \in \mathbb{R}$. Then, similar to the first part of the proof, again by the Hahn-Banach separation theorem for locally convex spaces (Theorem 7.331 (ii)) and Proposition 7.126, there are $x \in X$, $c \geq 0$, and $\gamma \in \mathbb{R}$ such that

$$\varphi(x) - cs > \gamma \geq \varphi'(x) - cs' \ , \qquad (\varphi', s') \in \mathrm{epi}(f) \ .$$

Assume that $c > 0$. Then,

$$\varphi(\bar{x}) - s > c^{-1}\gamma \geq \varphi'(\bar{x}) - f(\varphi') \ , \qquad \varphi' \in K \ ,$$

where $\bar{x} \doteq c^{-1}x$. Thus, by taking the supremum with respect to $\varphi' \in K$,

$$s < \varphi(\bar{x}) - f^*(\bar{x}) \ .$$

In particular, one has that, for all $s \in \mathbb{R}$,

$$\sup\{\varphi(x') - f^*(x') \ : \ x' \in X\} > s \ ,$$

that is, $\varphi \notin \mathrm{dom}(f^{**})$. Finally, suppose that $c = 0$. Then,

$$\varphi(x) > \gamma \geq \varphi'(x) , \qquad \varphi' \in K ,$$

that is, for some $\varepsilon > 0$,

$$(\varphi' - \varphi)(x) + \varepsilon \leq 0 , \qquad \varphi' \in K .$$

Observe from Young's inequality (Lemma 7.349) that, for any fixed $x' \in X$,

$$\varphi'(x') - f^*(x') - f(\varphi') \leq 0, \qquad \varphi' \in K .$$

Thus, for all $\alpha > 0$, $x' \in X$, and $\varphi' \in K$,

$$\varphi'(x') - f^*(x') - f(\varphi') + \alpha(\varphi' - \varphi)(x) + \alpha\varepsilon \leq 0 ,$$

which is equivalent to

$$\varphi'(x' + \alpha x) - f(\varphi') + \alpha\varepsilon \leq f^*(x') + \varphi(\alpha x) , \qquad \alpha > 0,\ x' \in X,\ \varphi' \in K .$$

Taking the supremum with respect to $\varphi' \in K$, for all $\alpha > 0$ and $x' \in X$, one arrives at

$$\varphi(x') - f^*(x') + \alpha\varepsilon \leq \varphi(x' + \alpha x) - f^*(x' + \alpha x) .$$

From the last inequality, we now deduce that, for every $s > 0$,

$$\sup\{\varphi(x') - f^*(x') \ : \ x' \in X\} \geq s .$$

Thus, $\varphi \notin \mathrm{dom}(f^{**})$.

□

Corollary 7.354 *Let K be any nonempty norm-bounded, compact, and convex set in the metric space (X^{td}, d). For any lower semicontinuous convex function $f : K \to \mathbb{R}$, $f^{**} = \gamma(f) = f$.*

Proof Combine Theorem 7.353 with Proposition 7.341. □

By Corollaries 7.342 and 7.352, note that $f^{**} \leq \gamma(f) \leq f$ for any lower semicontinuous function f. In fact, in general, the equality $f^{**} = \gamma(f)$ holds true. In other words, the γ-regularization of a function equals its double Legendre-Fenchel transform. See, for instance, [99, Paragraph 51.3].

7.5.6 Subdifferentials and Subgradients

In this brief subsection, we use Theorem 7.353 to show that the minimizers of any lower semicontinuous convex function $f : K \to \mathbb{R}$ on a norm-bounded, compact, and convex set in the metric space (X^{td}, d) are nothing else than the tangent functionals of the corresponding Legendre-Fenchel transform at $0 \in X$. Recall that this is a pivotal argument for the proof of existence of phase transitions in Chap. 5. The following notion is natural in this context:

Definition 7.355 (Subdifferentials and Gradients) Let X be a real normed space. Let M be any nonempty *convex* set in the vector space X^{td} and $f : M \to \mathbb{R}$ any *convex* function. For all $\varphi \in M$, we define the "subdifferential of f at $\varphi \in M$" by

$$\partial f(\varphi) \doteq \{x \in X \;:\; f(\varphi') - f(\varphi) \geq (\varphi' - \varphi)(x) \text{ for all } \varphi' \in M\} \subseteq X .$$

Equivalently,

$$\partial f(\varphi) = \left\{ x \in X \;:\; f(\varphi) - \varphi(x) = \min_{\varphi' \in K} \{f(\varphi') - \varphi'(x)\} \right\} \subseteq X .$$

The elements of $\partial f(\varphi)$ are called the "subgradients" of f at $\varphi \in M$. In a similar way, for an arbitrary convex function $f : M \to \mathbb{R}$ on some nonempty convex subset $M \subseteq X$ of X, at any $x \in M$, we define the subdifferential

$$\partial f(x) = \left\{ \varphi \in X^{\mathrm{td}} \;:\; f(x) - \varphi(x) = \min_{x' \in X} \{f(x') - \varphi(x')\} \right\} \subseteq X^{\mathrm{td}} .$$

Again, the elements of $\partial f(x)$ are called the "subgradients" of f at $x \in M$.

Observe that subgradients are nothing else than tangent functionals in the sense of Definition 3.18. One important characterization of the property of being a subgradient is related to Young's inequality (Lemma 7.349), as stated in the following lemma:

Lemma 7.356 *Let X be a separable real normed space. Let K be any nonempty norm-bounded, compact, and* convex *set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous* convex *function. For all $\varphi \in K$, $\partial f(\varphi)$ is a convex subset of X and $x \in \partial f(\varphi)$ iff*

$$f(\varphi) + f^*(x) = \varphi(x) ,$$

i.e., Young's inequality is satisfied with equality.

Proof Exercise. □

Now, by combining the last lemma with Theorem 7.353, i.e., the equality $f^{**} = f$, we arrive at the following important result:

Proposition 7.357 *Let X be a separable real normed space. Let K be any nonempty norm-bounded, compact, and* convex *set in the metric space (X^{td}, d) and $f : K \to \mathbb{R}$ any lower semicontinuous* convex *function. Then, for all $x \in X$ and $\varphi \in K$, $x \in \partial f(\varphi)$ iff $\varphi \in \partial f^*(x)$. In particular, for every $x \in X$, $\varphi \in K$ minimizes the functional $\varphi' \mapsto f(\varphi') - \varphi'(x)$ on K iff $\varphi \in \partial f^*(x)$.*

Proof Exercise. □

Notice that, by the assumptions of the proposition, for every $x \in X$, the functional $\varphi' \mapsto f(\varphi') - \varphi'(x)$ on K is lower semicontinuous. In particular, it has minimizers, and, thus, the subdifferential $\partial f^*(x)$ is nonempty for all $x \in X$.

7.5.7 Exposed Points of Convex Sets and the Bishop-Phelps Theorem

In this subsection, we present the Bishop-Phelps theorem, which is one of the main arguments for the proof of existence of phase transitions in Chap. 5. See Sect. 5.5. As already mentioned, Israel [70] realized that this theorem yields an elegant proof for the existence of phase transitions. In contrast to Israel's exposition, here we stick as much as possible to the original presentation of the Bishop-Phelps theorem. In fact, originally, this theorem is related to a question that is very different from the one regarding the existence of phase transitions:

The original question that motivated the theorem refers to the existence of norm-one vectors in a normed space that yield the (operator) norm of a given continuous linear functional on the normed space. In fact, observe that, for any *reflexive*[20] Banach space X and every continuous linear functional $\varphi \in X^{\text{td}}$, there is $x \in X$ with $\|x\| = 1$, such that $\|\varphi\|_{\text{op}} = \varphi(x)$: For all $\varphi \in X^{\text{td}}$, there is a net $(x_i)_{i\in I}$ of norm-one vector in X, such that

$$\lim_{i\in I} \varphi(x_i) = \|\varphi\|_{\text{op}} \ .$$

By Banach-Alaoglu's theorem [18, 3.15 Theorem], there is a subnet $(x_j)_{i\in J}$ in $X \subseteq (X^{\text{td}})^{\text{td}}$, which converges in $(X^{\text{td}})^{\text{td}}$, in the weak* sense, to some $\varphi_\infty \in (X^{\text{td}})^{\text{td}}$ with $\|\varphi_\infty\|_{\text{op}} \leq 1$. By reflexivity of X, there is some $x_\infty \in X$ with $\|x_\infty\| \leq 1$, such that $\|\varphi\|_{\text{op}} = \varphi(x_\infty)$. Clearly, by definition of the operator norm $\|\varphi\|_{\text{op}}$, it is not possible that $\|x_\infty\| < 1$, i.e., $\|x_\infty\| = 1$.

With this remark, the question of whether this property is true, or not, in normed spaces that are not reflexive (Banach) space naturally arises. The (original) Bishop-Phelps theorem directly yields that, in any (i.e., not necessarily reflexive) Banach space X, the property is true for continuous linear functionals in a dense subset of

[20] A Banach space X is "reflexive" if its double (topological) dual $(X^{\text{td}})^{\text{td}}$ is exactly X, recalling that any normed space X is naturally identified with a subspace of the bidual space $(X^{\text{td}})^{\text{td}}$.

X^{td}. See Corollary 7.362 below. Thus, the Bishop-Phelps theorem originally refers to a problem in pure mathematics.

We decided to spend some lines to explain how the original form of the theorem is related to a fundamental problem of statistical physics, in order to drive the attention of people from both (mathematical) physics and pure mathematics, in particular those from the domain of functional and convex analysis, to this kind of interesting and mutually enriching connection between different fields of mathematical sciences. At the end of the subsection, we then derive the precise version of the theorem that was used by Israel, as a corollary (Corollary 7.365) of the original one. For completeness, we also provide a direct proof of the corollary, stressing that we do not prove here the original Bishop-Phelps theorem. For more details on this theorem, along with a complete proof, we recommend [13].

In order to state the Bishop-Phelps theorem, we need the following notion:

Definition 7.358 (Exposed Points and Support Functionals) Let V be any real vector space and $M \subseteq V$ any nonempty convex subset of V. We say that $v \in M$ is a "exposed point" of M whenever there exists some linear functional $\varphi : V \to \mathbb{R}$ such that $\varphi(v) > \varphi(v')$ for all $v' \in M \backslash \{v\}$. In this case, we say that φ is a "support functional" for M, associated with v.

Observe that exposed points of a convex set are nothing else than special cases of extreme points:

Exercise 7.359 Let V be any real vector space and $M \subseteq V$ any nonempty convex subset of V. Prove that every exposed point v of M is an extreme point, i.e., $v \in \mathcal{E}(M)$.

It turns out that support functionals for the epigraph of a convex function are directly related to tangent functionals (Definition 3.18) or subgradients (Definition 7.355) for this function:

Lemma 7.360 *Let X be a real normed space. Let M be any nonempty convex set in the vector space X^{td} and $f : M \to \mathbb{R}$ any convex function. For arbitrary $c \in \mathcal{R}$ and $x \in X$, define the linear functional $\Theta : X^{\text{td}} \times \mathbb{R} \to \mathbb{R}$ by*

$$\Theta(\varphi', s') \doteq \varphi'(x) - cs' , \qquad \varphi' \in X^{\text{td}}, \; s' \in \mathbb{R} .$$

If Θ is a support functional for the convex set $\text{epi}(f) \subseteq X^{\text{td}} \times \mathbb{R}$ (see Exercise 7.329), associated with the exposed point $(\varphi, s) \in \text{epi}(f)$, then $c > 0$, $s = f(\varphi)$ and $c^{-1}x \in \partial f(\varphi)$.

Proof If Θ is a support functional for $\text{epi}(f) \subseteq X^{\text{td}} \times \mathbb{R}$, associated with the exposed point $(\varphi, s) \in \text{epi}(f)$, then, for all $(\varphi', s') \in \text{epi}(f_1)$ with $(\varphi', s') \neq (\varphi, s)$, one has that

$$\varphi(x) - cs > \varphi'(x) - cs' .$$

By the definition of the epigraph (Definition 7.147), $c \geq 0$, because, otherwise, the above inequality would be violated by large enough s'. Additionally, $c \neq 0$ since one would otherwise get $\varphi(x) > \varphi(x)$ by taking $\varphi' = \varphi$ in the above inequality. As a consequence, $c > 0$. Taking again $\varphi' = \varphi$, we thus conclude that $s < s'$. Again by the definition of the epigraph (Definition 7.147), it follows that $s = f(\varphi)$. Hence, for all $\varphi' \in M$ with $\varphi' \neq \varphi$,

$$c^{-1}\varphi(x) - f(\varphi) > c^{-1}\varphi'(x) - f(\varphi') ,$$

that is,

$$f(\varphi) - \varphi(c^{-1}x) = \min_{\varphi' \in M} (f(\varphi') - \varphi'(c^{-1}x)) .$$

□

The first version of the Bishop-Phelps theorem is as follows:

Theorem 7.361 (Bishop-Phelps 1967) *Let X be any real Banach space and $M \subseteq X$ any nonempty set that is convex, closed, and bounded. Then the continuous support functionals for M form a dense subset of X^{td} with respect to the operator norm for linear functionals on the normed space X.*

The above theorem applied to the closed unit ball of the Banach space X directly leads to the following corollary, which gives an answer to the question mentioned at the beginning of this subsection, about the existence of norm-one vectors yielding the norm of a given continuous linear functional on X:

Corollary 7.362 *Let X be any real Banach space. Then there is a dense subset $\widetilde{X}^{\mathrm{td}}$ of the topological dual space X^{td} such that, for all $\varphi \in \widetilde{X}^{\mathrm{td}}$, there is $x \in X$ with $\|x\| = 1$, such that $\|\varphi\|_{\mathrm{op}} = \varphi(x)$.*

Observe that Theorem 7.361 cannot be applied to epigraphs, for such special convex sets are always unbounded. In particular, this version of the theorem is not adequate for the application in statistical mechanics that we have in mind. However, about 20 years later, Bishop and Phelps provided a version of their theorem for unbounded convex sets:

Theorem 7.363 (Bishop-Phelps 1989) *Let X be any real Banach space and $M \subseteq X$ any nonempty set that is convex and closed (but not necessarily bounded). Then the support functionals for M in the convex cone*

$$C_K \doteq \{\varphi \in X^{\mathrm{td}} \ : \ \sup \varphi(M) < \infty\} \subseteq X^{\mathrm{td}}$$

form a dense subset of this cone.

Recall the epigraphs of convex lower semicontinuous functions are always convex and closed, thanks to Lemma 7.148 and Exercise 7.329. Applying the last theorem to this type of convex set, we arrive at the following proposition:

Proposition 7.364 *Let X be any real Banach space and $M \subseteq X$ any nonempty set that is convex and closed. Let $f : M \to \mathbb{R}$ be any convex lower semicontinuous function. Take a continuous linear functional $\varphi_0 \in X^{\text{td}}$ so that $f - \varphi_0$ is bounded from below in M. Then, for any $\varepsilon > 0$, there are $x \in M$ and a subgradient $\varphi \in \partial f(x) \subseteq X^{\text{td}}$, such that $\|\varphi - \varphi_0\|_{\text{op}} \leq \varepsilon$.*

Proof Consider the real Banach space $\tilde{X} \doteq X \times \mathbb{R}$ and recall that $\text{epi}(f)$ is convex and closed. Define the linear functional $\tilde{\varphi}_0 \in \tilde{X}^{\text{td}}$ by

$$\tilde{\varphi}_0(x, s) \doteq \varphi_0(x) - s , \qquad x \in X , \; s \in \mathbb{R} .$$

As the functional $f - \varphi_0$ is by assumption bounded from below in M, one has

$$\sup \tilde{\varphi}_0(\text{epi}(f)) = \sup\{\varphi_0(x) - s : x \in M , \; s \geq f(x)\} \leq \sup_{x \in M} \{\varphi_0(x) - f(x)\} < \infty .$$

Thus, by the Bishop-Phelps theorem (Theorem 7.363) and Lemma 7.360, there is a sequence of support functionals $\tilde{\varphi}_n \in \tilde{X}^{\text{td}}$, $n \in \mathbb{N}$, for $\text{epi}(f)$, respectively, associated with exposed points $(x_n, f(x_n)) \in \text{epi}(f)$, $x_n \in M$, $n \in \mathbb{N}$, such that

$$\lim_{n \to \infty} \|\tilde{\varphi}_n - \tilde{\varphi}_0\|_{\text{op}} = 0 .$$

Remark that $\tilde{\varphi}_n \in \tilde{X}^{\text{td}}$ is always of the form

$$\tilde{\varphi}_n(x, s) \doteq \varphi_n(x) - c_n s , \qquad x \in X , \; s \in \mathbb{R} ,$$

for some $\varphi_n \in X^{\text{td}}$ and $c_n > 0$. In particular,

$$\lim_{n \to \infty} \|\varphi_n - \varphi_0\|_{\text{op}} = 0 , \qquad \lim_{n \to \infty} |c_n - 1| = 0 ,$$

that is,

$$\lim_{n \to \infty} \left\| c_n^{-1} \varphi_n - \varphi_0 \right\|_{\text{op}} = 0 .$$

Finally, recall from Lemma 7.360 that $c_n^{-1} \varphi_n \in \partial f(x_n)$. □

For the special case $M = X$ with f being continuous, the last proposition yields the following corollary, which is the version of the Bishop-Phelps theorem needed for the proof of existence of phase transitions in Chap. 5. As mentioned above, for completeness, we provide a direct proof of the corollary, because it is one of the main arguments of the proof of existence of phase transitions, and we do not give any proof of Theorem 7.363 above.

Corollary 7.365 (Bishop-Phelps Theorem—Version for Convex Functions I) *Let X be any real Banach space and $f : X \to \mathbb{R}$ a* convex continuous *function. Let*

$\varphi_0 \in X^{\mathrm{td}}$ be any continuous linear functional such that, for some constant $c < \infty$ and all $x \in X$,

$$f(x) \geq \varphi_0(x) - c\ .$$

Then, for all $\varepsilon > 0$, there are $x_\varepsilon \in X$ and a linear functional $\varphi_\varepsilon \in X^{\mathrm{td}}$, which is tangent to f at x_ε, i.e., $\varphi_\varepsilon \in \partial f(x_\varepsilon)$, and satisfies $\|\varphi_\varepsilon - \varphi_0\|_{\mathrm{op}} \leq \varepsilon$.

Proof We give here a direct proof, without invoking Theorem 7.363 nor Proposition 7.364. For simplicity and without loss of generality, we assume that $\varphi_0 = 0$ and $c = 0$. For the more general case, it then suffices to consider the convex continuous function $f' \doteq f - \varphi_0 + c$ instead of f. Fix the parameter $\varepsilon > 0$.

1. For all $x \in X$, define the set

$$C_x \doteq \{x' \in X \ :\ f(x') \leq f(x) - \varepsilon\|x' - x\|\} \supseteq \{x\}\ .$$

By continuity of the function f, note that, for all $x \in X$, C_x is a closed subset of X. Additionally, for all $x_0 \in X$, $x_1 \in C_{x_0}$, and $x_2 \in C_{x_1}$,

$$\begin{aligned} f(x_2) &\leq f(x_1) - \varepsilon\|x_1 - x_2\| \\ &\leq (f(x_0) - \varepsilon\|x_0 - x_1\|) - \varepsilon\|x_1 - x_2\| \\ &\leq f(x_0) - \varepsilon\|x_0 - x_2\|\ , \end{aligned}$$

by the triangle inequality. Hence, $C_{x_1} \subseteq C_{x_0}$ for all $x_0 \in X$ and $x_1 \in C_{x_0}$.

2. There is $\tilde{x} \in X$ such that $C_{\tilde{x}} = \{\tilde{x}\}$. In order to prove this fact, choose any $x_0 \in X$ and a sequence $(x_n)_{n\in\mathbb{N}} \subseteq X$ such that, for all $n \in \mathbb{N}_0$, $x_{n+1} \in C_{x_n}$ and

$$f(x_{n+1}) < 2^{-n}\varepsilon + \inf\{f(x)\ :\ x \in C_{x_n}\}\ .$$

Observe that the above infimum exists, as f is (by assumption) a positive-valued function. For all $n \in \mathbb{N}_0$ and $x' \in C_{x_{n+1}} \subseteq C_{x_n}$,

$$f(x_{n+1}) - 2^{-n}\varepsilon < f(x') \leq f(x_{n+1}) - \varepsilon\|x' - x_{n+1}\|\ .$$

In particular, $\|x' - x_{n+1}\| < 2^{-n}$ for all $x' \in C_{x_{n+1}}$. From this, we conclude that $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, as $x_{n+2} \in C_{x_{n+1}}$. Let

$$\tilde{x} \doteq \lim_{n\to\infty} x_n\ .$$

As C_{x_n} is closed for all $n \in \mathbb{N}$, one has that $\tilde{x} \in C_{x_n}$ for all $n \in \mathbb{N}$. In particular,

$$\tilde{x} \in C_{\tilde{x}} \subseteq \bigcap_{n\in\mathbb{N}} C_{x_n}\ .$$

Let $x' \in C_{\tilde{x}}$. As $x' \in C_{x_{n+1}}$ for all $n \in \mathbb{N}$, one has $\|x' - x_{n+1}\| \leq 2^{-n}$ for all $n \in \mathbb{N}$. Therefore, $x' = \tilde{x}$ and, thus, $C_{\tilde{x}} = \{\tilde{x}\}$.

3. Consider the Banach space $X \times \mathbb{R}$. Define the nonempty convex subsets $M, M' \subseteq X \times \mathbb{R}$, respectively, by

$$M \doteq \{(x, s) \in X \times \mathbb{R} \ : \ s < f(\tilde{x}) - \varepsilon\|x - \tilde{x}\|\} ,$$
$$M' \doteq \{(x, s) \in X \times \mathbb{R} \ : \ s \geq f(x)\} .$$

Note that M is an open subset of $X \times \mathbb{R}$. The equality $C_{\tilde{x}} = \{\tilde{x}\}$ of the previous step implies that

$$f(x) \geq f(\tilde{x}) - \varepsilon\|x - \tilde{x}\| \tag{7.12}$$

for every $x \in X$. Hence,

$$M \cap M' \subseteq \{(x, s) \in X \times \mathbb{R} \ : \ f(x) < f(\tilde{x}) - \varepsilon\|x - \tilde{x}\|\} = \emptyset .$$

Therefore, by the Hahn-Banach separation theorem (Proposition 7.331 (i)), there is a continuous linear functional Θ on $X \times \mathbb{R}$ such that

$$\Theta(x, s) < \Theta(x', s') , \qquad (x, s) \in M, \ (x', s') \in M' . \tag{7.13}$$

In particular,

$$-\infty < \sup \Theta(M) \leq \inf \Theta(M') < \infty .$$

In other words, there is $T \in [\sup \Theta(M), \inf \Theta(M')]$ such that

$$\Theta(x, s) \leq T \leq \Theta(x', s') , \qquad (x, s) \in M, \ (x', s') \in M' .$$

4. General continuous linear functionals Θ on $X \times \mathbb{R}$ have the form

$$\Theta(x, s) = bs - \varphi(x) , \qquad x \in X, \ s \in \mathbb{R},$$

where $b \in \mathbb{R}$, while $\varphi \in X^{\mathrm{td}}$ is a continuous linear functional on X. Without loss of generality, we can assume that $\Theta(x, s) = s - \varphi(x)$ (i.e., $b = 1$), where $\varphi \in X^{\mathrm{td}}$. In fact, using $x' = x$, $s' = f(x)$, and the limit $s \to f(\tilde{x}) - \varepsilon\|x - \tilde{x}\|$, together with (7.12) and (7.13), it is easy to see that $b > 0$, and we can always replace Θ with $b^{-1}\Theta$ in order to have such a normalization. Doing this, we meanwhile arrive at the inequalities

$$f(\tilde{x}) - \varepsilon\|x - \tilde{x}\| - \varphi(x) \leq T \leq f(x) - \varphi(x)$$

for every $x \in X$. Setting $x = \tilde{x}$ in the above double inequality, one gets $T = f(\tilde{x}) - \varphi(\tilde{x})$, and it follows that

$$\varphi(\tilde{x}) - \varphi(x) = \varphi(\tilde{x} - x) \leq \varepsilon \|\tilde{x} - x\| ,$$

as well as

$$f(\tilde{x}) - \varphi(\tilde{x}) + \varphi(x) = f(\tilde{x}) + \varphi(x - \tilde{x}) \leq f(x) = f((x - \tilde{x}) + \tilde{x})$$

for every $x \in X$. The first inequality implies that $\|\varphi\|_{\mathrm{op}} \leq \varepsilon$ and the second one that φ is tangent to f at $\tilde{x}$. See Definition 3.18.

□

To conclude, we give below a more general version of the last corollary, used in Sect. 5.6 to prove the existence of spontaneous symmetry breaking, which is a physically very important case of phase transition.

Proposition 7.366 (Bishop-Phelps Theorem—Version for Convex Functions II) *Let X be any real Banach space, $f : X \to \mathbb{R}$ a* convex continuous *function, $C \subseteq X$ a* closed convex cone*, $\varepsilon > 0$, and $\varphi_0 \in X^{\mathrm{td}}$ a linear functional such that*

$$f(x) \geq \varphi_0(x) - c$$

for some $c \in \mathbb{R}$ and all $x \in X$. Then there are a vector $\tilde{x} \in C$ and a linear functional $\tilde{\varphi} \in X^{\mathrm{td}}$ that is tangent to f at $\tilde{x}$, i.e., $\tilde{\varphi} \in \partial f(\tilde{x})$, such that

$$\tilde{\varphi}(y) \geq \varphi_0(y) - \varepsilon \|y\|$$

for all $y \in C$.

Proof The proof of this proposition is an adaptation of the one of Corollary 7.365: In that proof, in Point 1, use

$$C_x \doteq \{x' \in C \ : \ f(x') \leq f(x) - \varepsilon \|x' - x\|\} \supseteq \{x\} , \qquad x \in X ,$$

and in Point 3, take

$$M \doteq \{(x, s) \in X \times \mathbb{R} \ : \ s < f(\tilde{x}) - \varepsilon \|x - \tilde{x}\|\} ,$$
$$M' \doteq \{(x, s) \in C \times \mathbb{R} \ : \ s \geq f(x)\} .$$

We omit the details. □

Notice that if the closed convex cone $C \subseteq X$ is a vector subspace, then we have

$$\tilde{\varphi}(y) \geq \varphi_0(y) - \varepsilon \|y\| \qquad \text{and} \qquad -\tilde{\varphi}(y) \geq -\varphi_0(y) - \varepsilon \|y\|$$

for all $y \in C$, which implies that

$$\|\varphi_\varepsilon|_C - \varphi_0|_C\|_{\mathrm{op}} \leq \varepsilon \ .$$

In particular, Corollary 7.365 is a special case of the last proposition.

References

1. J.-B. Bru, W. de Siqueira Pedra, Non-cooperative equilibria of Fermi systems with long range interactions. Memoirs of the AMS **224**(1052) (2013)
2. R.A. Ryan, *Introduction to Tensor Products of Banach Spaces* (Springer, New York, 2002)
3. C.D. Aliprantis, K.C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd edn. (Springer, New York, 2007)
4. C.D. Aliprantis, O. Burkinshaw, *Positive Operators (Pure and Applied Mathematics)*, vol. 119 (Academic Press, New York, 1985)
5. W.A.J. Luxemburg, A.C. Zaanen, *Riesz Spaces I* (Elsevier, Amsterdam, 1971)
6. S. Sherman, Order in operator algebras. Am. J. Math. **73**(1), 227–232 (1951)
7. A. Doering, C. Isham, 'What is a thing?': topos theory in the foundations of physics, in *New Structures for Physics*, ed. by B. Coecke. Springer Lecture Notes in Physics, vol. 813 (Springer, Berlin, 2011), pp. 753–940
8. O. van Gaans, A. Kalauch, Ideals and bands in pre-Riesz spaces. Positivity **12**, 591–611 (2008)
9. J. van Dobben de Bruyn, *Connections Between the General Theories of Ordered Vector Spaces and C^*-Algebras.* Master Thesis, Mathematical Institute, Leiden University (2018)
10. S. Awodey, *Category Theory*, 2nd edn. (Oxford University, Oxford, 2010)
11. L. Barreira, C. Holanda, Higher-dimensional nonlinear thermodynamic formalism. J. Stat. Phys. **187**(2), 18 (2022)
12. R.R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, 2nd edn. (Springer, Berlin, 1993)
13. R.R. Phelps, *Lectures on Choquet's Theorem*, in Lecture Notes in Mathematics, vol. 1757 (Springer, Berlin, 2001)
14. J.R. Ringrose, R.V. Kadison, *Fundamentals of the Theory of Operator Algebras*, in Advanced Theory (Graduate Studies in Mathematics, vol. 16), vol. 2 (American Mathematical Society, New York, 1997)
15. H. Araki, H. Moriya, Equilibrium statistical mechanics of fermion lattice systems. Rev. Math. Phys. **15**, 93–198 (2003)
16. E.T. Poulsen, A simplex with dense extreme boundary. Ann. Inst. Fourier (Grenoble) **11**, 83–87 (1961)
17. J. Lindenstrauss, G.H. Olsen, Y. Sternfeld, The Poulsen simplex. Ann. Inst. Fourier (Grenoble) **28**, 91–114 (1978)
18. W. Rudin, *Functional Analysis* (McGraw-Hill Science, New York, 1991)
19. W. Rudin, *Real and Complex Analysis*, 3rd edn. (McGraw-Hill Science, New York, 1987)

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1

20. E.M. Alfsen, *Compact Convex Sets and Boundary Integrals (Ergebnisse der Mathematik und ihrer Grenzgebiete. 2. Folge, 57)*, 1st edn. (Springer, Berlin, 1971)
21. J.-B. Bru, W. de Siqueira Pedra, *Classical Dynamics from Self-Consistency Equations in Quantum Mechanics—Extended Version*, arXiv:2009.04969 (2020)
22. J. Dixmier, *C^*-Algebras* (North-Holland Publishing Company, Amsterdam, 1977)
23. J.-B. Bru, W. de Siqueira Pedra, Weak* hypertopologies with application to genericity of convex sets. Journal of Convex Analysis **29**(1), 13–60 (2021)
24. T.C. Lim, The center of a convex set. Proc. AMS **81**(2), 345–346 (1981)
25. J.-B. Bru, W. de Siqueira Pedra, Remarks on the Γ-regularization of Non-convex and Non-Semi-Continuous functionals on topological vector spaces. J. Convex Analysis **19**(3), 467–483 (2012)
26. J.-B. Bru, W. de Siqueira Pedra, Effect of a locally repulsive interaction on s–wave superconductors. Rev. Math. Phys. **22**(3), 233–303 (2010)
27. H. Komiya, Elementary Proof For Sion's minimax theorem. Kodai Math. J. **11**(1), 5–7 (1988)
28. J.-B. Bru, W. de Siqueira Pedra, K. Rodrigues Alves, From Short-Range to Mean-Field Models in Quantum Lattices, submitted preprint (2022). See arXiv:2203.01021 [math-ph] (52 pages)
29. L.N. Cooper, Bound electron pairs in a degenerate Fermi gas. Phys. Rev. **104**, 1189–1190 (1956)
30. J. Bardeen, L.N. Cooper, J.R. Schrieffer, Microscopic theory of superconductivity. Phys. Rev. **106**, 162–164 (1957)
31. J. Bardeen, L.N. Cooper, J.R. Schrieffer, Theory of superconductivity. Phys. Rev. **108**, 1175–1204 (1957)
32. N.N. Bogoliubov, On the theory of superfluidity. J. Phys. (USSR) **11**, 23–32 (1947)
33. N.N. Bogoliubov, On some problems of the theory of superconductivity. Physica **26**, S1–S16 (1960)
34. N.N. Bogoliubov Jr., *A Method for Studying Model Hamiltonians* (Oxford, Pergamon, 1977)
35. N.N. Bogoliubov Jr., On model dynamical systems in statistical mechanics. Physica **32**, 933 (1966)
36. N.N. Bogoliubov Jr., J.G. Brankov, V.A. Zagrebnov, A.M. Kurbatov, N.S. Tonchev, Metod approksimiruyushchego gamil'toniana v statisticheskoi fizike (The Approximating Hamiltonian Method in Statistical Physics.). Sofia: Izdat. Bulgar. Akad. Nauk (Publ. House Bulg. Acad. Sci.) (1981)
37. N.N. Bogoliubov Jr., J.G. Brankov, V.A. Zagrebnov, A.M. Kurbatov, N.S. Tonchev, Some classes of exactly soluble models of problems in Quantum Statistical Mechanics: the method of the approximating Hamiltonian. Russ. Math. Surv. **39**, 1–50 (1984)
38. J.G. Brankov, D.M. Danchev, N.S. Tonchev, *Theory of Critical Phenomena in Finite–size Systems: Scaling and Quantum Effects* (Word Scientific, Singapore–New Jersey–London–Hong Kong, 2000)
39. V.A. Zagrebnov, J.-B. Bru, The Bogoliubov model of weakly imperfect Bose gas. Phys. Rep. **350**, 291–434 (2001)
40. J. Lebowitz, O. Penrose, A Rigorous treatment of the van der Waals-Maxwell theory of the vapor-liquid transition. J. Math. Phys. **7**, 98 (1966)
41. O. Penrose I, J.L. Lebowitz, Rigorous treatment of metastable states in the van der Waals-Maxwell theory. J. Stat. Phys. **3**(2), 211–236 (1971)
42. P.C. Hemmer, J.L. Lebowitz, Systems with weak long-range potentials, in *Phase Transitions and Critical Phenomena (Volume 5b)*, ed. by C. Domb, M.S. Green (Academic Press Inc, New York, 1976), pp 107–203
43. E. Lieb, Quantum-mechanical extension of the Lebowitz-Penrose theorem on the Van Der Waals theory. J. Math. Phys. **7**(6), 1016–1024 (1966)
44. E. Presutti, *Scaling Limits in Statistical Mechanics and Microstructures in Continuum Mechanics* (Springer, Berlin, 2009)
45. S. Franz, F.L. Toninelli, Kac limit for finite-range spin glasses. Phys. Rev. Lett. **92**, 030602 (3 pages) (2004)

46. S. Franz, F.L. Toninelli, Finite-range spin glasses in the Kac limit: free energy and local observables. J. Phys. A: Math. Gen. **37**, 7433–7446 (2004)
47. S. Franz, Spin glass models with Kac interactions. Eur. Phys. J. B **64**, 557–561 (2008)
48. J.-B. Bru, W. de Siqueira Pedra, Classical dynamics generated by long-range interactions for lattice fermions and quantum spins. J. Math. Anal. Appl. **493**(1), 124434 (pp 61) (2021)
49. J.-B. Bru, W. de Siqueira Pedra, Quantum dynamics generated by long-range interactions for lattice-fermion and quantum spins. J. Math. Anal. Appl. **493**(1), 124517 (pp 65) (2021)
50. J.-B. Bru, W. de Siqueira Pedra, Entanglement of classical and quantum short-range dynamics in mean-field systems. Ann. Phys. **434**, 168643 (pp 31) (2021)
51. O. Brattelli, D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, vol. I, 2nd edn. (Springer, New York, 1996)
52. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*. Functional Analysis, vol. I (Academic Press, New York-London, 1980)
53. S. Mazur, Über konvexe Menge in linearen normierten Raumen. Studia. Math. **4**, 70–84 (1933)
54. O.E. Lanford III, D.W. Robinson, Statistical mechanics of quantum spin systems. III. Commun. Math. Phys. **9**, 327–338 (1968)
55. O. Bratteli, D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, vol. II, 2nd ed. (Springer, New York, 1997)
56. J. Dixmier, *Les C^*-algè bres et leurs représentations* (Gauthier-Villars, Paris, 1969)
57. H. Bauer, Measure and integration theory, in *de Gruyter Studies in Mathematics*, vol. 26 (Walter de Gruyter GmbH & Co, Berlin, 2001)
58. S. Kochen, E.P. Specker, The problem of hidden variables in quantum mechanics. J. Math. Mech. **17**(1), 59–87 (1967)
59. J.S. Bell, On the problem of hidden variables in quantum mechanics. Rev. Mod. Phys. **38**(3), 447–452 (1966)
60. M. Takesaki, *Theory of Operator Algebras I* (Springer, Heidelberg, 2002)
61. M. Takesaki, *Theory of Operator Algebras II* (Springer, Heidelberg, 2003)
62. M. Takesaki, *Theory of Operator Algebras III* (Springer, Heidelberg, 2003)
63. O.A. Nielsen, *Direct Integral Theory*. Lecture Notes in Pure and Applied Mathematics, vol. 61 (Marcel Dekker, New York, 1980)
64. J.B. Conway, *A Course in Functional Analysis*, 2nd edn. (Springer, New York, 1997)
65. G. Emch, *Algebraic Methods in Statistical Mechanics and Quantum Field Theory* (Willey–Interscience, New York, 1972)
66. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*. Functional Analysis, vol. I (Academic Press, New York, 1980)
67. B. Blackadar, *Operator Algebras: Theory of C^*-Algebras and von Neumann Algebras*. Encyclopaedia of Mathematical Sciences, Operator Algebras and Non-Commutative Geometry III (Springer, Heidelberg, 2006)
68. B. Blackadar, Shape theory for C^*-Algebras. Math. Scand. **56**, 249–275 (1985)
69. H. Araki, On Quasifree states of CAR and Bogoliubov Automorphisms. Publ. RIMS, Kyoto Univ. **6**, 385–442 (1970/71)
70. R.B. Israel, *Convexity in the Theory of Lattice Gases* (Princeton Series in Physics, Princeton University Press, Princeton, 1979)
71. M. Sh. Birman, M.Z. Solomjak, *Spectral Theory of Self-Adjoint Operators in Hilbert Space*. Mathematics and Its Applications (Springer/D. Reidel Publishing Company, Netherlands/Dordrecht,1987)
72. H. Araki, On the diagonalization of a bilinear Hamiltonian by a Bogoliubov transformation. Publ. RIMS Kyoto Univ. Ser. A **4**, 387–412 (1968)
73. E.H. Lieb, Thomas-Fermi and Hartree-Fock theory, in *Proceedings International Congress Mathematicians* (Canadian Mathematical Society, Canada, 1975), pp. 383–386
74. E.H. Lieb, B. Simon, The Hartree-Fock theory for Coulomb systems. Commun. Math. Phys. **53**, 185–194 (1977)
75. E.H. Lieb, B. Simon. The Thomas-Fermi theory of atoms, molecules and solids. Adv. Math. **23**, 22–116 (1977)

76. E.H. Lieb, *The stability of matter: From Atoms to Stars*, 4th edn., ed. by W. Thirring (Springer, New York, 2005)
77. V. Bach, E.H. Lieb, J.P. Solovej, Generalized Hartree-Fock theory and the Hubbard model. J. Stat. Phys. **76**(1/2), 3–89 (1994)
78. V. Bach, Hartree–Fock Theory, Lieb's Variational Principle, and their Generalizations, arXiv:2209.10189 [quant-ph] (2022)
79. V. Bach, J. Poelchau, Accuracy of the Hartree-Fock approximation for the Hubbard model. J. Math. Phys. **38**, 2072–2083 (1997)
80. V. Bach, E.H. Lieb, M.V. Travaglia, Ferromagnetism of the hubbard model at strong coupling in the Hartree-Fock approximation. Rev. Math. Phys. **18**, 519–543 (2006)
81. J. Wojtkiewicz, P.H. Chankowski, Ground state energies of the Hubbard models and the Hartree Fock approximation, arXiv:2209.02361 [math-ph] (2022)
82. N. Benedikter, M. Porta, B. Schlein, *Effective Evolution Equations from Quantum Dynamics*. SpringerBriefs in Mathematical Physics, Series Volume **7** (2016)
83. C. Bardos, F. Golse, A.D. Gottlieb, N.J. Mauser, Mean-field dynamics of fermions and the time-dependent Hartree-Fock equation. J. Math. Pures Appl. **82**(6), 665–683 (2003)
84. C. Bardos, F. Golse, A.D. Gottlieb, N.J. Mauser, Accuracy of the time-dependent Hartree-Fock approximation for uncorrelated initial states. J. Stat. Phys. **115**(3–4), 1037–1055 (2004)
85. A. Elgart, L. Erdös, B. Schlein, H.T. Yau, Nonlinear Hartree equation as the mean-field limit of weakly coupled fermions. J. Math. Pures Appl. **83**(10), 1241–1273 (2004)
86. J. Fröhlich, A. Knowles, A microscopic derivation of the time-dependent Hartree-Fock equation with Coulomb two-body interaction. J. Stat. Phys. **145**(1), 23–50 (2011)
87. N. Benedikter, M. Porta, B. Schlein, Mean-field evolution of fermionic systems. Commun. Math. Phys. **331**, 1087–1131 (2014)
88. N. Benedikter, M. Porta, B. Schlein, Mean-field dynamics of fermions with relativistic dispersion. J. Math. Phys. **55**, 021901 (2014)
89. N. Benedikter, V. Jakšić, M. Porta, C. Saffirio, B. Schlein, Mean-field evolution of fermionic mixed states. Commun. Pure Appl. Math. **69**(12), 2250–2303 (2016)
90. S. Petrat, P. Pickl, A new method and a new scaling for deriving fermionic mean-field dynamics. Math. Phys. Anal. Geom. **19**(3), 1–51 (2016)
91. V. Bach, S. Breteaux, S. Petrat, P. Pickl, T. Tzaneteas, Kinetic energy estimates for the accuracy of the time-dependent Hartree-Fock approximation with Coulomb interaction. J. Math. Pures Appl. **105**(1), 1–30 (2016)
92. M. Porta, S. Rademacher, C. Saffirio, B. Schlein, Mean field evolution of fermions with Coulomb interaction. J. Stat. Phys. **166**(6), 1345–1364 (2017)
93. M.Ya. Amusia, A.Z. Msezane, V.R. Shaginyan, Density functional theory versus the Hartree-Fock method: comparative assessment. Physica Scripta. **T68**, C133–140 (2003)
94. J.-B. Bru, W. de Siqueira Pedra, *Lieb-Robinson Bounds for Multi-Commutators and Applications to Response Theory*. SpringerBriefs in Mathematical Physics, vol. 13 (2017)
95. F.A. Berezin, *The Method of Second Quantization* (Academic Press, New York, 1966)
96. J.-B. Bru, N.J.B. Aza, W. de Siqueira Pedra, L. Mü ssnich, Large deviations in weakly interacting fermions: Generating functions as Gaussian Berezin integrals and bounds on large Pfaffians. Rev. Math. Phys. **3**, 2150034 (73 pp) (2021). https://doi.org/10.1142/S0129055X21500343
97. J.-B. Bru, W. de Siqueira Pedra, K. Rodrigues Alves, From Short-Range to Mean-Field Models in Quantum Lattices, to be published (2022). See arXiv:2203.01021 [math-ph] (52 pages)
98. J. Benoist, J.-B. Hiriart-Urruty, *What is the subdifferential of the closed convex hull of a function?* SIAM J. Math. Anal. **27**, 1661–1679 (1996)
99. E. Zeidler, *Nonlinear Functional Analysis and Its Applications III: Variational Methods and Optimization* (Springer, New York, 1985)
100. R.V. Kadison, Order properties of bounded self-adjoint operators. Proc. Am. Math. Soc. **2**(3), 505–510 (1951)
101. J.D. Weston, The decomposition of a continuous linear functional into non-negative components. Math. Scand. **5**, 54–56 (1957)

Index

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

J.-B. Bru, W. Alberto de Siqueira Pedra, *C*-Algebras and Mathematical Foundations of Quantum Statistical Mechanics*, Latin American Mathematics Series – UFSCar subseries, https://doi.org/10.1007/978-3-031-28949-1