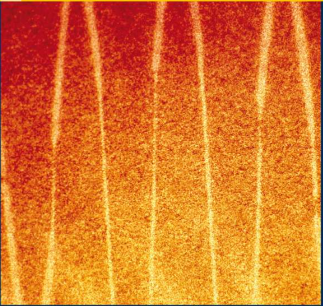
Valter Moretti

Spectral Theory and Quantum Mechanics

With an Introduction to the Algebraic Formulation



UNITEXT

Dedicated to those young and brilliant colleagues, mathematicians and physicists, forced to flee Italy for other countries in order to give their contribution, big or small, to scientific research.

UNITEXT – La Matematica per il 3+2

Volume 64

For further volumes:
http://www.springer.com/series/5418

Valter Moretti

Spectral Theory and Quantum Mechanics

With an Introduction to the Algebraic Formulation

Springer

Valter Moretti
Department of Mathematics
University of Trento

Translated by: Simon G. Chiossi, Department of Mathematics, Politecnico di Torino

Translated and extended version of the original Italian edition: V. Moretti, Teoria Spettrale e Meccanica Quantistica, © Springer-Verlag Italia 2010

UNITEXT – La Matematica per il 3+2
ISSN 2038-5722 ISSN 2038-5757 (electronic)

ISBN 978-88-470-2834-0 ISBN 978-88-470-2835-7 (eBook)
DOI 10.1007/978-88-470-2835-7

Library of Congress Control Number: 2012945983

Springer Milan Heidelberg New York Dordrecht London

9 8 7 6 5 4 3 2 1

Cover-Design: Beatrice Ꟊ, Milano

Typesetting with LaTeX: PTP-Berlin, Protago TeX-Production GmbH, Germany (www.ptp-berlin.de)
Printing and Binding: Grafiche Porpora, Segrate (MI)

Printed in Italy

Springer-Verlag Italia S.r.l., Via Decembrio 28, I-20137 Milano
Springer is a part of Springer Science+Business Media (www.springer.com)

Preface

I must have been 8 or 9 when my father, a man of letters but well-read in every discipline and with a curious mind, told me this story: "A great scientist named Albert Einstein discovered that any object with a mass can't travel faster than the speed of light". To my bewilderment I replied, boldly: "This can't be true, if I run almost at that speed and then accelerate a little, surely I will run faster than light, right?". My father was adamant: "No, it's impossible to do what you say, it's a known physics fact". After a while I added:"That bloke, Einstein, must've checked this thing many times … how do you say, he did many experiments?". The answer I got was utterly unexpected: "No, not even one I think, he used maths!".

What did numbers and geometrical figures have to do with the existence of a limit speed? How could one stand behind such an apparently nonsensical statement as the existence of a maximum speed, although certainly true (I trusted my father), just based on maths? How could mathematics have such big a control on the real world? And physics? What on earth was it, and what did it have to do with maths? This was one of the most beguiling and irresistible things I had ever heard till that moment … I had to find out more about it.

This is an extended and enhanced version of an existing textbook written in Italian (and published by Springer-Verlag). That edition and this one are based on a common part that originated, in preliminary form, when I was a *Physics* undergraduate at the University of Genova. The third-year compulsory lecture course called *Institutions of Theoretical Physics* was the second exam that had us pupils seriously climbing the walls (the first being the famous *Physics II*, covering thermodynamics and classical electrodynamics).

Quantum Mechanics, taught in that course, elicited a novel and involved way of thinking, a true challenge for craving students: for months we hesitantly faltered on a hazy and uncertain terrain, not understanding what was really key among the notions we were trying – struggling, I should say – to learn, together with a completely new formalism: linear operators on Hilbert spaces. At that time, actually, we did not realise we were using this mathematical theory, and for many mates of mine the matter would have been, rightly perhaps, completely futile; Dirac's *bra* vectors were what they were, and that's it! They were certainly not elements in the topological dual of the Hilbert space. The notions of *Hilbert space* and *dual topological space* had no right of abode in the mathematical toolbox of the majority of my fellows, even if they would soon come back in throught the back door, with the course *Mathematical Methods of Physics* taught by prof. G. Cassinelli. Mathematics, and the mathematical formalisation of physics, had always been my flagship to overcome the difficulties that studying physics presented me with, to the point that eventually (after a Ph.D. in theoretical physics) I officially became a mathematician. Armed with a maths background – learnt in an extracurricular course of study that I cultivated over the years, in parallel to academic physics – and eager to broaden my knowledge, I tried to formalise every notion I met in that new and riveting lecture course. At the same time I was carrying along a similar project for the mathematical formalisation of General Relativity, unaware that the work put into Quantum Mechanics would have been incommensurably bigger.

The formulation of the spectral theorem as it is discussed in § 8, 9 is the same I learnt when taking the Theoretical Physics exam, which for this reason was a dialogue of the deaf. Later my interest turned to *quantum field theory*, a topic I still work on today, though in the slightly more general framework of quantum field theory *in curved spacetime*. Notwithstanding, my fascination with the elementary formulation of Quantum Mechanics never faded over the years, and time and again chunks were added to the opus I begun writing as a student.

Teaching Master's and doctoral students in mathematics and physics this material, thereby inflicting on them the result of my efforts to simplify the matter, has proved to be crucial for improving the text; it forced me to typeset in LaTeX the pile of loose notes and correct several sections, incorporating many people's remarks.

Concerning this I would like to thank my colleagues, the friends from the *newsgroups it.scienza.fisica*, *it.scienza.matematica* and *free.it.scienza.fisica*, and the many students – some of which are now fellows of mine – who contributed to improve the preparatory material of the treatise, whether directly of not, in the course of time: S. Albeverio, P. Armani, G. Bramanti, S. Bonaccorsi, A. Cassa, B. Cocciaro, G. Collini, M. Dalla Brida, S. Doplicher, L. Di Persio, E. Fabri, C. Fontanari, A. Franceschetti, R. Ghiloni, A. Giacomini, V. Marini, S. Mazzucchi, E. Pagani, E. Pelizzari, G. Tessaro, M. Toller, L. Tubaro, D. Pastorello, A. Pugliese, F. Serra Cassano, G. Ziglio, S. Zerbini. I am indebted, for various reasons also unrelated to the book, to my late colleague Alberto Tognoli. My greatest appreciation goes to R. Aramini, D. Cadamuro and C. Dappiaggi, who read various versions of the manuscript and pointed out a number of mistakes.

I am grateful to my friends and collaborators R. Brunetti, C. Dappiaggi and N. Pinamonti for lasting technical discussions, suggestions on many topics covered and for pointing out primary references.

Lastly I would like to thank E. Gregorio for the invaluable and on-the-spot technical help with the LaTeX package.

In the transition from the original Italian to the expanded English version a massive number of (uncountably many!) typos and errors of various kind have been amended. I owe to E. Annigoni, M. Caffini, G. Collini, R. Ghiloni, A. Iacopetti, M. Oppio and D. Pastorello in this respect. Fresh material was added, both mathematical and physical, including a chapter, at the end, on the so-called *algebraic formulation*.

In particular, Chapter 4 contains the proof of Mercer's theorem for positive Hilbert–Schmidt operators. The now-deeper study of the first two axioms of Quantum Mechanics, in Chapter 7, comprises the algebraic characterisation of quantum states in terms of positive functionals with unit norm on the C^*-algebra of compact operators. General properties of C^*-algebras and *-morphisms are introduced in Chapter 8. As a consequence, the statements of the spectral theorem and several results on functional calculus underwent a minor but necessary reshaping in Chapters 8 and 9. I incorporated in Chapter 10 (Chapter 9 in the first edition) a brief discussion on abstract differential equations in Hilbert spaces. An important example concerning Bargmann's theorem was added in Chapter 12 (formerly Chapter 11). In the same chapter, after introducing the Haar measure, the Peter–Weyl theorem on unitary representations of compact groups is stated, and partially proved. This is then applied to the theory of the angular momentum. I also thoroughly examined the superselection rule for the angular momentum. The discussion on POVMs in Chapter 13 (formerly Chapter 12) is enriched with further material, and I included a primer on the fundamental ideas of non-relativistic scattering theory. Bell's inequalities (Wigner's version) are given considerably more space. At the end of the first chapter basic point-set topology is recalled together with abstract measure theory. The overall effort has been to create a text as self-contained as possible. I am aware that the material presented has clear limitations and gaps. Ironically – my own research activity is devoted to relativistic theories – the entire treatise unfolds at a non-relativistic level, and the quantum approach to Poincaré's symmetry is left behind.

I thank my colleagues F. Serra Cassano, R. Ghiloni, G. Greco, A. Perotti and L. Vanzo for useful technical conversations on this second version. For the same reason, and also for translating this elaborate opus to English, I would like to thank my colleague S.G. Chiossi.

Trento, September 2012 *Valter Moretti*

This page intentionally left blank

Contents

1

Introduction and mathematical backgrounds

"O frati", dissi "che per cento milia
perigli siete giunti a l'occidente,
a questa tanto picciola vigilia
d'i nostri sensi ch'è del rimanente
non vogliate negar l'esperienza,
di retro al sol, del mondo sanza gente".
Dante Alighieri, The Divine Comedy, Inferno XXVI[1]

1.1 On the book

1.1.1 Scope and structure

One of the aims of the present book is to explain the mathematical foundations of Quantum Mechanics, and Quantum Theories in general, in a mathematically rigorous way. That said, this is a treatise on Mathematics (or Mathematical Physics) rather than a text on Quantum Mechanics. Except for a few cases, the physical phenomenology is left in the background, to privilege the theory's formal and logical aspects. At any rate several examples of the physical formalism are presented, lest one lose touch with the world of physics.

In alternative to, and irrespective of, the physical content, the book should be considered as an introductory text, albeit touching upon rather advanced topics, on functional analysis on Hilbert spaces, including a few elementary yet fundamental results on C^*-algebras. Special attention is given to a series of results in spectral theory, such as the various formulations of the spectral theorem for bounded normal operators and not necessarily bounded, self-adjoint ones. This is, as a matter of fact, one further scope of the text. The mathematical formulation of Quantum Theories is "confined" to Chapters 6, 7, 11, 12, 13 and partly 14. The remaining chapters are logically independent of those, although the motivations for certain mathematical definitions are to be found in Chapters 7, 10, 11, 12, 13 and 14.

A third purpose is to collect in one place a number of rigorous and useful results on the mathematical structure of Quantum Mechanics and Quantum Theories. These are more advanced than what is normally encountered in quantum physics'

[1] "Brothers" I said, "who through a hundred thousand dangers have reached the channel to the west, to the short evening watch which your own senses still must keep, do not choose to deny the experience of what lies past the Sun and of the world yet uninhabited." Dante Alighieri, The Divine Comedy, translated by J. Finn Cotter, edited by C. Franco, Forum Italicum Publishing, New York, 2006.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_1, © Springer-Verlag Italia 2013

manuals. Many of these aspects have been known for a long time but are scattered in the specialistic literature. We should mention *Gleason's theorem*, *the theorem of Kochen and Specker*, the *theorems of Stone–von Neumann* and *Mackey*, *Stone's theorem* and *von Neumann's theorem* about one-parameter unitary groups, *Kadison's theorem*, besides the better known *Wigner*, *Bargmann* and *GNS theorems*; or, more abstract operator theory such as *Fuglede's theorem*, or the *polar decomposition for closed unbounded operators* (which is relevant in the *Tomita-Takesaki theory* and statistical Quantum Mechanics in relationship to the KMS condition); furthermore, self-adjoint properties for symmetric operators, due to Nelson, that descend from the existence of dense sets of analytical vectors, and finally, Kato's work (but not only his) on the essential self-adjointness of certain kinds of operators and their limits from the bottom of the spectrum (mostly based on the *Kato-Rellich theorem*).

Some chapters suffice to cover a good part of the material suitable for advanced courses on Mathematical Methods in Physics; this is common for Master's degrees in Physics or doctoral degrees, if we assume a certain familiarity with notions, results and elementary techniques of measure theory. The text may also be used for a higher-level course in Matematical Physics that includes foundational material on Quantum Mechanics. In the attempt to reach out to Master or Ph.D. students, both in physics with an interest in mathematical methods or in mathematics with an inclination towards physical applications, the author has tried to prepare a self-contained text, as far as possible: hence a primer on general topology and abstract measure theory was included, together with an appendice on differential geometry. Most chapters are accompanied by exercises, many of which solved explicitly.

The book could, finally, be useful to scientists when organising and presenting accurately the profusion of advanced material disseminated in the literature.

At the end of this introductory chapter some results from topology and measure theory are recalled, much needed throughout the whole treatise. The rest of the book is ideally divided into three parts. The first part, up to Chapter 5, regards the general theory of operators on Hilbert spaces, and introduces several fairly general notions, like Banach spaces. Core results are proved, such as the theorems of Baire, Hahn–Banach and Banach–Steinhaus, as well as the fixed-point theorem of Banach-Caccioppoli, the Arzelà–Ascoli theorem and Fredholm's alternative, plus some elementary consequences. In this part basic topological notions are summarised, in the belief that this might benefit physics' students. The latter's training on general topology is at times disparate and often presents gaps, because this subject is, alas, usually taught sporadically in physics' curricula, and not learnt in an organic way like students in mathematics do.

Part two of the book ends in Chapter 10. Beside setting out the quantum formalism, it develops spectral theory in terms of projector-valued measures, up to the spectral decomposition theorems for unbounded self-adjoint operators on Hilbert spaces. This includes the features of maps of operators (functional analysis) for measurable maps that are not necessarily bounded, whose general spectral aspects and domain properties are investigated. A great emphasis is placed on the structure of C^*-algebras and the relative functional calculus, including an elementary study of *Gelfand's transform* and the *commutative Gelfand–Najmark theorem*. The technical results leading to the spectral theorem are stated and proven in a completely abstract manner in Chapter 8, forgetting that the algebras in question are actually operator algebras, and thus showing their broader validity. In Chapter 10 spectral theory is applied to several practical and completely abstract contexts, both quantum and not.

Chapter 6 treats, from a physical perspective, the motivation underlying the theory. The general mathematical formulation of Quantum Mechanics concerns Chapter 7. The mathematical starting point is the idea, going back to von Neumann, that the propositions of physical quantum systems are described by the lattice of orthogonal projectors on a complex Hilbert space. Maximal sets of physically compatible propositions (in the quantum sense) are described by distributive and orthocomplemented, bounded and σ-complete lattices. From this standpoint the quantum definition of an observable in terms of a self-adjoint operator is extremely natural, as is, on the other hand, the formulation of the spectral decomposition theorem. Quantum states are defined as measures on the lattice of *all* orthogonal projectors, which is no longer distributive (due to the presence, in the quantum world, of *incompatible* propositions and observables). Using *Gleason's theorem* states are characterised as positive operators of trace class with unit trace. Pure states (rays in the Hilbert space of the physical system) arise as extreme elements of the convex body of states.

The third part of the book is devoted to formulating axiomatically the mathematical foundations of Quantum Mechanics and investigating more advanced topics like *quantum symmetries* and the *algebraic formulation of quantum theories*. A comprehensive study is reserved to the notions of quantum symmetry and symmetry group (both Wigner's and Kadison's definitions are discussed). Dynamical symmetries and the quantum version of *Nöther's theorem* are covered as well. The *Galilean group* is employed repeatedly, together with its subgroups and central extensions, as reference symmetry group, to explain the theory of projective unitary representations. *Bargmann's theorem* on the existence of unitary representations of simpy connected Lie groups whose Lie algebra obeys a certain cohomology constraint is proved, and Bargmann's *rule of superselection of the mass* is discussed in detail. Then the useful theorems of Gårding and Nelson for projective unit ary representations of Lie groups of symmetries are considered. Important topics are examined that are often neglected in manuals, like the formulation of the uniqueness of unitary representations of the canonical commutation relations (theorems of Stone–von Neumann and Mackey), or the theoretical difficulties in defining time as the conjugate operator to energy (the Hamiltonian). The mathematical hurdles one must overcome in order to make the statement of *Ehrenfest's theorem* precise are briefly treated. Chapter 14 offers an introduction to the ideas and methods of the abstract formulation of observables and algebraic states via C^*-algebras. Here one finds the proof to the *GNS theorem* and some consequences of purely mathematical flavour, like the general *theorem of Gelfand–Najmark*. This closing chapter contains also material on quantum symmetries in an algebraic setting. As example the notion of *C^*-algebra of Weyl* associated to a symplectic space (usually infinite-dimensional) is discussed.

The appendices at the end of the book recap elementary notions about partially ordered sets, group theory and differential geometry.

The author has chosen not to include topics, albeit important, such as the theory of rigged Hilbert spaces (the famous *Gelfand triples*); doing so would have meant adding further preparatory material, in particular regarding the theory of distributions.

1.1.2 Prerequisites

Essential requisites to understand the book's contents (apart from firm backgrounds on linear algebra, plus some group- and representation theory) are the basics of calculus in one and several real variables, measure theory on σ-algebras [Coh80, Rud82] (summarised at the end of the chapter), and a few notions about complex functions. Imperative, on the physics' side, is the acquaintance with undergraduate Physics. More precisely, Analytical Mechanics (the groundwork of Hamilton's formulation of dynamics) and Electromagnetism (the key features of electromagnetic waves and the crucial wavelike phenomena like interference, diffraction, scattering).

Lesser elementary, useful facts will be recalled where needed (examples included) to allow for a solid understanding. One section of Chapter 12 will use the notion of *Lie group* and elemental facts from the corresponding theory. For these we refer to the book's epilogue: the last appendix summarises some useful differential geometry rather thoroughly. More details should be seeked in [War75, NaSt82].

1.1.3 General conventions

1. The symbol $:=$ means "equal, by definition, to".
2. The *inclusion* symbols $\subset, \supset$ allow for equality $=$.
3. The symbol $\bigsqcup$ denotes the disjoint union.
4. $\mathbb{N}$ is the set of natural numbers *including nought*, and $\mathbb{R}_+ := [0, +\infty)$.
5. Unless otherwise stated, the *field of scalars of a* normed/Banach/Hilbert *vector space* is the field of complex numbers $\mathbb{C}$. *Inner product* always means *Hermitian* inner product, unless specified differently.
6. The *complex conjugate* of a number c will be denoted by $\overline{c}$. The same symbol is used for the *closure* of a set of operators; should there be confusion, we will comment on the meaning.
7. The *inner/scalar product* of two vectors ψ, ϕ in a Hilbert space H will be indicated by $(\psi|\phi)$ to distinguish it from the *ordered pair* (ψ, ϕ). The product's *left* entry is antilinear: $(\alpha\psi|\phi) = \overline{\alpha}(\psi|\phi)$.
   If $\psi, \phi \in \mathsf{H}$, the symbols $\psi(\phi|\ )$ and $(\phi|\ )\psi$ denote the *same* linear operator $\mathsf{H} \ni \chi \mapsto (\phi|\chi)\psi$.
8. Complete orthonormal systems in Hilbert spaces will be called *Hilbert bases*, or *bases* for short.
9. The word *operator* tacitly implies *linear* operator, even though this will be often understated.
10. A linear operator $U : \mathsf{H} \to \mathsf{H}'$ between Hilbert spaces H and H' that is isometric and onto will be called *unitary*, even if elsewhere in the literature the name is reserved for the case $\mathsf{H} = \mathsf{H}'$.

11. By *vector subspace* we will mean a subspace *for the vector-space operations*, even in presence of additional structures on the ambient space (e.g. Hilbert, Banach *etc.*).
12. For the Hermitian conjugation we will always use the symbol *. *Hermitian operator*, *symmetric operator*, and *self-adjoint operator* will *not* be considered synonyms.
13. *One-to-one*, 1-1, and injective are synonyms, just like *onto* and surjective. *Bijective* means simultaneously one-to-one and onto. Beware that a *one-to-one correspondence* is a bijective mapping. An *isomorphism*, irrespective of the algebraic structure at stake, is a 1-1 map onto its image, hence a bijective homomorphism.
14. **Boldface** typeset (within a definition or elsewhere) is typically used when *defining* a term for the first time.
15. Corollaries, definitions, examples, lemmas, notations, remarks, propositions, and theorems are all labelled sequentially to simplify their retrieval.
16. At times we will use the shorthand "*iff*", instead of 'if and only if', to say that two statements imply one another, i.e. they are logically equivalent.

1.2 On Quantum Mechanics

1.2.1 Quantum Mechanics as a mathematical theory

From a mathematical point of view Quantum Mechanics (QM) represents a rare blend of mathematical elegance and descriptive insight into the physical world. The theory essentially makes use of techniques of functional analysis mixed with incursions and overlaps with measure theory, probability and mathematical logic.

There are (at least) two possible ways to formulate precisely (i.e. mathematically) elementary QM. The eldest one, historically speaking, is due to von Neumann (1932) in essence, and is formulated using the language of Hilbert spaces and the spectral theory of unbounded operators. A more recent and mature formulation was developed by several authors in the attempt to solve quantum-field-theory problems in mathematical physics; it relies on the theory of abstract algebras (*-algebras and C^*-algebras) built mimicking operator algebras that were defined and studied, again, by von Neumann (nowadays known as *W^*-algebras* or *von Neumann algebras*), but freed from the Hilbert-space structure (for instance, [BrRo02] is a classic on operator algebras). The core result is the celebrated *GNS theorem* (from Gelfand, Najmark and Segal) [Haa96, BrRo02] that we will prove in Chapter 14. The newer formulation can be considered an extension of the former one, in a very precise sense that we shall not go into here, also by virtue of the novel physical context it introduces and by the possibility of treating physical systems with infinitely many degrees of freedom, i.e. quantum fields. In particular, this second formulation makes precise sense of the demand for *locality* and *covariance* of relativistic quantum field theories [Haa96], and allows to extend quantum field theories to curved spacetime.

The algebraic formulation in elementary QM is not strictly necessary, even though it can be achieved and is very elegant (see for example [Str05a] and parts of [DA10]).

Given the relatively basic nature of our book we shall treat almost exclusively the first formulation, which displays an impressive mathematical complexity together with a manifest formal elegance. We will introduce the algebraic formulation in the last chapter only, and stay within the general framework rather than consider QM as a physical object.

A crucial mathematical tool to develop a Hilbert-space formulation for QM is the so-called *spectral theorem for self-adjoint operators* (unbounded, usually) defined on dense subspaces of a Hilbert space. This theorem, which can be extended to normal operators, was first proved by von Neumann in [Neu32] apropos the mathematical structure of QM: this fundamental work ought to be considered a XX century milestone of mathematical physics and pure mathematics. The definition of abstract Hilbert spaces and much of the relative general theory, as we know it today, are also due to von Neumann and his formalisation of QM. Von Neumann built the modern and axiomatic notion of an abstract Hilbert space by considering, in [Neu32, Chapter 1], the two approaches to QM known at that time: the one relying on Heisenberg matrices, and the one using Schrödinger's wavefunctions.

The relationship between QM and *spectral theory* depends upon the following fact. The standard way of interpreting QM dictates that physical quantities that are measurable over quantum systems can be associated to unbounded self-adjoint operators on a suitable Hilbert space. The spectrum of each operator coincides with the collection of values the associated physical quantity can attain. The construction of a physical quantity from easy properties and propositions of the type "the value of the quantity falls in the interval $(a,b]$", which correspond to orthogonal projectors in the adopted mathematical scheme, is nothing else but an integration procedure with respect to an appropriate projector-valued spectral measure. In practice the spectral theorem is just a means to construct complicated operators starting from projectors, or conversely, decompose operators in terms of projector-valued measures.

The contemporary formulation of spectral theory is certainly different from von Neumann's, although the latter already contained all basic constituents. Von Neumann's treatise (dating back 1932) discloses still today an impressive depth, especially in the most difficult sides of the physical interpretation of QM's formalism: by reading that book it becomes clear that von Neumann was well aware of these issues, as opposed to many colleagues of his. It would be interesting to juxtapose his opus to the much more renowned book by Dirac [Dir30] on QM's fundamentals, a comparison that we leave to the interested reader. At any rate, the great interpretative profundity given to QM by von Neumann begins to be recognised by experimental physicists as well, in particular those concerned with quantum measurements [BrKh95].

The so-called *Quantum Logics* arise from the attempt to formalise QM from the most radical position: endowing the same *logic* used to treat quantum systems with properties different from those of ordinary logic, and modifying the semantic theory. For example, more than two truth values are possible, and the Boolean lattice of propositions is replaced by a more complicated non-distributive structure. In the first formulation of quantum logic, known as *Standard Quantum Logic* and introduced by Von Neumann and Birkhoff in 1936, the role of the Boolean algebra of propositions is taken by an *orthomodular lattice*: this is modelled, as a matter of fact, on the set of orthogonal projectors on a Hilbert space, or the collection of closed projection spaces [Bon97], plus some composition rules. Despite its sophistication, that model is known to contain many flaws when one tries to translate it in concrete (*operational*) physical terms. Beside the various formulations of Quantum Logic [Bon97, DCGi02, EGL09], there are also other foundational fomulations based on alternative viewpoints (e.g., *topos* theory).

1.2.2 QM in the panorama of contemporary Physics

Quantum Mechanics, roughly speaking the physical theory of the atomic and subatomic world, and *General and Special Relativiy* (GSR) – the physical theory of gravity, the macroscopic world and cosmology, represent the two paradigms through which the physics of the XX and XXI centuries developed. These two paradigms coalesced, in several contexts, to give rise to relativistic quantum theories. *Relativistic Quantum Field Theory* [StWi00, Wei99], in particular, has witnessed a striking growth and a spectacular predictive and explanatory success in relationship to the theory of elementary particles and fundamental interactions. One example for all: regarding the so-called *standard model of elementary particles*, that theory predicted the unification of the *weak* and *electromagnetic forces* which was confimed experimentally at the end of the '80s during a memorable experiment at the C.E.R.N., in Geneva, where the particles Z_0 and $W^{\pm}$, expected by electro-weak unification, were first observed.

The best-ever accuracy in the measurement of a physical quantity in the whole history of Physics was predicted by *quantum* electrodynamics. The quantity is the so-called gyro-magnetic ratio of the electron g, a dimensionless number. The value expected by quantum electrodynamics for $a := g/2 - 1$ was

$$0.001159652359 \pm 0.000000000282\,,$$

and the experimental result turned out to be

$$0.001159652209 \pm 0.000000000031\,.$$

Many physicists believe QM to be the fundamental theory of the universe (more than relativistic theories), also owing to the impressive range of linear scales where it holds: from $1m$ (Bose-Einstein condensates) to *at least* $10^{-16}m$ (inside nucleons: quarks). QM has had an enormous success, both theoretical and experimental, in materials' science, optics, electronics, with several key technological repercussions: every technological object of common use that is complex enough to contain a *semiconductor* (childrens' toys, mobiles, remote controls ...) exploits the quantum properties of matter.

Going back to the two major approaches of the past century – QM and GSR – there remain a number of obscure points where the paradigms seem to clash; in particular the so-called "quantisation of gravity" and the structure of spacetime at *Planck scales* ($10^{-33}cm$, $10^{-43}s$, the length and time intervals obtained from the fundamental constants of the two theories: the speed of light, the universal constant of gravity and Planck's constant). The necessity of a discontinuous spacetime at ultra-microscopic scales is also reinforced by certain mathematical (and conceptual) hurdles that the so-called theory of quantum *Renormalisation* has yet to overcome, as consequence of the infinite values arising in computing processes due to the interaction of elementary particles. From a purely mathematical perspective the existence of infinite values is actually related to the problem, already intrinsically ambiguous, of defining the product of two distributions: infinites are not the root of the problem, but a mere manifestation of it.

These issues, whether unsolved or partially solved, have underpinned important theoretical advancements of late, which in turn influenced the developments of pure mathematics itself. Examples include (super-)*string* theory, and the various *Noncommutative Geometries*, first of all A. Connes' version and the so-called *Loop Quantum Gravity*. The difficulty in deciding which of these theories makes any physical sense and is apt to describe the universe at very small scales is also practical: today's technology is not capable of preparing experiments that enable to distinguish among all available theories. However, it is relevant to note that recent experimental observations of the so-called γ-bursts, conducted with the telescope "Fermi Gamma-ray", have lowered the threshold for detecting quantum-gravity phenomena (e.g. the violation of Lorentz's symmetry) well below Planck's length[2]. Other discrepancies between QM and GSR, about which the debate is more relaxed today than it was in the past, have to do with QM vs the notions of *locality of relativistic nature* (Einstein-Podolsky-Rosen paradox [Bon97]) in relationship to QM's *entanglement* phenomena. This is due in particular to Bell's study of the late '60s, and to the famous experiments of Aspect that first disproved Einstein's expectations, secondly confirmed the Copenhagen interpretation, and eventually proved that nonlocality is a character istic of Nature, independent of whether one accepts the standard interpretation of QM or not. The vast majority of physicists seems to agree that the existence of nonlocal physical processes, as QM forecasts, does not imply any concrete violation of Relativity's core (quantum entanglement does not involve superluminal transmission of information nor the violation of causality [Bon97]).

In the standard interpretation of QM that is called *of Copenhagen* there are parts that remain physically and mathematically unintelligible, yet still very interesting conceptually. In particular, and despite several appealing attempts, it it still not clear how standard mechanics may be seen as a subcase, or limiting case, of QM, nor how to demarcate (even roughly, or temporarily) the two worlds. Further, the question remains of the physical and mathematical description of the so-called *process of quantum measurement*, of which more later, that is strictly related to the classical limit of QM. From this fact, as well, other interpretations of the QM formalisms were born that differ deeply from Copenhagen's interpretation. Among these more recent interpretations, once considered heresies, Bohm's interpretation relies on *hidden variables* [Bon97, Des99] and is particularly intriguing.

[2] Abdo A.A. et al.: A limit on the variation of the speed of light arising from quantum gravity effects. Nature **462**, 331–334 (2009).

Doubts are sometimes raised about the formulation of QM and on it being not truly clear, but just a list of procedures that "actually work", whereas its true nature is something inaccessible, sort of "noetic". In the author's opinion a dangerous epistemological mistake hides behind this point of view. The misconception is based on the belief that "explaining" a phenomenon means reducing it to the categories of daily life, as if everyday experience reached farther than reality itself. Quite the contrary: those categories were built upon conventional wisdom, and hence without any alleged metaphysical insight. Behind that simple "actually works" a deep philosophical landscape could unfold and draw us closer to reality rather than pushing us farther away. Quantum Mechanics taught us to think in a different fashion, and for this reason it has been (is, actually) an incredible opportunity for the human enterprise. To turn our backs to QM and declare we do not understand it because it refuses to befit our familiar mental categories means to lock the door that separates us from something huge. This is the author's stance, who does indeed consider *Heisenberg's uncertainty principle* (a theorem in today's formulation, despite the name) one of the highest achievements of the human being.

Mathematics is the most accurate of languages invented by man. It allows to create formal structures corresponding the possible worlds that may or not exist. The plausibility of these hypothetical realities is found solely in the logical or syntactical coherence of the corresponding mathematical structure. In this way semantic "chimaeras" might arise, that turn out to be syntactically coherent though. These creatures are sometimes consistent with worlds or states that do exist, albeit not yet discovered. A feature that is attributable to an existing entity can only either be present or not, according to the classical ontological view. Quantum Mechanics, in particular, leads to say that any such property may not just obey the true/false pattern, but also be "uncertain", despite being inherent to the object itself. This tremendous philosophical leap can be entirely managed within the mathematical foundations of QM, and represents the most profound philosophical legacy of Heisenberg's principle.

There remain open at least two general issues, of gnoseological nature essentially, common to the entire formulation of modern science. The first is the relationship between theoretical entities and the objects we have experience of. The problem is particularly delicate in QM, where the notion of measuring instrument has not of yet been fully clarified. Generally speaking, the relationship of a theoretical entity with an experimental object is not direct, and still based on often understated theoretical assumptions. But this is the case in classical theories as well, when one for example tackles problems such as the geometry of the physical space: there, it is necessary to identify, inside the physical reality, objects that correspond to the idea of a point, a segment, and so on, and to do that we use other assumptions, like the fact that the geometry of the straightedge is the same obtained when inspecting space with light beams. The second issue is the hopelessness of trying to prove the syntactic coherence of a mathematical construction. We may attempt to reduce the latter to the coherence of set theory, or category theory; that this reduction should prove the construction's solidity has more to do with psychology than of it being a real fact, due to the profusion of well-known paradoxes disseminated along the history of the formalisation of mathematics, and eventually to Gödel's famous theorem.

1.3 Backgrounds on general topology

For the reader's sake we sum up here notions of point-set topology that will be used by and large in the book. All statements are elementary and classical, and can be easily found in any university treatise, so for brevity we will prove almost nothing. The practiced reader may skip this section completely and return to it at subsequent stages for reference.

1.3.1 Open/closed sets and basic point-set topology

The notions of *open set* and *closed set* are defined as follows [Ser94II], in the greatest generality.

Definition 1.1. *The pair* $(\mathsf{X}, \mathscr{T})$*, where* X *is a set and* $\mathscr{T}$ *a collection of subsets of* X*, is said a* **topological space** *if:*

(i) $\varnothing, \mathsf{X} \in \mathscr{T}$;
(ii) *the union of (arbitrarily many) elements of* $\mathscr{T}$ *is an element of* $\mathscr{T}$;
(iii) *the intersection of a finite number of elements of* $\mathscr{T}$ *belongs to* $\mathscr{T}$.

$\mathscr{T}$ *is called a* **topology** *on* X *and the elements of* $\mathscr{T}$ *are the* **open sets** *of* X.

Definition 1.2. *On a topological space* $(\mathsf{X}, \mathscr{T})$*:*

(a) *A* **basis** *for the topology of* $(\mathsf{X}, \mathscr{T})$ *is a subset* $\mathscr{B} \subset \mathscr{T}$ *such that each element in* $\mathscr{T}$ *is the union of elements of* $\mathscr{B}$.
(b) *An* **open neighbourhood** *of* $p \in \mathsf{X}$ *is an element* $A \in \mathscr{T}$ *such that* $p \in A$.
(c) $x \in S \subset \mathsf{X}$ *is an* **interior point** *of* S *if there exists an open neighbourhood* A *of* x *contained in* S.
The **interior** *of a set* $S \subset \mathsf{X}$ *is the set:*

$$Int(S) := \{x \in \mathsf{X} \mid x \text{ is an interior point of } S\}\,.$$

The **exterior** *of a set* $S \subset \mathsf{X}$ *is the set:*

$$Ext(S) := \{x \in \mathsf{X} \mid x \text{ is an interior point of } \mathsf{X} \setminus S\}\,.$$

The **frontier** *of a set* $S \subset \mathsf{X}$ *is the difference set:*

$$\partial S := \mathsf{X} \setminus (Int(S) \cup Ext(S))\,.$$

(d) $C \subset \mathsf{X}$ *is called* **closed** *if* $\mathsf{X} \setminus C$ *is open.*

A subset $S \subset \mathsf{X}$ in a topological space $(\mathsf{X}, \mathscr{T})$ inherits the structure of a topological space from $(\mathsf{X}, \mathscr{T})$ by defining a topology on S as $\mathscr{T}_S := \{S \cap A \mid A \in \mathscr{T}\}$. This topology (the definition is easily satisfied) is called the **induced topology** on S by $(\mathsf{X}, \mathscr{T})$.

Most of the topological spaces we will see in this text are *Hausdorff spaces*, on which open sets "separate" points.

Definition 1.3. *A topological space* $(\mathsf{X}, \mathscr{T})$ *and its topology are called* **Hausdorff** *if they satisfy the* **Hausdorff property***: for every* $x, x' \in \mathsf{X}$ *there exist* $A, A' \in \mathscr{T}$*, with* $x \in A$*,* $x' \in A'$*, such that* $A \cap A' = \varnothing$.

Remark 1.4. **(1)** Both X and $\varnothing$ are open and closed sets.
(2) Closed sets satisfy properties that are "dual" to open sets, as follows straightforwardly from their definition. Hence:

(i) $\varnothing, \mathsf{X}$ are closed;
(ii) the intersection of (infinitely many) closed sets is closed;
(iii) the finite union of closed sets is a closed set.

(3) The simplest example of Hausdorff topology is given by the collection of subsets of $\mathbb{R}$ containing the empty set and arbitrary unions of open intervals. This is thus a *basis* for the topology in the sense of Definition 1.1. It is called **Euclidean topology** or **standard topology** of $\mathbb{R}$ or $\mathbb{C}$.
(4) A slightly more complicated example of Hausforff topology is the **Euclidean topology**, or **standard topology**, of $\mathbb{R}^n$, $\mathbb{C}^n$. It is the usual topology one refers to in elementary calculus, and is built as follows. If $\mathbb{K} := \mathbb{R}$ or $\mathbb{C}$, the **standard norm** of $(c_1, \ldots, c_n) \in \mathbb{K}^n$ is, by definition:

$$||(c_1, \ldots, c_n)|| := \sqrt{\sum_{k=1}^{n} |c_k|^2}, \quad (c_1, \ldots, c_n) \in \mathbb{K}^n . \tag{1.1}$$

The set:

$$B_\delta(x_0) := \{x \in \mathbb{K}^n \,|\, ||x|| < \delta\} \tag{1.2}$$

is, hence, the usual **open ball** of $\mathbb{K}^n$ of radius $\delta > 0$ and centre $x_0 \in \mathbb{K}^n$. The open sets in the standard topology of $\mathbb{K}^n$ are, empty set aside, the unions of open balls of any given radius and centre. These balls constitute a basis for the standard topology of $\mathbb{R}^n$ and $\mathbb{C}^n$. ■

Here are notions that will come up often in the sequel.

Definition 1.5. *If $(\mathsf{X}, \mathscr{T})$ is a topological space, the* **closure** *of $S \subset \mathsf{X}$ is the set:*

$$\overline{S} := \cap\{C \supset S, C \subset \mathsf{X} \,|\, C \text{ is closed}\}.$$

S is called **dense** *in X if $\overline{S} = \mathsf{X}$.*
$(\mathsf{X}, \mathscr{T})$ is said to be **separable** *if there exists a dense and countable subset $S \subset \mathsf{X}$.*

From the definition follow these properties.

Proposition 1.6. *If $(\mathsf{X}, \mathscr{T})$ is a topological space and $S \subset \mathsf{X}$:*

(a) *$\overline{S}$ is closed;*
(b) *$\overline{\overline{S}} = \overline{S}$;*
(c) *if $T \subset \mathsf{X}$, then $S \subset T$ implies $\overline{S} \subset \overline{T}$;*
(d) *S is closed if and only if $\overline{S} = S$.*

Definition 1.7. *A topological space $(\mathsf{X}, \mathscr{T})$ has a* **countable basis**, *or is* **second-countable**, *if there is a countable subset $\mathscr{T}_0 \subset \mathscr{T}$ (the "countable basis") such that every open set is the union of elements of $\mathscr{T}_0$.*

If $(\mathsf{X}, \mathscr{T})$ has a countable basis then *Lindelöf's lemma* holds:

Theorem 1.8 (Lindelöf's lemma). *Let $(\mathsf{X}, \mathscr{T})$ be a second-countable topological space. Then any open covering of a given subset in X admits a countable sub-covering: if $B \subset \mathsf{X}$ and $\{A_i\}_{i\in I} \subset \mathscr{T}$ with $\cup_{i\in I} A_i \supset B$, then $\cup_{i\in J} A_i \supset B$ for some countable $J \subset I$.*

Remarks 1.9. $\mathbb{R}^n$ and $\mathbb{C}^n$, equipped with the standard topology, are second-countable: for $\mathbb{R}^n$, $\mathscr{T}_0$ can be taken to be the collection of open balls with rational radii and centred at rational points. The generalisation to $\mathbb{C}^n$ is obvious. ■

In conclusion, we recall the definition of product topology.

Definition 1.10. *If $\{(\mathsf{X}_i, \mathscr{T}_i)\}_{i\in F}$ is a collection of topological spaces indexed by a finite set F, the* **product topology** *on $\times_{i\in F}\mathsf{X}_i$ is the topology whose open sets are $\varnothing$ and the unions of Cartesian products $\times_{i\in F} A_i$, with $A_i \in \mathscr{T}_i$ for any $i \in F$.*

If F has arbitrary cardinality, the previous definition cannot be generalised directly. If we did so in the obvious way we would not maintain important properties, such as Tychonoff's theorem, that we will discuss later. Nevertheless, a natural topology on $\times_{i\in F}\mathsf{X}_i$ can be defined, still called **product topology** because is extends Definition 1.10.

Definition 1.11. *If $\{(\mathsf{X}_i, \mathscr{T}_i)\}_{i\in F}$ is a collection of topological spaces with F of arbitrary cardinality, the* **product topology** *on $\times_{i\in F}\mathsf{X}_i$ has as open sets $\varnothing$ and the unions of Cartesian products $\times_{i\in F} A_i$, with $A_i \in \mathscr{T}_i$ for any $i \in F$, such that on each $\times_{i\in F} A_i$ we have $A_i = \mathsf{X}_i$ but for a finite number of indices i.*

Remark 1.12. (1) The standard topology of $\mathbb{R}^n$ is the product topology obtained by endowing the single factors $\mathbb{R}$ with the standard topology. The same happens for $\mathbb{C}^n$.
(2) Either in case of finitely many, or infinitely many, factors, the canonical projections:

$$\pi_i : \times_{j\in F}\mathsf{X}_j \ni \{x_j\} \mapsto x_i \in \mathsf{X}_i$$

are clearly continuous if we have the product topology on the domain. It can be proved that the product topology is the coarsest among all topologies making the canonical projections continuous (coarsest means it is included in any such topology). ■

1.3.2 Convergence and continuity

Let us pass to convergence and continuity. First of all we need to recall the notions of *convergence of a sequence* and *limit point*.

Definition 1.13. *Let $(\mathsf{X}, \mathscr{T})$ be a topological space.*

(a) *A sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$* **converges** *to a point $x \in \mathsf{X}$, called the* **limit of the sequence***:*

$$x = \lim_{n\to+\infty} x_n \quad \text{and also} \quad x_n \to x \quad \text{as } n \to +\infty$$

if, for any open neighbourhood A of x there exists an $N_A \in \mathbb{R}$ such that $x_n \in A$ whenever $n > N_A$.

(b) *$x \in \mathsf{X}$ is a* **limit point** *of a subset $S \subset \mathsf{X}$ if any neighbourhood A of x contains a point of $S \setminus \{x\}$.*

Remarks 1.14. It should be patent from the definitions that in a Hausdorff space the limit of a sequence is unique, if it exists. ■

The relationship between limit points and closure of a set is sanctioned by the following classical and elementary result:

Proposition 1.15. *Let $(\mathsf{X}, \mathscr{T})$ be a topological space and $S \subset \mathsf{X}$.*
$\overline{S}$ coincides with the union of S and the set of its limit points.

The definition of continuous map and continuous map at one point is recalled below.

Definition 1.16. *Let $f : \mathsf{X} \to \mathsf{X}'$ be a function between topological spaces $(\mathsf{X}, \mathscr{T})$, $(\mathsf{X}', \mathscr{T}')$.*
(a) *f is called* **continuous** *if $f^{-1}(A') \subset \mathscr{T}$ for any $A' \in \mathscr{T}'$.*
(b) *f is said* **continuous at** *$p \in \mathsf{X}$ if, for any open neighbourhood $A'_{f(p)}$ of $f(p)$, there is an open neighbourhood A_p of p such that $f(A_p) \subset A'_{f(p)}$.*
(c) *f is called a* **homeomorphism** *if:*

(i) *f is continuous;*
(ii) *f is bijective;*
(iii) *$f^{-1} : \mathsf{X}' \to \mathsf{X}$ is continuous.*

In this case X and X' are said to be **homeomorphic** *(under f).*

Remark 1.17. (1) It is easy to check that $f : \mathsf{X} \to \mathsf{X}'$ is continuous if and only if it is continuous at every point $p \in \mathsf{X}$.
(2) The notion of continuity at p as of (b) reduces to the more familiar "ε-δ" definition when the spaces X and X' are $\mathbb{R}^n$ (or $\mathbb{C}^n$) with the standard topology; to see this bear in mind that: (a) open neighbourhood can always be chosen to be open balls of radii δ and ε, centred at p and $f(p)$ respectively; (b) every open neighbourhood contains an open ball centred at that point. ■

Let us mention a useful result concerning the standard real line $\mathbb{R}$. One defines the **limit supremum** (also **superior limit**, or simply **limsup**) and the **limit infimum** (**inferior limit** or just **liminf**) of a sequence $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$ as follows:

$$\limsup_{n\in\mathbb{N}} s_n := \inf_{k\in\mathbb{N}} \sup_{n\geq k} s_k \left(= \lim_{k\to+\infty} \sup_{n\geq k} s_k \right), \quad \liminf_{n\in\mathbb{N}} s_n := \sup_{k\in\mathbb{N}} \inf_{n\geq k} s_k \left(= \lim_{k\to+\infty} \inf_{n\geq k} s_k \right).$$

Note how these numbers exist for any given sequence $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, possibly being infinite, as they arise as limits of monotone sequences, whereas the limit of $\{s_n\}_{n\in\mathbb{N}}$ might not exist (neither finite nor infinite). However, it is not hard to prove the following elementary fact.

Proposition 1.18. *If* $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, *then* $\lim_{n\to+\infty} s_n$ *exists, possibly infinite, if and only if*

$$\limsup_{n\in\mathbb{N}} s_n = \liminf_{n\in\mathbb{N}} s_n \,.$$

In such a case:

$$\lim_{n\to+\infty} s_n = \limsup_{n\in\mathbb{N}} s_n = \liminf_{n\in\mathbb{N}} s_n \,.$$

1.3.3 Compactness

Let us briefly recall some easy facts about *compact* sets.

Definition 1.19. *Let* $(\mathsf{X}, \mathscr{T})$ *be a topological space and* $K \subset \mathsf{X}$.

(a) *K is called* **compact** *if any open covering of it admits a finite sub-covering: if* $\{A_i\}_{i\in I} \subset \mathscr{T}$ *with* $\cup_{i\in I} A_i \supset K$ *then* $\cup_{i\in J} A_i \supset K$ *for some finite* $J \subset I$.
(b) *K is said* **relatively compact** *if* $\overline{K}$ *is compact.*
(c) X *is* **locally compact** *if any point in* X *has a relatively compact open neighbourhood.*

Compact sets satisfy a host of properties [Ser94II] and we will not be concerned with them much more (though returning to them in Chapter 4). Let us recall a few results at any rate.

Let us begin with the relationship with calculus on $\mathbb{R}^n$. If X is $\mathbb{R}^n$ (or $\mathbb{C}^n$ identified with $\mathbb{R}^{2n}$), the celebrated *Heine–Borel theorem* holds [Ser94II].

Theorem 1.20 (Heine–Borel). *If* $\mathbb{R}^n$ *is equipped with the standard topology,* $K \subset \mathbb{R}^n$ *is compact if and only if K is simultaneously closed and bounded (meaning* $K \subset B_\delta(x)$ *for some* $x \in \mathbb{R}^n$, $\delta > 0$*).*

In calculus, the *Weierstrass theorem*, which deals with continuous maps defined on compact subsets of $\mathbb{R}^n$ (or $\mathbb{C}^n$), can be proved directly without the definition of compactness. Actually one can prove a more general statement on $\mathbb{R}^n$-valued ($\mathbb{C}^n$-valued) continuous maps defined on compact subsets.

Proposition 1.21. *If* $\mathbb{K} = \mathbb{C}$ *or* $\mathbb{R}$, *let* $|| \; ||$ *denote the standard norm of* $\mathbb{K}^n$ *as in (1.1), and endow* $\mathbb{K}^n$ *with the standard topology.*
If $f : K \to \mathbb{K}^n$ *is continuous on the compact subset K of a topological space, then it is bounded, i.e. there is an* $M \in \mathbb{R}$ *such that* $||f(x)|| \le M$ *for any* $x \in K$.

Proof. Since f is continuous at any point $p \in K$, we have $||f(x)|| \le M_p \in \mathbb{R}$ for all $x \in A_p$ open neighbourhood of p. As K is compact, we may extract a finite sub-covering $\{A_{p_k}\}_{k=1,\ldots,N}$ from $\{A_p\}_{p\in K}$ that covers K. The number $M := \max_{k=1,\ldots,N} M_k$ satisfies the request. □

Remark 1.22. **(1)** It is easily proved that if X is a Hausdorff space and $K \subset \mathsf{X}$ is compact then K is closed.
(2) Similarly, if K is compact in X, then every closed subset $K' \subset K$ is compact.
(3) Continuous functions map compact sets to compact sets.

(4) By definition of compatness and of induced topology it is clear that a set $K \subset Y$, with the induced topology on $Y \subset \mathsf{X}$, is compact in Y if and only if K is compact in X. ■

The properties of being compact and Hausdorff bear an interesting relationship. One such property is expressed by the following useful statement.

Proposition 1.23. *Let $f : M \to N$ be a continuous map from the compact space M to the compact Hausdorff space N. If f is bijective then it is a homeomorphism.*

On locally compact Hausdorff spaces an important technical result, known as **Urysohn's lemma**, holds. To state it, we first need to define the **support of a map** $f : \mathsf{X} \to \mathbb{C}$:

$$supp(f) := \overline{\{x \in \mathsf{X} \,|\, f(x) = 0\}}\,,$$

where the bar is the topological closure in the space X.

Theorem 1.24 (Urysohn's lemma). *If $(\mathsf{X}, \mathscr{T})$ is a Hausdorff, locally compact space, for any compact $K \subset \mathsf{X}$ and any open set $U \supset K$ there exists a continuous map $f : \mathsf{X} \to [0,1]$ such that:*

(i) *$supp(f) \subset U$;*
(ii) *$supp(f)$ is compact;*
(iii) *$f(x) = 1$ if $x \in K$.*

Eventually, the following key theorem relates the product topology to compactness.

Theorem 1.25 (Tychonoff). *The Cartesian product of (arbitrarily many) compact spaces is compact in the product topology.*

1.3.4 Connectedness

Definition 1.26. *A topological space* X *is said to be* **connected** *if it cannot be written as the union of two disjoint open sets. A subset $A \subset \mathsf{X}$ is* **connected** *if it is connected in the induced topology.*

By defining the equivalence relation:

$$x \sim x' \text{ iff } x, x' \in C, \text{ where } C \text{ is a connected set in } \mathsf{X},$$

the resulting equivalence classes are maximal connected subsets in X called the **connected components** of X. Consequently, the connected components of X are disjoint and cover X. Connected components are clearly closed.

Definition 1.27. *A subset A in a topological space X is* **path-connected** *if for any pair of points $p, q \in A$ there is a continuous map (a path) $\gamma : [0,1] \to A$ such that $\gamma(0) = p$, $\gamma(1) = q$.*

Definition 1.28. *A subset A in a topological space* X *is called* **simply connected** *if, for any $p,q \in A$ and any (continuous) paths $\gamma_i : [0,1] \to A$, $i = 0,1$, such that $\gamma_i(0) = p$, $\gamma_i(1) = q$, there exists a continuous map $\gamma : [0,1] \times [0,1] \to A$, called a* **homotopy**, *satisfying $\gamma(s,0) = p$, $\gamma(s,1) = q$ for all $s \in [0,1]$ and $\gamma(0,t) = \gamma_0(t)$, $\gamma(1,t) = \gamma_1(t)$ for all $t \in [0;1]$.*

Remark 1.29. (1) Directly from the definition we have that continuous functions map connected spaces to connected spaces and path-connected spaces to path-connected spaces.
(2) A path-connected space is connected, but not conversely in general. A non-empty, open connected subset of $\mathbb{R}^n$ is always path-connected. This is a general property that holds in **locally path-connected** spaces, in which each point has a path-connected open neighbourhood.
(3) It can be proved that the product of two simply connected spaces, if equipped with the product topology, is simply connected.
(4) There is an equivalent definition of simply connected space, based on the important notion of homotopy group [Ser94II]. We shall not make use of that notion in this book. ■

1.4 Round-up on measure theory

This section contains, for the reader's sake, basic notions and elementary results on abstract measure theory, plus fundamental facts from real analysis on Lebesgue's measure on the real line. To keep the treatise short we will not prove any statement, for these can be found in the classics [Hal69, Coh80, Rud82]. Well-read users might want to skip this part entirely, and refer to it for explanations on conventions or notations used throughout.

1.4.1 Measure spaces

The modern theory of integration is rooted in the notion of σ-*algebra* of sets: this is a collection $\Sigma(\mathsf{X})$ of subsets of a given 'universe' set X that can be "measured" by an arbitrary "measuring" function μ that we will fix later. The definition of a σ-algebra specifies which are the good properties that subsets should possess in relationship to the operations of union and intersection. The "σ" in the name points to the closure property (property 1.30(c)) of $\Sigma(\mathsf{X})$ under countable unions. The *integral of a function* defined on X with respect to a measure μ on the σ-algebra is built step by step.

We begin by defining σ-algebras, and a weaker version (*algebras of sets*) where unions are allowed only finite cardinality, which has an interest of its own.

Definition 1.30. *A* **σ-algebra** *over the set* X *is a collection* $\Sigma(\mathsf{X})$ *of subsets of* X *such that:*

(a) $\mathsf{X} \in \Sigma(\mathsf{X})$.
(b) $E \in \Sigma(\mathsf{X})$ *implies* $\mathsf{X} \setminus E \in \Sigma(\mathsf{X})$.

(c) *if* $\{E_k\}_{k\in\mathbb{N}} \subset \Sigma(\mathsf{X})$ *then* $\bigcup_{k\in\mathbb{N}} E_k \in \Sigma(\mathsf{X})$.

A **measurable space** *is a pair* $(\mathsf{X}, \Sigma(\mathsf{X}))$*, where* X *is a set and* $\Sigma(\mathsf{X})$ *a* σ*-algebra on* X.
A collection $\Sigma_0(\mathsf{X})$ *of subsets* X *is called an* **algebra** *(of sets) on* X *in case (a), (b) hold (replacing* $\Sigma(\mathsf{X})$ *by* $\Sigma_0(\mathsf{X})$*), and (c) is weakened to:*

(c)' *if* $\{E_k\}_{k\in F} \subset \Sigma_0(\mathsf{X})$*, with F finite, then* $\bigcup_{k\in F} E_k \in \Sigma_0(\mathsf{X})$.

Remark 1.31. (1) From (a) and (b) follows $\varnothing \in \Sigma(\mathsf{X})$. (c) includes *finite* unions in $\Sigma(\mathsf{X})$: a σ-algebra is an algebra of sets. This is a consequence of (c) if one takes finitely many distinct E_k. (b) and (c) imply $\Sigma(\mathsf{X})$ is also closed under countable intersections (at most).
(2) By definition of σ-algebra it follows that the intersection of σ-algebras on X is a σ-algebra on X. Moreover, the collection of all subsets of X is a σ-algebra on X. ■

Remark (2) prompts us to introduce a relevant technical notion. If $\mathscr{A}$ is a collection of subsets in X, there always is at least one σ-algebra containing all elements of $\mathscr{A}$. Since the intersection of all σ-algebras on X containing $\mathscr{A}$ is still a σ-algebra, the latter is well defined and called σ-algebra **generated by** $\mathscr{A}$.

Now let us define a notion, crucial for our purposes, where topology and measure theory merge.

Definition 1.32. *If* X *is a topological space with topology* $\mathscr{T}$*, the* σ*-algebra on* X *generated by* $\mathscr{T}$*, denoted* $\mathscr{B}(\mathsf{X})$*, is said* **Borel σ-algebra** *on* X.

Remark 1.33. (1) The notation $\mathscr{B}(\mathsf{X})$ is slightly ambiguous since $\mathscr{T}$ does not appear. We shall use that symbol anyway, unless confusion arises.
(2) If X coincides with $\mathbb{R}$ or $\mathbb{C}$ we shall assume in the sequel that $\Sigma(\mathsf{X})$ is the Borel σ-algebra $\mathscr{B}(\mathsf{X})$ determined by the standard topology on X (that of $\mathbb{R}^2$ if we are talking of $\mathbb{C}$).
(3) By definition of σ-algebra it follows immediately that $\mathscr{B}(\mathsf{X})$ contains in particular open and closed subsets, intersections of (at most countably many) open sets and unions of (at most countably many) closed sets. ■

The mathematical concept we are about to present is that of a *measurable* function. We can say, in a manner of speaking, that this notion corresponds to that of a continuous function in topology.

Definition 1.34. *Let* $(\mathsf{X}, \Sigma(\mathsf{X}))$*,* $(\mathsf{Y}, \Sigma(\mathsf{Y}))$ *be measurable spaces. A function* $f : \mathsf{X} \to \mathsf{Y}$ *is said to be* **measurable** *(with respect to the two* σ*-algebras) whenever* $f^{-1}(E) \in \Sigma(\mathsf{X})$ *for any* $E \in \Sigma(\mathsf{Y})$*. In particular, if we take* $\Sigma(\mathsf{X}) = \mathscr{B}(\mathsf{X})$*, and* $Y = \mathbb{R}$ *or* $\mathbb{C}$*, measurable functions from* X *to* Y *are called* **(Borel) measurable functions***, respectively real or complex.*

Remarks 1.35. Let X and Y be topological spaces with topologies $\mathscr{T}(\mathsf{X})$ and $\mathscr{T}(\mathsf{Y})$. It is easily proved that an $f : \mathsf{X} \to \mathsf{Y}$ is measurable with respect to the Borel σ-algebras $\mathscr{B}(\mathsf{X})$, $\mathscr{B}(\mathsf{Y})$ if and only if $f^{-1}(E) \in \mathscr{B}(\mathsf{X})$ for any $E \in \mathscr{T}(\mathsf{Y})$. Immediately, then, *every continuous map* $f : \mathsf{X} \to \mathbb{C}$ *or* $f : \mathsf{X} \to \mathbb{R}$ *is Borel measurable.* ■

Let us summarise the main features of measurable maps from X to $\mathsf{Y} = \mathbb{R}$, $\mathbb{C}$.

Proposition 1.36. *Let* $(\mathsf{X}, \Sigma(\mathsf{X}))$ *be a measurable space and* $M_{\mathbb{R}}(\mathsf{X})$, $M(\mathsf{X})$ *the sets of measurable maps from* X *to* $\mathbb{R}$, $\mathbb{C}$ *respectively. The following results hold.*

(a) $M_{\mathbb{R}}(\mathsf{X})$ *and* $M(\mathsf{X})$ *are vector spaces, respectively real and complex, with respect to pointwise linear combinations*

$$(\alpha f + \beta g)(x) := \alpha f(x) + \beta g(x), \quad x \in \mathsf{X},$$

for any measurable maps f, g *from* X *to* $\mathbb{R}$, $\mathbb{C}$ *and any real or complex* α, β.
(b) *If* $f, g \in M_{\mathbb{R}}(\mathsf{X})$, $M(\mathsf{X})$ *then* $f \cdot g \in M_{\mathbb{R}}(\mathsf{X})$, $M(\mathsf{X})$, *with* $(f \cdot g)(x) := f(x)g(x)$ *for all* $x \in \mathsf{X}$.
(c) *The following facts are equivalent:*

(i) $f \in M(\mathsf{X})$;
(ii) $\overline{f} \in M(\mathsf{X})$;
(iii) $Re\, f, Im\, f \in M_{\mathbb{R}}(\mathsf{X})$,

where $\overline{f}(x) := \overline{f(x)}$, $(Re\, f)(x) := Re(f(x))$, *and* $(Im\, f)(x) := Im(f(x))$, *for all* $x \in \mathsf{X}$.
(d) *If* $f \in M_{\mathbb{R}}(\mathsf{X})$ *or* $f \in M(\mathsf{X})$ *then* $|f| \in M_{\mathbb{R}}(\mathsf{X})$, *where* $|f|(x) := |f(x)|$, $x \in \mathsf{X}$.
(e) *If* $f_n \in M(\mathsf{X})$, *or* $f_n \in M_{\mathbb{R}}(\mathsf{X})$, *for any* $n \in \mathbb{N}$ *and* $f_n(x) \to f(x)$ *for all* $x \in \mathsf{X}$ *as* $n \to +\infty$, *then* $f \in M(\mathsf{X})$, *or* $f \in M_{\mathbb{R}}(\mathsf{X})$.
(f) *If* $f_n \in M_{\mathbb{R}}(\mathsf{X})$ *and* $\sup_{n\in\mathbb{N}} f(x)$ *is finite for any* $x \in \mathsf{X}$, *then the function* $\mathsf{X} \ni x \mapsto \sup_{n\in\mathbb{N}} f(x)$ *belongs to* $M_{\mathbb{R}}(\mathsf{X})$.
(g) *If* $f_n \in M_{\mathbb{R}}(\mathsf{X})$ *and* $\limsup_{n\in\mathbb{N}} f(x)$ *is finite for all* $x \in \mathsf{X}$, *the function* $\mathsf{X} \ni x \mapsto \limsup_{n\in\mathbb{N}} f(x)$ *is an element of* $M_{\mathbb{R}}(\mathsf{X})$.
(h) *If* $f, g \in M_{\mathbb{R}}(\mathsf{X})$ *the map* $\mathsf{X} \ni x \mapsto \sup\{f(x), g(x)\}$ *is in* $M_{\mathbb{R}}(\mathsf{X})$.
(i) *If* $f \in M_{\mathbb{R}}(\mathsf{X})$ *and* $f \geq 0$, *then the map* $\mathsf{X} \ni x \mapsto \sqrt{f(x)}$ *is in* $M_{\mathbb{R}}(\mathsf{X})$.

From now on, as is customary in measure theory, we will work with the **extended real line**:

$$[-\infty, \infty] := \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}.$$

Here $\mathbb{R}$ is widened by adding the symbols $\pm\infty$; the ordering of the reals is extended by declaring $-\infty < r < +\infty$ for any $r \in \mathbb{R}$ and defining on $\overline{\mathbb{R}}$ the topology whose basis consists of real open interval and the sets (the notation should be obvious) $[-\infty, a)$, $(a, +\infty]$ for any $a \in \mathbb{R}$. Moreover one defines: $|-\infty| := |+\infty| =: +\infty$.

Now a standard result.

Proposition 1.37. *If* $(\mathsf{X}, \Sigma(\mathsf{X}))$ *is a measurable space,* $f : \mathsf{X} \to \overline{\mathbb{R}}$ *is measurable if and only if* $f^{-1}((a, +\infty]) \in \Sigma(\mathsf{X})$ *for any* $a \in \mathbb{R}$. *Furthermore, statements (d), (e), (f), (g), (h) of Proposition 1.36 still hold when* f_n *and* f *are* $\overline{\mathbb{R}}$*-valued, with the proviso that one drops finiteness in (f) and (g).*

Remark 1.38. **(1)** In (f), (g) and (h) of propostion 1.36 we may substitute inf to sup and obtain valid statements.
(2) As far as the first statement of 1.37, the analogous statements with $(a, +\infty]$ replaced by $[a, +\infty]$, $[-\infty, a)$, or $[-\infty, a]$ hold. ■

1.4.2 Positive σ-additive measures

We pass to define σ*-additive, positive measures.*

Definition 1.39. *If* $(\mathsf{X}, \Sigma(\mathsf{X}))$ *is a measurable space, a* **(σ-additive) positive measure** *on X (with respect to* $\Sigma(\mathsf{X})$*), is a function* $\mu : \Sigma(\mathsf{X}) \to [0, +\infty]$ *satisfying:*

(a) $\mu(\varnothing) = 0$*;*
(b) $\mu\left(\bigcup_{n\in\mathbb{N}} E_n\right) = \sum_{n\in\mathbb{N}} \mu(E_n)$ *if* $\{E_n\}_{n\in\mathbb{N}} \subset \Sigma(\mathsf{X})$*, and* $E_n \cap E_m = \varnothing$ *if* $n \neq m$ *(***σ-additivity***).*
The triple $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ *is called a* **measure space**.

Remark 1.40. (1) The series in (b), having non-negative terms, is well defined and can be rearranged at will.
(2) Easy consequences of the definition are the following properties.
Monotonicity: if $E \subset F$ with $E, F \in \Sigma(\mathsf{X})$,

$$\mu(E) \leq \mu(F).$$

Sub-additivity: if $\{E_n\}_{n\in\mathbb{N}} \subset \Sigma(\mathsf{X})$:

$$\mu\left(\cup_{n\in\mathbb{N}} E_n\right) \leq \sum_{n\in\mathbb{N}} \mu(E_n).$$

Inner continuity: if $E_1 \subset E_2 \subset E_3 \subset \cdots$ for $E_n \in \Sigma(\mathsf{X})$, $n = 1, 2, \ldots$, then:

$$\mu\left(\cup_{n=1}^{+\infty} E_n\right) = \lim_{n\to+\infty} \mu(E_n).$$

Outer continuity: if $E_1 \supset E_2 \supset E_3 \supset \cdots$ for $E_n \in \Sigma(\mathsf{X})$, $n = 1, 2, \ldots$, and $\mu(E_m) < +\infty$ for some m, then:

$$\mu\left(\cap_{n=1}^{+\infty} E_n\right) = \lim_{n\to+\infty} \mu(E_n).$$ ∎

Measures on σ-algebras can be constructed using extension techniques, by starting with measures on algebras (hence not closed under countable unions). We will employ such recipes later in the text. An important extension theorem for measures [Hal69] goes like this.

Theorem 1.41. *Let* $\Sigma_0(\mathsf{X})$ *be an algebra of sets on* X *and suppose* $\mu_0 : \Sigma_0(\mathsf{X}) \to [0, +\infty]$ *is a map such that:*

(i) *Definition 1.39(a) holds;*
(ii) μ_0 *satisfies 1.39(b) whenever* $\cup_{n\in\mathbb{N}} E_n \in \Sigma_0(\mathsf{X})$ *for* $E_k \in \Sigma_0(\mathsf{X})$*,* $k \in \mathbb{N}$*.*

If $\Sigma(\mathsf{X})$ *denotes the σ-algebra generated by*$\Sigma_0(\mathsf{X})$*, we have*
(i)

$$\Sigma(\mathsf{X}) \ni R \mapsto \mu(R) := \inf\left\{ \sum_{n\in\mathbb{N}} \mu_0(S_n) \;\middle|\; \{S_n\}_{n\in\mathbb{N}} \subset \Sigma_0(\mathsf{X}), \cup_{n\in\mathbb{N}} S_n \supset R \right\} \quad (1.3)$$

is a σ-additive positive measure on X *with respect to* $\Sigma(\mathsf{X})$ *that restricts to* μ_0 *on* $\Sigma_0(\mathsf{X})$.
(ii) *If* $\mathsf{X} = \cup_{n\in\mathbb{N}} \mathsf{X}_n$*, with* $\mathsf{X}_n \in \Sigma_0(\mathsf{X})$ *and* $\mu_0(\mathsf{X}_n) < +\infty$ *for any* $n \in \mathbb{N}$*, then* μ *is the unique σ-additive positive measure on* X*, with respect to* $\Sigma(\mathsf{X})$*, restricting to* μ_0 *on* $\Sigma_0(\mathsf{X})$.

As we shall use several kinds of positive measures and measure spaces henceforth, we need to gather some special instances in one place.

Definition 1.42. *A measure space* $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ *and its (positive, σ-additive) measure* μ *are called:*

(i) **finite**, *if* $\mu(\mathsf{X}) < +\infty$ *;*
(ii) **σ-finite**, *if* $\mathsf{X} = \cup_{n\in\mathbb{N}} E_n$, $E_n \in \Sigma(\mathsf{X})$ *and* $\mu(E_n) < +\infty$ *for any* $n \in \mathbb{N}$*;*
(iii) **probability space** *and* **probability measure**, *if* $\mu(\mathsf{X}) = 1$*;*
(iv) **Borel space** *and* **Borel measure**, *if* $\Sigma(\mathsf{X}) = \mathscr{B}(\mathsf{X})$ *with* X *locally compact Hausdorff space.*
In case μ *is a Borel measure, and more generally if* $\Sigma(\mathsf{X}) \supset \mathscr{B}(\mathsf{X})$, *with* X *locally compact and Hausdorff,* μ *is called:*
(v) **inner regular**, *if:*

$$\mu(E) = \sup\{\mu(K) \mid K \subset E\,, \quad K \quad \textit{is compact}\}$$

for any $E \in \Sigma(\mathsf{X})$*;*
(vi) **outer regular**, *if:*

$$\mu(E) = \inf\{\mu(V) \mid V \supset E\,, \quad V \quad \textit{is open}\}$$

for any $E \in \Sigma(\mathsf{X})$*;*
(vii) **regular**, *when simultaneously inner and outer regular.*
In the general case the measure μ *is* **concentrated on** $E \in \Sigma(\mathsf{X})$ *when:*

$$\mu(S) = \mu(E \cap S) \quad \textit{for any } S \in \Sigma(\mathsf{X}).$$

Remarks 1.43. Inner regularity requires that compact sets belong to the σ-algebra of sets on which the measure acts. In case of measures on σ-algebras including Borel's, this fact is true on locally compact Hausdorff spaces because compact sets are closed in Hausdorff spaces (Remark 1.22(1)) and hence they belong in the Borel σ-algebra. ■

A key notion, very often used in the sequel, is that of the *support* of a measure on a Borel σ-algebra.

Definition 1.44. *Let* $(\mathsf{X}, \mathscr{T}(\mathsf{X}))$ *be a topological space and* $\Sigma(\mathsf{X}) \supset \mathscr{B}(\mathsf{X})$. *The* **support** *of a (positive, σ-additive) measure* μ *on* $\Sigma(\mathsf{X})$ *is the closed subset of* X*:*

$$supp(\mu) := \mathsf{X} \setminus \bigcup_{O \in \mathscr{T}(\mathsf{X}),\, \mu(O)=0} O\,.$$

Note how the open set $\mathsf{X} \setminus supp(\mu)$ does not necessarily have zero measure. Still, the following is useful.

Proposition 1.45. *If* $\mu : \Sigma(\mathsf{X}) \to [0, +\infty]$ *is a σ-additive positive measure on* X *and* $\Sigma(\mathsf{X}) \supset \mathscr{B}(\mathsf{X})$, *then* μ *is concentrated on* $supp(\mu)$ *if at least one of the following conditions holds:*

(i) X *has a countable basis for its topology;*
(ii) X *is Hausdorff, locally compact and* μ *is inner regular.*

Proof. Let $A := \mathsf{X} \setminus supp(\mu)$ be the union (usually not countable) of all open sets in X with zero measure. Decompose $S \in \Sigma(\mathsf{X})$ into the disjoint union $S = (A \cap S) \cup (supp(\mu) \cap S)$; μ's additivity implies $\mu(S) = \mu(A \cap S) + \mu(supp(\mu) \cap S)$. By positivity and monotonicity $0 \leq \mu(A \cap S) \leq \mu(A)$, so the result holds provided $\mu(A) = 0$. Let us then prove $\mu(A) = 0$. In case (i), Lindelöf's lemma guarantees we can write A as a countable union of open sets of zero measure $A = \cup_{i \in \mathbb{N}} A_i$, and positivity plus sub-additivity force $0 \leq \mu(A) \leq \sum_{i \in \mathbb{N}} \mu(A_i) = 0$. Therefore $\mu(A) = 0$.

In case (ii), by inner regularity we have $\mu(A) = 0$ if $\mu(K) = 0$, for any compact set $K \subset A$. Since A is a union of zero-measure sets by construction, K will be covered by open sets of zero measure. By compactness then we may extract from there a finite covering $A_1, \ldots, A_n$. Again by positivity and sub-additivity, $0 = \mu(K) \leq \mu(A_1) + \ldots + \mu(A_n) = 0$, whence $\mu(K) = 0$, as requested. □

In conclusion we briefly survey zero-measure sets [Coh80, Rud82].

Definition 1.46. *If $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ is a measure space, a set $E \in \Sigma(\mathsf{X})$ has* **zero measure** *if $\mu(E) = \varnothing$. Then E is called a* **zero-measure set***, (more rarely, a null or negligible set). $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ and μ are called* **complete** *if, given any $E \in \Sigma(\mathsf{X})$ of zero measure, every subset in E belongs to $\Sigma(\mathsf{X})$ (so it has zero measure, by monotonicity). A property P is said to hold* **almost everywhere (with respect to** μ**)***, shortened to* **a.e.***, if P is true everywhere on X minus a set E of* **zero measure***.*

Remark 1.47. (1) Every measure space $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ can be made complete in the following manner.

Proposition 1.48. *If $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ is a (σ-additive, positive) measure space, there is a measure space $(\mathsf{X}, \Sigma'(\mathsf{X}), \mu')$, called the* **completion** *of $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$, such that:*

(i) $\Sigma'(\mathsf{X}) \supset \Sigma(\mathsf{X})$;
(ii) $\mu' \restriction_{\Sigma(\mathsf{X})} = \mu$;
(iii) $(\mathsf{X}, \Sigma'(\mathsf{X}), \mu')$ *is complete.*

The completion can be constructed in the two ensuing ways (yielding the same measure space).

(a) *Take the collection $\Sigma'(\mathsf{X})$ of $E \subset \mathsf{X}$ for which there exist $A_E, B_E \in \Sigma(\mathsf{X})$ with $B_E \subset E \subset A_E$ and $\mu(A_E \setminus B_E) = 0$. Then $\mu'(E) := \mu(A_E)$.*
(b) *Let $\Sigma'(\mathsf{X})$ be defined as the collection of subsets of X of the form $E \cup Z$, where $E \in \Sigma(\mathsf{X})$ and $Z \subset N_Z$ for some $N_Z \in \Sigma(\mathsf{X})$ with $\mu(N_Z) = 0$. Then $\mu'(E \cup Z) := \mu(E)$.*

It is quite evident from (b) that if $(\mathsf{X}, \Sigma_1(\mathsf{X}), \mu_1)$ is a complete measure space such that, once again, $\Sigma_1(\mathsf{X}) \supset \Sigma(\mathsf{X})$, $\mu_1 \restriction_{\Sigma(\mathsf{X})} = \mu$, then necessarily $\Sigma_1(\mathsf{X}) \supset \Sigma'(\mathsf{X})$ and $\mu_1 \restriction_{\Sigma'(\mathsf{X})} = \mu'$. In this sense the completion of a measure space is the *smallest* complete extension. When the initial measure space is already complete, the completion is the space itself.

Notice that the completion depends heavily on μ: in general, distinct measures on the same σ-algebra give rise to different completions.

Moreover, measurable functions for the completed σ-algebra are, generally speaking, no longer measurable for the initial one, whereas the converse is true: by completing the measurable space we enlarge the class of measurable functions.

(2) If (X,Σ,μ) is a measure space and $E \in \Sigma$, we may restrict Σ and μ to E like this: first of all we define $\Sigma\restriction_E := \{S \cap E \mid S \in \Sigma\}$ and $\mu\restriction_E(S) := \mu(S)$ for any $S \in \Sigma\restriction_E$. It should be clear that $(E,\Sigma\restriction_E,\mu\restriction_E)$ is a measure space corresponding to the natural restriction of the initial measure on E.

If $g : \mathsf{X} \to \mathbb{C}$ (respectively $\mathbb{R}$, $[-\infty,+\infty]$, $[0,+\infty]$) is a measurable function with respect to Σ, then by construction the restriction $g\restriction_E$ of g to E is measurable with respect to $\Sigma\restriction_E$.

Conversely, if $f : E \to \mathbb{C}$ ($\mathbb{R}$, $[-\infty,+\infty]$, $[0,+\infty]$) is measurable with respect to $\Sigma\restriction_E$, it is simple to show that its extension $\widetilde{f} : \mathsf{X} \to \mathbb{C}$ ($\mathbb{R}$, $[-\infty,+\infty]$, $[0,+\infty]$), with $\widetilde{f}(x) = f(x)$ if $x \in E$ and $\widetilde{f}(x) = 0$ otherwise, is measurable with respect to Σ.

(3) Keeping the above remark in mind one easily proves that if every $f_n : \mathsf{X} \to \mathbb{R}$, or $\mathbb{C}$, is measurable for $n \in \mathbb{N}$, $f(x) = \lim_{n\to+\infty} f_n(x)$ a.e. with respect to μ on X and we set $f(x) = c$ for some constant $c \in \mathbb{R}$, or $\mathbb{C}$, on the set N where $f(x)$ does not coincide with the limit of the sequence $f_n(x)$ (as this might not exist), then f is measurable.

If μ is complete, f turns out to be measurable irrespective of how we define it on N. ■

1.4.3 Integration of measurable functions

We are now ready to define the integral of a measurable function with respect to a σ-additive positive measure μ defined on a measurable space $(\mathsf{X},\Sigma(\mathsf{X}))$. We proceed in steps, defining the integral on a special class of functions first, and then extending to the measurable case.

The starting point are functions with values in $[0,+\infty] := [0,+\infty) \cup \{+\infty\}$. For technical reasons it is convenient to extend the notion of sum and product of non-negative real numbers so that $+\infty \cdot 0 := 0$, $+\infty \cdot r := +\infty$ if $r \in (0,+\infty]$, and $+\infty \pm r := +\infty$ if $r \in [0,+\infty)$.

A (non-negative) map $s : \mathsf{X} \to [0,+\infty]$ is called **simple** if it is measurable and its range is finite in $[0,+\infty]$. Such a function can be written, for certain $s_1,\ldots,s_n \in [0,+\infty) \cup \{+\infty\}$, as:

$$s = \sum_{i=1,\ldots,n} s_i \chi_{E_i}$$

where $E_1,E_2,\ldots E_n$ are pairwise-disjoint elements of $\Sigma(\mathsf{X})$ and χ_{E_i} the corresponding characteristic functions. The decomposition is not unique. Every function that can be written like this is simple. The **integral** of the simple map s with respect to μ is defined as the number in $[0,+\infty]$:

$$\int_{\mathsf{X}} s(x)d\mu(x) := \sum_{i=1,2,\ldots,n} s_i \mu(E_i)\,.$$

It is not difficult to show that the definition does not depend on the choice of decomposition of $s = \sum_{i=1,\ldots,n} s_i \chi_{E_i}$.

This notion can then be generalised to non-negative measurable functions in the obvious way: if $f : \mathsf{X} \to [0,+\infty]$ is measurable, let the **integral** of f with respect to

μ be:

$$\int_{\mathsf{X}} f(x)d\mu(x) := \sup\left\{ \int_{\mathsf{X}} s_n(x)d\mu(x) \;\middle|\; s \geq 0 \quad \text{is simple and} \quad s \leq f \right\}.$$

Note the integral may equal $+\infty$.

To justify the definition, we must remark that simple functions approximate with arbitrary accuracy non-negative measurable functions, as implied by the ensuing classical technical result [Rud82] (which we will state for complex functions and prove in Proposition 7.49).

Proposition 1.49. *If $f : \mathsf{X} \to [0,+\infty]$ is measurable, there exists a sequence of simple maps $0 \leq s_1 \leq s_2 \leq \cdots \leq s_n \leq f$ with $s_n \to f$ pointwise. The convergence is uniform if there is a $C \in [0,+\infty)$ such that $f(x) \leq C$ for all $x \in \mathsf{X}$.*

Note that the definition implies an elementary, yet important property of the integral.

Proposition 1.50. *If $f, g : \mathsf{X} \to [0,+\infty]$ are measurable and $f(x) \leq g(x)$ a.e. on X with respect to μ, then the integrals (in $[0,+\infty]$) satisfy:*

$$\int_{\mathsf{X}} f(x)d\mu(x) \leq \int_{\mathsf{X}} g(x)d\mu(x)\,.$$

To finish the contruction we define the integral of a complex-valued measurable function in the most natural way: by writing, that is, the function as sum of its real and imaginary parts and then decomposing the latter two real functions into their respective positive and negative parts. To overcome having to deal with awkward differences of infinites we must restrict the class of definition, which we do now by introducing *μ-integrable functions*.

Definition 1.51. *If $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ is a (σ-additive, positive) measure space, a measurable map $f : \mathsf{X} \to \mathbb{C}$ is* **integrable with respect to** μ *or* μ**-integrable**, *if:*

$$\int_{\mathsf{X}} |f(x)|d\mu(x) < +\infty\,.$$

Then we define the **integral of** f *on* X *with respect to* μ *as the complex number:*

$$\int_{\mathsf{X}} f(x)d\mu(x) = \int_{\mathsf{X}} Re(f)_+ d\mu(x) - \int_{\mathsf{X}} Re(f)_- d\mu(x)$$
$$+ i\left(\int_{\mathsf{X}} Im(f)_+ d\mu(x) - \int_{\mathsf{X}} Im(f)_- d\mu(x) \right),$$

where, if $g : \mathsf{X} \to \mathbb{R}$, we defined non-negative maps:

$$g_+(x) := \sup\{g(x), 0\} \quad \textit{and} \quad g_-(x) := -\inf\{g(x), 0\} \quad \textit{for any } x \in \mathbb{R}.$$

The set of μ-integrable functions on X will be indicated by $\mathscr{L}^1(\mathsf{X}, \mu)$.

If $f \in \mathscr{L}^1(\mathsf{X},\mu)$ and $E \subset \mathsf{X}$ is in the σ-algebra of X, we set:

$$\int_E f(x)d\mu(x) := \int_{\mathsf{X}} f(x)\chi_E(x)\, d\mu(x)\,,$$

where χ_E is the characteristic function of E.

It is no problem to check that the integral of $f : \mathsf{X} \to \mathbb{C}$ on X with respect to μ generalises the integral of a measurable function from X to $[0,+\infty)$. Also not hard is the following proposition, that clarifies the elementary features of the integral with respect to the measure μ.

Proposition 1.52. *If $(\mathsf{X},\Sigma(\mathsf{X}),\mu)$ is a (σ-additive, positive) measure space, then the measurable maps $f,g : \mathsf{X} \to \mathbb{C}$ satisfy:*

(a) *If $|f(x)| \leq |g(x)|$ a.e. on X, $g \in \mathscr{L}^1(\mathsf{X},\mu)$ implies $f \in \mathscr{L}^1(\mathsf{X},\mu)$.*
(b) *If $f = g$ a.e. on X then f and g are either both μ-integrable or neither is. In the former case*

$$\int_{\mathsf{X}} f(x)d\mu(x) = \int_{\mathsf{X}} g(x)d\mu(x)\,.$$

(c) *If f,g are μ-integrable, for any chosen $a,b \in \mathbb{C}$, so is $af+bg$; moreover,*

$$\int_{\mathsf{X}} af(x)+bg(x)d\mu(x) = a\int_{\mathsf{X}} f(x)d\mu(x) + b\int_{\mathsf{X}} g(x)d\mu(x)\,.$$

(d) *If $f \geq 0$ a.e. is μ-integrable, then:*

$$\int_{\mathsf{X}} f(x)d\mu(x) \geq 0$$

and the integral is null only if $f = 0$ a.e.
(e) *If f is μ-integrable, then:*

$$\left|\int_{\mathsf{X}} f(x)d\mu(x)\right| \leq \int_{\mathsf{X}} |f(x)|d\mu(x)\,.$$

Remarks 1.53. Consider the restrition $(E,\Sigma\restriction_E,\mu\restriction_E)$ of the measure space (X,Σ,μ) to the subset $E \in \Sigma$ as explained in 1.47(2). The extension of $f \in \mathscr{L}(E,\mu\restriction_E)$ to X, say $\widetilde{f}$, defined as the zero map outside E, satisfies $\widetilde{f} \in \mathscr{L}^1(\mathsf{X},\mu)$. Additionally,

$$\int_E f(x)d\mu\restriction_E(x) = \int_{\mathsf{X}} \widetilde{f}(x)d\mu(x) = \int_E \widetilde{f}(x)d\mu(x)\,. \qquad \blacksquare$$

The three central theorems of measure theory are listed below [Rud82].

Theorem 1.54 (Beppo-Levi's monotone convergence). *Let $(\mathsf{X},\Sigma(\mathsf{X}),\mu)$ be a (positive and σ-additive) measure space and $\{f_n\}_{n\in\mathbb{N}}$ a sequence of measurable functions $\mathsf{X} \to [0,+\infty]$ such that, a.e. at $x \in \mathsf{X}$, $0 \leq f_n(x) \leq f_{n+1}(x)$ for all $n \in \mathbb{N}$.*
Then:

$$\int_{\mathsf{X}} \lim_{n\to+\infty} f_n(x)d\mu(x) = \lim_{n\to+\infty}\int_{\mathsf{X}} f_n(x)d\mu(x) \leq +\infty\,,$$

where the map $\lim_{n\to+\infty} f_n(x)$ is zero where the limit does not exist, and the integral is the one defined for functions with values in $[0,+\infty]$.

Theorem 1.55 ("Fatou's lemma"). *Let* $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ *be a (σ-additive, positive) measure space and* $\{f_n\}_{n\in\mathbb{N}}$ *a sequence of measurable maps* $f_n : \mathsf{X} \to [0,+\infty]$.
Then:

$$\int_{\mathsf{X}} \liminf_{n\to+\infty} f_n(x)d\mu(x) \le \liminf_{n\to+\infty} \int_{\mathsf{X}} f_n(x)d\mu(x) \le +\infty\,,$$

the integral being the one defined for functions with values in $[0,+\infty]$.

Theorem 1.56 (Lebesgue's dominated convergence). *Let* $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ *be a (positive, σ-additive) measure space,* $\{f_n\}_{n\in\mathbb{N}}$ *a sequence of measurable maps* $f_n : \mathsf{X} \to \mathbb{C}$, *with* $f_n(x) \to f(x)$ *a.e. at* $x \in \mathsf{X}$ *as* $n \to +\infty$. *If there is a* μ*-integrable map* $g : \mathsf{X} \to \mathbb{C}$ *such that* $|f_n(x)| \le |g(x)|$ *a.e. at* $x \in \mathsf{X}$ *for any* $n \in \mathbb{N}$, *then* f *(set to zero where* $f_n(x) \not\to f(x)$*) is* μ*-integrable, and furthermore:*

$$\int_{\mathsf{X}} f(x)d\mu(x) = \lim_{n\to+\infty} \int_{\mathsf{X}} f_n(x)d\mu(x) \text{ and } \lim_{n\to+\infty} \int_{\mathsf{X}} |f(x) - f_n(x)|\, d\mu(x) = 0\,.$$

The next proposition (cf. Remark 1.47(1)) shows that the completion does not really affect integration, as we expect.

Proposition 1.57. *Let* $(\mathsf{X}, \Sigma, \mu)$ *be a measure space and* $(\mathsf{X}, \Sigma', \mu')$ *its completion. If* $f : \mathsf{X} \to \mathbb{C}$ *is measurable with respect to* Σ' *there exists a measurable* $g : \mathsf{X} \to \mathbb{C}$ *with respect to* Σ *with* $f = g$ *a.e. with respect to* μ. *If, moreover,* $f \in \mathscr{L}^1(\mathsf{X}, \mu')$, *then* $g \in \mathscr{L}^1(\mathsf{X}, \mu)$ *and*

$$\int_{\mathsf{X}} f(x)d\mu'(x) = \int_{\mathsf{X}} g(x)d\mu(x)\,.$$

Proof. Splitting f into real and imaginary parts and these into their positive and negative parts with the aid of 1.49, we can construct a sequence $s'_n := \sum_{i=1}^{M_n} c_i^{(n)} \chi_{E_i^{\prime(n)}}$ where $E_i^{\prime(n)} \in \Sigma'$, $E_i^{\prime(n)} \cap E_j^{\prime(n)} = \varnothing$ if $i \neq j$, $|s'_n(x)| \le |s'_{n+1}(x)| \le |f(x)|$ and $s'_n(x) \to f(x)$ everywhere on X as $n \to +\infty$. Because of Remark 1.47(1), we can write $E_i^{\prime(n)} = E_i^{(n)} \cup Z_i^{(n)}$ where $E_i^{(n)} \in \Sigma$, while $Z_i^{(n)} \subset N_i^{(n)} \in \Sigma$ with $\mu(N_i^{(n)}) = 0$. Define the maps, measurable with respect to Σ, $s_n := \sum_{i=1}^{M_n} c_i^{(n)} \chi_{E_i^{(n)} \setminus N_i^{(n)}}$. By construction $N := \cup_{n,i} N_i^{(n)}$ has zero μ-measure, being a countable union of zero-measure sets. Then set $g(x) = \lim_{n\to+\infty} s_n(x)$, measurable with respect to Σ as limit of measurable functions. The limit exists for any x, for it equals, by construction, 0 on N and $f(x)$ on $\mathsf{X} \setminus N$. Therefore we proved g is Σ-measurable and $g(x) = f(x)$ a.e. with respect to μ, as required. Now to the last statement. By construction $|s_n(x)| \le |s_{n+1}(x)| \le |g(x)|$, $|s'_n(x)| \le |s'_{n+1}(x)| \le |f(x)|$, $|s_n(x)| \to |g(x)|$, $|s'_n(x)| \to |f(x)|$ and $\int |s_n| d\mu = \int |s_n| d\mu' = \int |s'_n| d\mu'$. Therefore the monotone convergence theorem applied to the sequence $|s_n|$, with respect to both measures μ and μ', warrants that $g \in \mathscr{L}^1(\mathsf{X}, \mu)$ if $f \in \mathscr{L}^1(\mathsf{X}, \mu')$. By dominated convergence we finally have $\int_{\mathsf{X}} f d\mu' = \int_{\mathsf{X}} g d\mu$. □

1.4.4 Riesz's theorem for positive Borel measures

Moving on to Borel measures, we mention two important theorems. The first is the well-known *theorem of Riesz for positive Borel measures* [Coh80], which we shall often use in the sequel; it tells that every linear and positive functional on the space of continuous maps with compact support on a locally compact, Hausdorff space is actually an integral. From now on, given a topological space X, $C_c(\mathsf{X})$ will be the complex vector space of continuous maps $f : X \to \mathbb{C}$ with compact support. The vector-space structure of $C_c(\mathsf{X})$ comes from pointwise linear combinations of $f, g \in C_c(\mathsf{X})$, with $\alpha, \beta \in \mathbb{C}$:

$$(\alpha f + \beta g)(x) := \alpha f(x) + \beta g(x) \quad \text{for all } x \in \mathsf{X}.$$

Theorem 1.58 (Riesz's theorem for positive Borel measures). *Take a locally compact Hausdorff space X and consider a linear functional $\Lambda : C_c(\mathsf{X}) \to \mathbb{C}$ such that $\Lambda f \geq 0$ whenever $f \in C_c(\mathsf{X})$ satisfies $f \geq 0$. Then there exists a σ-additive, positive measure μ_Λ on the Borel σ-algebra $\mathscr{B}(\mathsf{X})$ such that:*

$$\Lambda f = \int_{\mathsf{X}} f d\mu_\Lambda \quad \textit{if } f \in C_c(X) \qquad \textit{and} \qquad \mu_\Lambda(K) < +\infty \textit{ when } K \subset \mathsf{X} \textit{ is compact.}$$

The measure μ_Λ can be chosen to be regular, in which case it is uniquely determined.

This result can be strengthened to produce a complete measure representing Λ, by extending $(\mathsf{X}, \mathscr{B}(\mathsf{X}), \mu_\Lambda)$ to its completion, and in particular enlarging the Borel σ-algebra in a way that depends on μ_Λ (regular, we may assume). In this way it is far from evident that the extended measure is still regular. But this is precisely what happens, because of the following, useful, fact [Coh80].

Proposition 1.59. *Let $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ be a measure space, where X is locally compact and Hausdorff and $\Sigma(\mathsf{X}) \supset \mathscr{B}(\mathsf{X})$. If μ is regular, the measure obtained by completing $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ is regular.*

A second valuable comment is that under certain assumptions on X, μ_Λ becomes automatically regular and hence uniquely determined by Λ. This is a consequence of a technical fact [Rud82, Theorem 2.18], which we recall here.

Proposition 1.60. *If ν is a positive Borel measure on a locally compact Hausdorff space X, and each open set is a countable union of compact sets of finite measure, then ν is regular.*

The second pivotal result is *Luzin's theorem* [Rud82], according to which on locally compact Hausdorff spaces, the functions of $C_c(\mathsf{X})$ approximate, so to speak, measurable functions when we work with measures on σ-algebras large enough to contain $\mathscr{B}(\mathsf{X})$ and satisfy further conditions (this happens in spaces with Lebesgue measure, that we shall see in short).

Theorem 1.61 (Luzin). *Let X be a locally compact Hausdorff space, μ a measure on a σ-algebra $\Sigma(\mathsf{X})$ such that:*

(i) $\Sigma(\mathsf{X}) \supset \mathscr{B}(\mathsf{X})$;
(ii) $\mu(K) < +\infty$ *if $K \subset \mathsf{X}$ is compact;*
(iii) μ *is regular;*
(iv) μ *is complete.*

Suppose $f : \mathsf{X} \to \mathbb{C}$ *is measurable such that* $f(x) = 0$ *if* $x \in \mathsf{X} \setminus A$*, for some* $A \in \Sigma(\mathsf{X})$ *with* $\mu(A) < +\infty$*. Then for any* $\varepsilon > 0$ *there is a* $g \in C_c(\mathsf{X})$ *such that:*

$$\mu(\{x \in \mathsf{X} \mid f(x) \neq g(x)\}) < \varepsilon .$$

Moreover, g can be chosen so that

$$\sup_{x\in X} |g(x)| \leq \sup_{x\in \mathsf{X}} |f(x)|.$$

Corollary 1.62. *Under the same assumptions of Theorem 1.61, if* $|f| \leq 1$ *there exists a sequence* $\{g_n\}_{n\in\mathbb{N}} \subset C_c(X)$ *with* $|g_n| \leq 1$ *for any* $n \in \mathbb{N}$ *and such that:*

$$f(x) = \lim_{n\to+\infty} g_n(x) \quad \textit{almost everywhere on } \mathsf{X}.$$

1.4.5 Differentiating measures

Definition 1.63. *If* μ, ν *are positive* σ*-additive measures defined on the same* σ*-algebra* Σ*:*

(a) ν *is called* **absolutely continuous** *with respect to* μ *(or* **dominated** *by* μ*), written* $\nu \prec \mu$*, whenever* $\nu(E) = 0$ *if* $\mu(E) = 0$ *with* $E \in \Sigma$*.*
(b) ν *is* **singular** *with respect to* μ *when there are* $A, B \in \Sigma$*,* $A \cap B = \varnothing$*, such that* μ *is concentrated on A and* ν *concentrated on B.*

Note μ is singular with respect to ν if and only if ν is singular with respect to μ. The paramount Radon–Nikodym theorem holds [Coh80, Rud82]. Recall that given a subset A of B, $\chi_A : B \to \mathbb{R}$ is the **characteristic function** of A if $\chi_A(x) = 1$ for $x \in A$ and $\chi_A(x) = 0$ otherwise.

Theorem 1.64 (Radon–Nikodým). *Suppose* μ *and* ν *are positive,* σ*-additive and* σ*-finite measures on the same* σ*-algebra* Σ *over* X*. If* $\nu \prec \mu$ *there exists a function* $\frac{d\nu}{d\mu} : \mathsf{X} \to [0, +\infty]$ *measurable such that:*

$$\nu(E) = \int_{\mathsf{X}} \chi_E \frac{d\nu}{d\mu} \, d\mu \quad \textit{for any } E \in \Sigma.$$

$\frac{d\nu}{d\mu}$ *is called the* **Radon–Nikodým derivative** *of* ν *in* μ*, and is determined by* μ *and* ν *up to sets of zero* μ*-measure. Furthermore,* $f \in \mathscr{L}^1(\mathsf{X}, \nu)$ *iff* $f \cdot \frac{d\nu}{d\mu} \in \mathscr{L}^1(\mathsf{X}, \mu)$*, in which case:*

$$\int_{\mathsf{X}} f d\nu = \int_{\mathsf{X}} f \frac{d\nu}{d\mu} d\mu .$$

1.4.6 Lebesgue's measure on $\mathbb{R}^n$

Lebesgue's measure on $\mathbb{R}^n$ is the prototype of all abstract positive measures. We define it, *a posteriori*, remembering what we proved in the previous sections. The starting point is the following proposition, itself a corollary of [Rud82, Theorem 2.20].

Proposition 1.65. *Fix $n = 1,2,\ldots$. There exists a unique σ-additive, positive Borel measure μ_n on $\mathbb{R}^n$ satisfying:*

(i) $\mu_n\left(\times_{k=1}^n [a_k, b_k]\right) = (b_1 - a_1)(b_2 - a_2)\cdots(b_n - a_n)$ *if* $a_k \leq b_k$, $a_k, b_k \in \mathbb{R}$*;*
(ii) μ_n *is invariant under translations:* $\mu_n(E + \mathbf{t}) = \mu_n(E)$ *for* $E \in \mathscr{B}(\mathbb{R}^n)$, $\mathbf{t} \in \mathbb{R}^n$.

It is possible to extend $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \mu_n)$ to a measure space $(\mathbb{R}^n, \mathscr{M}(\mathbb{R}^n), m_n)$, so that the measure m_n:

(i) *maps compact sets to finite values;*
(ii) *is complete;*
(iii) *is regular;*
(iv) *is translation-invariant.*

The extension is characterised as follows. If $A \subset \mathbb{R}^n$ then $A \in \mathscr{M}(\mathbb{R}^n)$ if and only if $F \subset A \subset G$ with $\mu_n(F \setminus G) = 0$, where $F, G \in \mathscr{B}(\mathbb{R}^n)$ are, respectively, an (at most) countable union and intersection of open sets. In such a case $m_n(A) := \mu_n(G)$.

Remarks 1.66. As a consequence, $\mathscr{M}(\mathbb{R}^n)$ is included in the completion of $\mathscr{B}(\mathbb{R}^n)$ with respect to μ_n (cf. Remark 1.47(1)). Since $\mathscr{M}(\mathbb{R}^n)$ is complete and the completion is the smallest complete extension, we conclude that $(\mathbb{R}^n, \mathscr{M}(\mathbb{R}^n), m_n)$ is nothing but the completion of $(\mathbb{R}^n, \mathscr{B}(\mathbb{R}^n), \mu_n)$. ■

Definition 1.67. *The measure m_n and the σ-algebra $\mathscr{M}(\mathbb{R}^n)$ determined by Proposition 1.65 are called* **Lebesgue measure on** $\mathbb{R}^n$ *and* **Lebsgue σ-algebra on** $\mathbb{R}^n$.
A function $f : \mathbb{R}^n \to \mathbb{C}$ (or $\mathbb{R}$) that is measurable with respect to $\mathscr{M}(\mathbb{R}^n)$ is said **Lebesgue measurable**.

Notation 1.68. From now on we shall often denote Lebesgue's measure by dx and not only m_n. For example,

$$m_n(E) = \int_{\mathbb{R}^n} \chi_E(x)dx \quad \text{if } E \in \mathscr{M}(\mathbb{R}^n)\,.$$

Sometimes we shall speak of **Lebesgue measure on a measurable subset**, like in *Lebesgue measure on* $[a,b]$. This will mean the restriction of Lebesgue's measure on $\mathbb{R}$ to $[a,b]$ in the sense of Remark 1.47(2). In such cases we shall tacitly follow Remark 1.53. In relation to the restricted Lebesgue we will drop the sign $\restriction_E$. Hence $\mathscr{L}^1([a,b], dx)$ will denote $\mathscr{L}^1([a,b], dx\restriction_{[a,b]})$, for example. ■

Remark 1.69. (1) Lebesgue's measure m_n is invariant under the whole isometry group of $\mathbb{R}^n$, not just under translations: therefore it is also invariant under rotations, reflections and any composition of these, translations included.
(2) Borel measurable maps $f : \mathbb{R}^n \to \mathbb{C}$ are thus Lebesgue measurable, but the converse is generally false. Continuous maps $f : \mathbb{R}^n \to \mathbb{C}$ trivially belong to both categories.
(3) The restriction of m_n to $\mathscr{B}(\mathbb{R}^n)$ is just the μ_n of Proposition 1.65, hence a *regular* Borel measure.
(4) Condition (i) in 1.65 implies on the one hand that already the Borel measure μ_n assigns finite values to compact sets, these being bounded in $\mathbb{R}^n$; on the other hand it

immediately implies, by monotonicity, that both μ_n and m_n give *non-zero measure* to non-empty open sets. This fact has an important consequence, expressed by the next useful, albeit simple, proposition.

Proposition 1.70. *Let $\mu : \mathscr{B}(\mathsf{X}) \to [0,+\infty]$ be a σ-additive positive measure on X such that $\mu(B) > 0$ if $B \neq \varnothing$ is open. (In particular μ can be Lebesgue's measure on $\mathbb{R}^n$ restricted to an open set $\mathsf{X} \subset \mathbb{R}^n$.) If $f : \mathsf{X} \to \mathbb{C}$ is continuous and $f(x) = 0$ a.e. with respect to μ, then $f(x) = 0$ for any $x \in \mathsf{X}$.*

Proof. $B := f^{-1}(\mathbb{C} \setminus \{0\})$ is open because f is continuous and $\mathbb{C} \setminus \{0\}$ is open. If we had $\mu(B) > 0$, then f would not be zero almost everywhere. Hence $\mu(B) = 0$ and we must have $B = \varnothing$, i.e. $f(x) = 0$ for all $x \in \mathsf{X}$. □

(5) Invariance under translations in 1.65 is extremely strong a requirement: one can prove [Rud82] that if $\nu : \mathscr{B}(\mathbb{R}^n) \to [0,+\infty]$ maps compact sets to finite values and is invariant under translations, there exists a constant $c \geq 0$ such that $\nu(E) = c m_n(E)$ for every $E \in \mathscr{B}(\mathbb{R}^n)$. ■

An established result, crucial in computations, is the following.

Proposition 1.71. *If $K = [a_1,b_1] \times \cdots \times [a_n,b_n] \subset \mathbb{R}^n$, with $-\infty < a_i < b_i < +\infty$ for $i = 1,\ldots,n$, consider $f : \mathbb{R}^n \to \mathbb{R}$ bounded on K with $f(x) = 0$ if $x \in \mathbb{R}^n \setminus K$.*

(a) *If $n = 1$, f is Riemann integrable on K if and only if it is continuous on K almost everywhere with respect to Lebesgue's measure on $\mathbb{R}$.*
(b) *If $n \geq 1$ and f is Riemann integrable on K, then it is Lebesgue measurable and Lebesgue integrable with respect to Lebesgue's measure on $\mathbb{R}^n$. Moreover,*

$$\int_{\mathbb{R}^n} f(x)dx = \int_K f(x)dx_R(x)\,,$$

where on the left is the Lebesgue integral, on the right the Riemann integral.

The two pivotal theorems of calculus, initially formulated for the Riemann integral, generalise to Lebesgue's integral on the real line as follows. Before that, we need some definitions.

Definition 1.72. *If $a,b \in \mathbb{R}$, a map $f : [a,b] \to \mathbb{C}$ has* **bounded variation** *on $[a,b]$ if, however we choose a finite number of points $a = x_0 < x_1 < \cdots < x_n = b$ in the interval, we have:*

$$\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leq C$$

where $C \in \mathbb{R}$ does not depend on the choice of points x_k, nor their number.

A subclass of functions of bounded variation is that of *absolutely continuous* maps.

Definition 1.73. *If $a,b \in \mathbb{R}$, $f : [a,b] \to \mathbb{C}$ is* **absolutely continuous** *on $[a,b]$ if for any $\varepsilon > 0$ there is a $\delta > 0$ such that, for any finite family of pairwise-disjoint, open subintervals (a_k,b_k), $k = 1,2,\ldots,n$,*

$$\sum_{k=1}^{n} (b_k - a_k) < \delta \quad \textit{implies} \quad \sum_{k=1}^{n} |f(b_k) - f(a_k)| < \varepsilon\,.$$

Remark 1.74. **(1)** Absolutely continuous functions on $[a,b]$ have bounded variation and are uniformly continuous (not conversely).
(2) Maps with bounded variation on $[a,b]$ and absolutely continuous ones on $[a,b]$ form vector spaces. The product of absolutely continuous maps on the compact interval $[a,b]$ is absolutely continuous.
(3) It is not hard to see that a differentiable map $f : [a,b] \to \mathbb{C}$ (admitting, in particular, left and right derivatives at the endpoints) with bounded derivative is absolutely continuous, hence also of bounded variation on $[a,b]$. A weaker version is this: $f : [a,b] \to \mathbb{C}$ is absolutely continuous if **Lipschitz**, i.e. if there is $L > 0$ such that $|f(x) - f(y)| \leq L|x-y|$, $x, y \in [a,b]$. ■

Now we are in the position to state [KoFo99] two classical results in real analysis, due to Lebesgue, that generalise the fundamental theorems of Riemann integration to Lebesgue's integral. Below, dx (and dt) are Lebesgue measures.

Theorem 1.75. *Fix $a, b \in \mathbb{R}$, $a < b$.*
(a) *If $f : [a,b] \to \mathbb{C}$ is absolutely continuous then it admits derivative $f'(x)$ for almost every $x \in [a,b]$ with respect to Lebesgue's measure. Defining, say, $f'(x) := 0$ where the derivative does not exist, f' becomes Lebesgue measurable, $f' \in \mathscr{L}^1([a,b], dx)$ and*

$$\int_a^x f'(t)dt = f(x) - f(a) \quad \text{for all } x \in [a,b].$$

(b) *If $f \in \mathscr{L}^1([a,b], dx)$ the map $[a,b] \ni x \mapsto \int_a^x f(t)dt$ is absolutely continuous, and*

$$\frac{d}{dx}\int_a^x f(t)dt = f(x) \quad \text{a.e. on } [a,b] \text{ with respect to Lebesgue's measure.}$$

To end the section, we mention a famous decomposition theorem for Borel measures on $\mathbb{R}$ that plays a certain role in spectral theory.

Let μ be a regular Borel measure (σ-additive, positive) on $\mathbb{R}$ with $\mu(K) < +\infty$ for any compact $K \subset \mathbb{R}$.

(i) The set $P_\mu := \{x \in \mathbb{R} \,|\, \mu(\{x\}) \neq 0\}$ is called set of **atoms** of μ (notice P_μ is either finite or countable);
(ii) μ is said **continuous** if $P_\mu = \varnothing$;
(iii) μ is a **purely atomic** measure if $\mu(S) = \sum_{p \in S} \mu(\{p\})$, $S \in \mathscr{B}(\mathbb{R})$.

A (σ-additive, positive) regular Borel measure μ on $\mathbb{R}$ with $\mu(K) < +\infty$ for any compact set $K \subset \mathbb{R}$ can be decomposed *uniquely* into:

$$\mu = \mu_{pa} + \mu_c\,,$$

where μ_{pa} is purely atomic and μ_c continuous, by setting:

$$\mu_{pa}(S) := \sum_{x \in P \cap S} \mu(P_\mu \cap S) \quad \forall S \subset \mathscr{B}(\mathbb{R}) \quad \text{and so } \mu_c := \mu - \mu_{pa}.$$

More precisely, a key decomposition theorem due to Lebesgue holds.

Theorem 1.76 (Lebesgue). *Let μ be a (σ-additive, positive) regular Borel measure on $\mathbb{R}$ with $\mu(K) < +\infty$ for any compact set $K \subset \mathbb{R}$. Then μ decomposes in a unique way as*

$$\mu = \mu_{ac} + \mu_{pa} + \mu_{sing} ,$$

where the regular Borel measures on $\mathbb{R}$ on the right are: (μ_{ac}) an absolutely continuous measure for Lebesgue's measure on $\mathbb{R}$, (μ_{pa}) a purely atomic measure (hence singular for Lebesgue's measure on $\mathbb{R}$) and (μ_{sing}) a continuous and singular measure for Lebesgue's measure on $\mathbb{R}$.

1.4.7 The product measure

If $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ and $(\mathsf{Y}, \Sigma(\mathsf{Y}), \nu)$ are measure spaces, we indicate with $\Sigma(\mathsf{X}) \otimes \Sigma(\mathsf{Y})$ the σ-algebra on $\mathsf{X} \times \mathsf{Y}$ generated by the family of rectangles $E \times F$ with $E \in \Sigma(\mathsf{X})$ and $F \in \sigma(\mathsf{Y})$.

If μ, ν are σ-finite, one can define uniquely a σ-finite measure on $\Sigma(\mathsf{X}) \otimes \Sigma(\mathsf{Y})$, written $\mu \otimes \nu$, by imposing

$$\mu \otimes \nu(E \times F) = \mu(E)\nu(F) \quad \text{if } E \in \Sigma(\mathsf{X}) \text{ and } F \in \Sigma(\mathsf{Y}).$$

$\mu \otimes \nu$ is called the **product** measure of μ, ν.

Remark 1.77. (1) We have the following fact [Rud82].

Proposition 1.78. *If f is measurable with respect to $\Sigma(\mathsf{X}) \otimes \Sigma(\mathsf{Y})$, then $\mathsf{Y} \ni y \mapsto f(x,y)$ and $\mathsf{X} \ni x \mapsto f(x,y)$ are measurable for any $x \in \mathsf{X}$ and $y \in \mathsf{Y}$, respectively.*

(2) The completion of the product of Lebesgue measures on $\mathbb{R}^n$ and $\mathbb{R}^m$ coincides with the Lebesgue measure on $\mathbb{R}^{n+m}$ [Rud82]. ■

The theorems of Fubini and Tonelli, which we state as one, hold.

Theorem 1.79 (Fubini and Tonelli). *Let $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$, $(\mathsf{Y}, \Sigma(\mathsf{Y}), \nu)$ be spaces with σ-finite measures, and consider a map $f : \mathsf{X} \times \mathsf{Y} \to \mathbb{C}$.*

(a) *If f is $\mu \otimes \nu$-integrable:*

(i) *$\mathsf{Y} \ni y \mapsto f(x,y)$ is ν-integrable for almost every $x \in \mathsf{X}$, and $\mathsf{X} \ni x \mapsto f(x,y)$ is μ-integrable for almost every $y \in \mathsf{Y}$;*

(ii) *$\mathsf{X} \ni x \mapsto \int_Y f(x,y)d\nu(y)$ and $\mathsf{Y} \ni y \mapsto \int_\mathsf{X} f(x,y)d\mu(x)$ (set to zero where the integrals do not exist) are integrable on X and on Y respectively. Moreover:*

$$\int_{\mathsf{X}\times\mathsf{Y}} f(x,y)d\mu \otimes d\nu(x,y) = \int_\mathsf{X} \left(\int_\mathsf{Y} f(x,y)d\nu(y) \right) d\mu(x)$$
$$= \int_\mathsf{Y} \left(\int_\mathsf{X} f(x,y)d\mu(x) \right) d\nu(y) .$$

(b) *Suppose f is measurable with respect to $\Sigma(\mathsf{X}) \otimes \Sigma(\mathsf{Y})$. Then:*

(i) *if $\mathsf{Y} \ni y \mapsto f(x,y)$ is ν-integrable for almost every $x \in \mathsf{X}$, or $\mathsf{X} \ni x \mapsto f(x,y)$ is μ-integrable for almost every $y \in \mathsf{Y}$, then the corresponding maps $\mathsf{X} \ni x \mapsto$*

$\int_Y |f(x,y)|d\nu(y)$ and $\mathsf{Y} \ni y \mapsto \int_\mathsf{X} |f(x,y)|d\mu(x)$ (null where the integrals are not defined) are measurable;

(ii) *if, additionally:*

$$\int_\mathsf{X}\left(\int_\mathsf{Y}|f(x,y)|d\nu(y)\right)d\mu(x) < +\infty$$

$$or \quad \int_\mathsf{Y}\left(\int_\mathsf{X}|f(x,y)|d\mu(x)\right)d\nu(y) < +\infty$$

respectively, then f is $\mu \otimes \nu$-integrable.

1.4.8 Complex (and signed) measures

We recall a few definitions and elementary results from the theory of complex functions [Rud82].

Definition 1.80. *A* **complex measure** *on X [Rud82], $\mu : \Sigma \to \mathbb{C}$, is a map associating a complex number to every element in a σ-algebra Σ on X so that:*

(i) *$\mu(\varnothing) = 0$ and*
(ii) *$\mu(\cup_{n\in\mathbb{N}}E_n) = \sum_{n=0}^{+\infty}\mu(E_n)$, independently of the summing order, for any collection $\{E_n\}_{n\in\mathbb{N}} \subset \Sigma$ with $E_n \cap E_m = \varnothing$ if $n \neq m$.*

Under (i)–(ii), if $\mu(\Sigma) \subset \mathbb{R}$, then μ is called a **signed measure** *or* **charge** *on X.*

Remark 1.81. (1) Requirement (ii) is equivalent to asking absolute convergence of the series $\sum_{n=0}^{+\infty}\mu(E_n)$ to $\mu(\cup_{n\in\mathbb{N}}E_n)$, by virtue of a generalisation of a classical result of Riemann on re-ordering real series that do not converge absolutely ([Rud64, Theorem 3.56]).

Theorem 1.82. *If $\{z_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$, the series $\sum_{n=0}^{+\infty} z_n$ converges absolutely ($\sum_{n=0}^{+\infty}|z_n| < +\infty$) if and only if there is $S \in \mathbb{C}$ such that $\sum_{n=0}^{+\infty} z_{P(n)} = S$ for any bijection $P : \mathbb{N} \to \mathbb{N}$.*

(2) There is a way to generate a finite positive measure from any complex (or signed) measure as follows. If $E \in \Sigma$, we shall say $\{E_i\}_{i\in I} \subset \Sigma$ is a **partition** of E if I is finite or countable, $\cup_{i\in I}E_i = E$ and $E_i \cap E_j = \varnothing$ for $i \neq j$. The *σ-additive positive measure* $|\mu|$ on Σ, called **total variation di** μ, is by definition:

$$|\mu|(E) := \sup\left\{\sum_{i\in I}|\mu(E_i)| \;\middle|\; \{E_i\}_{i\in I} \text{ partition of } E\right\} \quad \text{for any } E \in \Sigma.$$

$|\mu|$ clearly satisfies $|\mu|(E) \geq |\mu(E)|$ if $E \in \Sigma$. Moreover, $|\mu|(\mathsf{X}) < +\infty$ [Rud82]. $|\mu|$ is therefore a (σ-additive, positive) *finite* measure on Σ for any given complex measure μ. ■

In analogy to the real case, the *support* of a complex (or signed) measure defined on a Borel σ-algebra is defined herebelow.

Definition 1.83. *If* $(\mathsf{X}, \mathscr{T}(\mathsf{X}))$ *is a topological space and* $\Sigma(\mathsf{X}) \supset \mathscr{B}(\mathsf{X})$, *the* **support** *of a complex (or signed) measure* μ *on* $\Sigma(\mathsf{X})$ *is the closed subset of* X*:*

$$supp(\mu) := \mathsf{X} \setminus \bigcup_{O \in \mathscr{T}(\mathsf{X}),\, |\mu|(O)=0} O \,.$$

The definition of absolutely continuous measure with respect to a given measure generalises straighforwardly to complex measures.

Definition 1.84. *A complex (or signed) measure* ν *is* **absolutely continuous** *with respect to a given* σ*-additive, positive measure* μ*, or* **dominated by** μ, $\nu \prec \mu$, *whenever both are defined over one* σ*-algebra* Σ *on* X *and* $\mu(E) = 0$ *implies* $\nu(E) = 0$ *for* $E \in \Sigma$.

The theorem of Radon–Nikodým (Theorem 1.64) can be generalised to the case of complex/signed measures [Rud82]:

Theorem 1.85 (Radon–Nikodym theorem for complex and signed measures). *Let* ν *be a complex (or signed) measure,* μ *a* σ*-additive, positive and* σ*-finite measure, both defined on the* σ*-algebra* Σ *over* X*. If* $\nu \prec \mu$ *there exists a map* $\frac{d\nu}{d\mu} \in \mathscr{L}^1(\mathsf{X}, \mu)$ *such that:*

$$\nu(E) = \int_{\mathsf{X}} \chi_E \frac{d\nu}{d\mu}\, d\mu \quad \textit{for any } E \in \Sigma$$

where χ_E *is the characteristic function of* $E \subset \mathsf{X}$*. Such map* $\frac{d\nu}{d\mu}$ *is unique up to sets of zero* μ*-measure, and is called the* **Radon–Nikodým derivative** *of* ν *in* μ.

The following important result is a corollary of the above [Coh80, Rud82].

Theorem 1.86 (Characterisation of complex measures). *For any complex measure* μ *on a* σ*-algebra* Σ *on* X *there exists a measurable function* $h : \mathsf{X} \to \mathbb{C}$ *with* $|h| = 1$ *on* X*, unique up to redefinition on zero-measure sets, that belongs in* $\mathscr{L}^1(\mathsf{X}, |\mu|)$ *and such that* $\mu(E) = \int_E h\, d|\mu|$ *for all* $E \in \Sigma$.

The same result holds, with the obvious changes, for signed measures.

According to it, if $f \in \mathscr{L}^1(\mathsf{X}, |\mu|)$ we define the **integral of** f **with respect to the complex measure** μ by:

$$\int_{\mathsf{X}} f d\mu := \int_{\mathsf{X}} f h\, d|\mu| \,.$$

In Chapter 2 (Example 2.45(1)) we shall state a general version of Riesz's representation theorem for complex measures.

1.4.9 Exchanging derivatives and integrals

In this final section we state the pivotal theorem that allows to differentiate inside an integral for a general positive measure. The proof is an easy consequence of the dominated convergence theorem plus Lagrange's mean value theorem.

Theorem 1.87 (Differentiation inside an integral). *In relation to the measure space* $(\mathsf{X}, \Sigma(\mathsf{X}), \mu)$ *(σ-additive, positive), consider the family of maps* $\{h_t\}_{t\in A} \subset \mathscr{L}^1(\mathsf{X}, \mu)$ *where* $A \subset \mathbb{R}^m$ *is open and* $t = (t_1, \dots, t_m)$*. Assume that*

(i) *for some* $k \in \{1, 2, \dots, m\}$ *the derivative*

$$\frac{\partial h_t(x)}{\partial t_k}$$

exists for any $x \in \mathsf{X}$ *and* $t \in A$*;*

(ii) *there is a* $g \in \mathscr{L}^1(\mathsf{X}, \mu)$ *such that:*

$$\left|\frac{\partial h_t(x)}{\partial t_k}\right| \leq |g(x)| \text{ for any } t \in A, \text{ a.e. on } \mathsf{X}.$$

Then:

(a) $\mathsf{X} \ni x \mapsto \frac{\partial h_t}{\partial t_k} \in \mathscr{L}^1(\mathsf{X}, \mu)$*;*
(b) *for any* $t \in A$*, integral and derivative can be swapped:*

$$\frac{\partial}{\partial t_k}\int_{\mathsf{X}} h_t(x)d\mu(x) = \int_{\mathsf{X}} \frac{\partial h_t(x)}{\partial t_k} d\mu(x)\,. \tag{1.4}$$

Eventually:

(iii) *if, for a given g, condition (ii) holds simultaneously for all* $k = 1, 2, \dots, m$*, almost everywhere at* $x \in \mathsf{X}$*, and every function (for any fixed* $t \in A$*)*

$$A \ni t \mapsto \frac{\partial h_t(x)}{\partial t_k}$$

is continuous, then

(c) *the function:*

$$A \ni t \mapsto \int_{\mathsf{X}} h_t(x)d\mu(x)$$

belongs to $C^1(A)$*.*

Remarks 1.88. Theorem 1.87 is true also when taking a complex (or signed) measure μ and replacing $\mathscr{L}^1(\mathsf{X}, \mu)$ with $\mathscr{L}^1(\mathsf{X}, |\mu|)$ in the statement. The proof is direct, and relies on Theorem 1.86. ■

2

Normed and Banach spaces, examples and applications

I'm convinced mathematics is the most important investigating tool of the legacy of the human enterprise, it being the sorurce of everything.

René Descartes

In the book's first proper chapter, we will discuss the fundamental notions and theorems about normed and Banach spaces. We will introduce certain algebraic structures modelled on natural algebras of operators on Banach spaces. Banach operator algebras play a relevant role in modern formulations of Quantum Mechanics.

The chapter will, in essence, introduce the working language and the elementary topological instruments of the theory of linear operators. Even if mostly self-cointained, the chapter is by no means exhaustive if compared to the immense literature on the basic properties of normed and Banach spaces. The texts [Rud82, Rud91] should be consulted in this respect. In due course we shall specialise to operators on complex Hilbert spaces, with a short detour in Chapter 4 into the more general features of compact operators.

The most important notions of the present chapter are without any doubt *bounded operators* and the various *topologies (induced by norms or seminorms) on spaces of operators*. The relevance of these mathematical tools descends from the fact that the language of linear operators on linear spaces is the language used to formulate QM. Here the class of bounded operators plays a central technical part, even though in QM one is forced, on physical grounds, to introduce and work with unbounded operators too, as we shall see in the second part of the book.

The chapter's first part is devoted to the elementary notions of normed space, Banach space and their basic topological properties. We will discuss examples, like the space of continuous maps $C(\mathsf{K})$ over a compact space K, and prove the crucial theorem of Arzelà–Ascoli in this setup. In the examples will also prove key results such as the completeness of L^p spaces (Fischer–Riesz theorem).

In the second part we will define the norm of an operator and establish its main features.

Part three will discuss the fundamental theorems of Banach spaces, in their simplest versions. These are the theorems of Hahn–Banach, Banach–Steinhaus, plus the open mapping theorem a corollary of Baire's category theorem. We will prove a few useful technical consequences (the inverse operator theorem and the closed graph theorem). Then we will introduce the various operator topologies that come into play,

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_2, © Springer-Verlag Italia 2013

prove the theorem of Banach–Alaoglu and briefly recall the Krein–Milman theorem and Fréchet spaces.

Part four will be devoted to projection operators in normed spaces. This we will specialise in the next chapter to that – more useful for our purposes – of an orthogonal projector.

In the final two sections we will treat two elementary but important topics: equivalent norms (including a proof that n-dimensional normed spaces are Banach and homeomorphic to the standard $\mathbb{C}^n$) and the theory of contractions in complete normed spaces (including, in the examples, a proof of the local existence and uniqueness of solutions to first-order ODEs on $\mathbb{R}^n$ or $\mathbb{C}^n$). The latter will be the only instance of *nonlinear* functional analysis present in the book.

From now onwards we assume the reader is familiar with vector spaces and linear mappings [Ser94I].

2.1 Normed and Banach spaces and algebras

After reducing to normed spaces the notions of the previous chapter we will introduce Banach spaces. Then, by enriching the algebraic structure with an inner product, we will study normed and Banach algebras.

2.1.1 Normed spaces and essential topological properties

The first definitions we present are those of *norm*, *normed space* and *continuous map* between normed spaces.

Examples of normed spaces, very common in functional analysis and its physical applications, will be presented later, especially in the next section.

Definition 2.1. *Let* X *be a vector space over the field* $\mathbb{K} = \mathbb{C}$ *or* $\mathbb{R}$*. A map* $N : \mathsf{X} \to \mathbb{R}$ *is called* **norm** *on* X*, and* (X, N) *is a* **normed space***, if:*
N0. $N(u) \geq 0$ *for any* $u \in \mathsf{X}$*;*
N1. $N(\lambda u) = |\lambda| N(u)$ *for any* $\lambda \in \mathbb{K}$ *and* $u \in \mathsf{X}$*;*
N2. $N(u+v) \leq N(u) + N(v)$*, for any* $u, v \in \mathsf{X}$*;*
N3. $N(u) = 0 \Rightarrow u = 0$*, for any* $u \in \mathsf{X}$.
When **N0***,* **N1***,* **N2** *are valid but* **N3** *does not necessarily hold, N is called a* **seminorm**.

Remark 2.2. (1) Clearly, from **N1** descends $N(0) = 0$, while **N2** implies the inequality:

$$|N(u) - N(v)| \leq N(u-v) \quad \text{se } u, v \in \mathsf{X}. \tag{2.1}$$

(2) N1 is called *homogeneity* property, **N2** is known as *triangle inequality* or *subadditivity*. Together, **N0** and **N3** are referred to as *positive definiteness*, whereas **N0** alone is sometimes called *semi-definiteness*. ■

Notation 2.3. Henceforth $|| \ ||$ and $p(\ )$, with subscripts if necessary, will denote a norm and a seminorm respectively. Other symbols might be used as well. ■

An elementary yet fundamental notion is that of *open ball*.

Definition 2.4. *Let* $(\mathsf{X}, ||\ ||)$ *be a normed space.*
The **open ball of centre** $x_0 \in \mathsf{X}$ **and radius** $r > 0$ *is the set:*

$$B_r(x_0) := \{x \in \mathsf{X} \mid ||x - x_0|| < r\}\,.$$

A set $A \subset \mathsf{X}$ *is* **bounded** *if there exists an open ball* $B_r(x_0)$ *(of finite radius!) such that* $B_r(x_0) \supset A$.

We could define the same object using a seminorm p instead of a norm $||\ ||$: this will be done later.

Two useful properties of open balls (valid with seminorms, too), that follow immediately from **N2** and the definition, are:

$$B_\delta(y) \subset B_r(x) \qquad \text{if } y \in B_r(x) \text{ and } 0 < \delta + ||y - x|| < r, \tag{2.2}$$

$$B_r(x) \cap B_{r'}(x') = \varnothing \qquad \text{if } 0 < r + r' < ||x - x'||. \tag{2.3}$$

Let us introduce the natural topology of a normed space.

Definition 2.5. *Consider a normed space* $(\mathsf{X}, ||\ ||)$.

(a) $A \subset \mathsf{X}$ *is* **open** *if* $A = \varnothing$ *or* A *is the union of open balls.*
(b) *The* **norm topology** *of* X *is the family of open sets in* X.

Remark 2.6. (1) By (2.2) we have:

$$A \subset \mathsf{X} \textit{ is open iff } \forall x \in A,\ \exists r_x > 0 \textit{ such that } B_{r_x}(x) \subset A. \tag{2.4}$$

(2) By definition of open set and (2.2), (2.4), it follows that open sets as defined in 2.5 are also open according to 1.1. Hence the collection of open sets in a normed space is indeed an honest *topology*. The normed space X equipped with the above family of open sets is a true *topological space*. The collection of open balls with arbitrary centres and radii is a *basis* for the norm topology of the normed space $(\mathsf{X}, ||\ ||)$.
(3) Each normed space $(\mathsf{X}, ||\ ||)$ satisfies *Hausdorff's property*, cf. Definition 1.3 (and as such is a *Hausdorff space*). The proofs follows from (2.3) by choosing $A = B_r(x)$, $A' = B_{r'}(x')$ with $r + r' < ||x - x'||$; the latter is non-zero if $x \neq x'$, by property **N3**. Had we defined the topology using a *seminorm* (rather than a norm), the Hausdorff property would not have been guaranteed. ∎

Consider these facts, valid in any normed space: (a) open neighbourhoods can be chosen to be open balls (of radii ε and δ) and (b) each open neighbourhood of a point contains an open ball centred at that point (this follows from the definition of open set in a normed space and (2.2)). A straighforward consequence of (a) and (b) is that the continuity of a map, see (1.16), can be equivalently expressed as follows in normed spaces.

Definition 2.7. *A map* $f : \mathsf{X} \to \mathsf{Y}$ *between normed spaces* $(\mathsf{X}, ||\ ||_\mathsf{X})$ *and* $(\mathsf{Y}, ||\ ||_\mathsf{Y})$ *is* **continuous at** $x_0 \in \mathsf{X}$ *if for any* $\varepsilon > 0$ *there exists a number* $\delta > 0$ *such that* $||f(x) - f(x_0)||_\mathsf{Y} < \varepsilon$ *whenever* $||x - x_0||_\mathsf{X} < \delta$.
A map $f : \mathsf{X} \to \mathsf{Y}$ *is* **continuous** *if it is continuous at each point of* X.

Analogously, in normed spaces, convergent sequences (Definition 1.13) become:

Definition 2.8. *If* $(\mathsf{X}, ||\ ||)$ *is a normed space, the sequence* $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ **converges** *to* $x \in \mathsf{X}$*:*

$$x_n \to x \quad as \quad n \to +\infty \quad or \quad \lim_{n\to+\infty} x_n = x$$

if and only if, for any $\varepsilon > 0$ *there is* $N_\varepsilon \in \mathbb{R}$ *such that* $||x_n - x|| < \varepsilon$ *whenever* $n > N_\varepsilon$*; equivalently*

$$\lim_{n\to+\infty} ||x_n - x|| = 0\,.$$

The point x is the **limit of the sequence**.

Remarks 2.9. If $(\mathsf{X}, ||\ ||)$ is a normed space and $A \subset \mathsf{X}$, $x \in X$ is a limit point of A if and only if there is a sequence $\{x_n\}_{n\in\mathbb{N}} \subset A \setminus \{x\}$ converging to x.

In fact if x is a limit point for A, every open ball $B_{1/n}(x)$, $n = 1, 2, \ldots$, contains at least one $x_n \in A \setminus \{x\}$, and by construction $x_n \to x$ as $n \to +\infty$. Conversely, let $\{x_n\}_{n\in\mathbb{N}} \subset A \setminus \{x\}$ tend to x. Since every open neighbourhood B of x contains a ball $B_\varepsilon(x)$ centred at x by (2.4), the definition of convergence implies $B_\varepsilon(x)$, and so B, contains every x_n with $n > N_\varepsilon$ for some $N_\varepsilon \in \mathbb{R}$. Thus x is a limit point. ∎

A nice class of *continuous* linear functions is that of *isometries*.

Definition 2.10. *If* $(\mathsf{X}, ||\ ||_\mathsf{X})$*,* $(\mathsf{Y}, ||\ ||_\mathsf{Y})$ *are normed spaces over the same field* $\mathbb{C}$*, or* $\mathbb{R}$*, a linear map* $L : \mathsf{X} \to \mathsf{Y}$ *is called* **isometric***, or an* **isometry***, if* $||L(x)||_\mathsf{Y} = ||x||_\mathsf{X}$ *for all* $x \in \mathsf{X}$*. If the isometry* $L : \mathsf{X} \to \mathsf{Y}$ *is onto, it is an* **isomorphism of normed spaces***. Given an isomorphism L of normed spaces, the domain and codomain are said* **isomorphic (under** L**)**.

Remark 2.11. (1) It is obvious that an isometry $L : \mathsf{X} \to \mathsf{Y}$ is injective, by **N3**, but it may *not* be onto. If $\mathsf{X} = \mathsf{Y}$ and L is not surjective, then X is infinite-dimensional.
(2) Since the pre-image of an open ball under an isometry is an open ball, each isometry $f : \mathsf{X} \to \mathsf{Y}$ between normed spaces X, Y is continuous in the two topologies.
(3) If an isometry $f : \mathsf{X} \to \mathsf{Y}$ is onto (an isomorphism), its inverse $f^{-1} : \mathsf{Y} \to \mathsf{X}$ is still linear and isometric, hence an isomorphism from Y to X. An isomorphism of normed spaces is clearly a (linear) homeomorphisms of the two spaces.
(4) Other textbooks may provide a different definition, *not equivalent* to ours, of isomorphism of normed spaces. This usually requires an isomorphism be only a linear continuous map with continuous inverse (i.e. a linear homeomorphism). An isomorphism according to Definition 2.10 is also such in this second meaning, but not conversely. Having f, f^{-1} both continuous is much weaker a condition than preserving norms. For instance $f : \mathsf{X} \ni x \mapsto ax \in \mathsf{X}$, with $a \neq 0$ fixed, is an isomorphism from X to itself for the second definition, but not in our sense. ∎

A further technical result we mention about normed spaces is the direct analogue of something that happens in $\mathbb{R}$, normed with the absolute value.

Proposition 2.12. *A function* $f : \mathsf{X} \to \mathsf{Y}$ *between normed spaces* X, Y *is continuous at* $x \in \mathsf{X}$ *iff it is* **sequentially continuous** *at x, i.e.* $f(x_n) \to f(x)$*,* $n \to +\infty$*, for any sequence* $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ *such that* $x_n \to x$ *as* $n \to +\infty$*.*

Proof. If f is continuous at x, let $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ tend to x. By continuity, for any $\varepsilon > 0$ there is $\delta > 0$ such that $||f(x_n) - f(x)||_{\mathsf{Y}} < \varepsilon$ when $||x_n - x||_{\mathsf{X}} < \delta$. Since $||x_n - x|| \to 0$, then for any $\varepsilon > 0$ there is $N_\varepsilon \in \mathbb{R}$ such that $||f(x_n) - f(x)||_{\mathsf{Y}} < \varepsilon$ whenever $n > N_\varepsilon$. Thus f is sequentially continuous at x. Now assume f is *not* continuous at x_0, and let us show it cannot be sequentially continuous at x. With these assumptions there must be $\varepsilon > 0$ such that for any $n = 1, 2, \ldots$, there exists $x_n \in \mathsf{X}$ with $||x_n - x||_{\mathsf{X}} < 1/n$ but $||f(x) - f(x_n)||_{\mathsf{Y}} > \varepsilon$. The sequence $\{x_n\}_{n=1,2,\ldots}$ tends to x, but the corresponding $\{f(x_n)\}_{n=1,2,\ldots}$ does *not* converge to $f(x)$ in Y. Therefore f is not sequentially continuous at x. □

At last we want to discuss *continuity properties* of the *vector-space operations* with respect to the *norm topologies* on normed spaces.

If $(\mathsf{X}, || \, ||_{\mathsf{X}})$, $(\mathsf{Y}, || \, ||_{\mathsf{Y}})$ are normed spaces over the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, we can form the Cartesian product $\mathsf{Y} \times \mathsf{X}$ and its *product topology*, induced by the topologies of the factors X, Y (cf. 1.10). This topology has as open sets the empty set and the unions of Cartesian products of open balls in X and Y. In case $\mathsf{Y} = \mathsf{X}$, we can study the continuity of the *sum of two vectors* in $\mathsf{X} \times \mathsf{X}$:

$$+ : \mathsf{X} \times \mathsf{X} \ni (u, v) \mapsto u + v \in \mathsf{X},$$

where $\mathsf{X} \times \mathsf{X}$ has the product topology. From **N2**

$$||u + v|| \le ||u|| + ||v||,$$

making $+$ **jointly continuous** in its two arguments in the norm topologies; said otherwise, it is *continuous* in the product topology of the domain and the standard topology of the range.

In fact, the triangle inequality implies that given $(u_0, v_0) \in \mathsf{X} \times \mathsf{X}$ and $\varepsilon > 0$, then $u + v \in B_\varepsilon(u_0 + v_0)$ provided $(u, v) \in B_\delta(u_0) \times B_\delta(v_0)$ with $0 < \delta < \varepsilon/2$.

If Y is the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, thought of as normed by the absolute value norm, we can consider the continuity of the *product between a scalar and a vector* in $\mathbb{K} \times \mathsf{X}$:

$$\mathbb{K} \times \mathsf{X} \ni (\alpha, u) \mapsto \alpha u \in \mathsf{X},$$

where the left-hand side has the product topology. From **N2** and **N1**,

$$||\alpha u|| = |\alpha| \, ||u||$$

implies that the product scalar–vector is a **jointly continuous** operation in its arguments in the norm topologies; that is to say, it is *continuous* with respect to the product topology on the domain and the standard one on the range. Here, too, the proof is easy: from the above identity and **N2**, given $(\alpha_0, u_0) \in \mathbb{K} \times \mathsf{X}$ and $\varepsilon > 0$, then $\alpha u \in B_\varepsilon(\alpha_0 u_0)$ if we take $(\alpha, u) \in B^{(\mathbb{K})}_{\delta'}(\alpha) \times B_\delta(u_0)$ with $0 < \delta = \varepsilon/(2|\alpha_0| + 1)$ and $0 < \delta' < \varepsilon/(2(||u_0|| + \delta))$. ($B^{(\mathbb{K})}_{\delta'}(\alpha)$ is an open ball in $\mathbb{K}$ seen as normed space).

2.1.2 Banach spaces

Some of the material presented above can be adapted to completely general topological spaces. At the same time there are properties, like *completeness* (which we treat below), that befit the theory of normed spaces (and more generally *metrisable spaces*, which we will only mention elsewhere, in passing).

A well-known fact from the elementary theory on $\mathbb{R}^n$ is that convergent sequences $\{x_n\}_{n\in\mathbb{N}}$ in a normed space $(\mathsf{X}, ||\ ||)$ satisfy the so-called **Cauchy property**:

for any $\varepsilon > 0$ *there exists* $N_\varepsilon \in \mathbb{R}$ *such that* $||x_n - x_m|| < \varepsilon$ *whenever* $n, m > N_\varepsilon$.

In fact, if $||x_n - x|| \to 0$ as $n \to +\infty$, then $||x_k - x|| < \varepsilon$ for $k > N_\varepsilon$, so $||x_n - x_m|| \leq ||x_n - x|| + ||x_m - x|| < \varepsilon$ for $n, m > N_{\varepsilon/2}$.

A sequence satisfying this property is called a **Cauchy sequence**.

The idea of the above proof is that if a sequence converges to some point, its terms get closer to each other. It is interesting to see whether the converse holds as well: does a sequence of vectors that become closer always admit a limit?

As is well known from elementary calculus, the answer is yes on $\mathsf{X} = \mathbb{R}$ with the absolute value norm. Therefore also on $\mathbb{C}$ and on any vector space built over Cartesian products of standard copies of $\mathbb{R}^n$, $\mathbb{C}^n$. This is guaranteed by the fact that $\mathbb{R}$ satisfies the so-called *completeness axiom*.

Normed spaces in which every Cauchy sequence is convergent are called **complete** normed spaces. In general a normed space is not complete: complete normed spaces are scarce, hence interesting by default. They present relevant and useful features, especially for the physical applications that will be the object of the book.

Definition 2.13. *A normed space is called a* **Banach space** *if it is* **complete***, i.e. if any Cauchy sequence inside the space converges to a point of the space.*

Remarks 2.14. (1) The property of completeness is invariant under isomorphisms of normed spaces, but not under homeomorphisms (continuous maps with continuous inverses, not necessarily linear). A countrexample is given by $\mathbb{R}$ and $(0,1)$, both normed by the absolute value. Although the pair is homeomorphic, the line is complete, the interval not.
(2) It is easy to prove that any closed subspace M in a Banach space B is itself a Banach space for the restricted norm: each Cauchy sequence in M is Cauchy for B too, so it must converge to a point in B. But this point must belong to M because M is closed and contains its limit points. ∎

The spaces $\mathbb{C}^n$ and $\mathbb{R}^n$ with standard norm:

$$||(c_1, \ldots, c_n)|| = \sqrt{\sum_{k=1}^{n} |c_k|^2}$$

are the simplest instances of *finite-dimensional* Banach spaces, respectively complex and real. Actually it can be proved that every *finite-dimensional* complex Banach space is homeomorphic to a standard $\mathbb{C}^n$. We will prove this fact in Section 2.5, and show explicit examples of Banach spaces starting from the next section. At any rate, any normed space satisfies a nice property: it can be *completed* to a Banach space determined by it, in which it is *dense*.

Theorem 2.15 (Completion of Banach spaces). *Let* $(\mathsf{X}, ||\ ||)$ *be a normed vector space over* $\mathbb{K} = \mathbb{C}$ *or* $\mathbb{R}$.

(a) *There exists a Banach space* (Y, N) *over* $\mathbb{K}$, *called* **completion** *of* X, *such that* X *is isometrically identified to a dense subspace of* Y *under a linear injective mapping* $J : \mathsf{X} \to \mathsf{Y}$.
Put otherwise, there is a linear 1-1 *map* $J : \mathsf{X} \to \mathsf{Y}$ *with*

$$\overline{J(\mathsf{X})} = \mathsf{Y} \quad \textit{and} \quad N(J(x)) = ||x|| \ \textit{ for any } x \in \mathsf{X}.$$

(b) *If the triple* (J_1, Y_1, N_1), *with* $J_1 : \mathsf{X} \to \mathsf{Y}_1$ *a linear isometry and* (Y_1, N_1) *Banach on* $\mathbb{K}$, *is such that* $(\mathsf{X}, ||\ ||)$ *is isometric to a dense subspace of* Y_1 *under* J_1, *then there is a unique isomorphism of normed spaces* $\phi : \mathsf{Y} \to \mathsf{Y}_1$ *such that* $J_1 = \phi \circ J$.

Sketch of the proof. The idea is rather similar to the procedure for completing the set of rational numbers to the reals.

(a) Let C denote the space of Cauchy sequences in X and define the equivalence relation on C:

$$x_n \sim x'_n \quad \Leftrightarrow \quad \lim_{n\to\infty} ||x_n - x'_n|| = 0\,.$$

Clearly $\mathsf{X} \subset C/\sim$ by identifying each x of X with the equivalence class of the constant sequence $x_n = x$. Let J be the identification map. Then $C/\sim$ is easily a $\mathbb{K}$-vector space with norm induced by the structure of X. Now one must prove $C/\sim$ is complete, J is linear, isometric (hence 1-1) and $J(\mathsf{X})$ is dense in $\mathsf{Y} := C/\sim$.
(b) $J_1 \circ J^{-1} : J(\mathsf{X}) \to \mathsf{Y}_1$ is a linear and continuous isometry from a dense set $J(\mathsf{X}) \subset \mathsf{Y}$ to a Banach space Y_1, so it extends uniquely to a linear and continuous isometry ϕ on Y (see Proposition 2.44). As ϕ is isometric, it is injective. The same is true about the extension ϕ' of $J \circ J_1^{-1} : J_1(\mathsf{X}) \to \mathsf{Y}$, and by construction $(J \circ J_1^{-1}) \circ (J_1 \circ J^{-1}) = id_{J(\mathsf{X})}$. Extending to $\overline{J(\mathsf{X})} = \mathsf{Y}$ by continuity, we see $\phi' \circ \phi = id_{\mathsf{Y}}$, and similarly $\phi \circ \phi' = id_{\mathsf{Y}_1}$. In conclusion ϕ and ϕ' are onto, so in particular ϕ is an isomorphism of normed spaces and by construction $J_1 = \phi \circ J$. The uniqueness of an isomorphism $\phi : \mathsf{Y} \to \mathsf{Y}'$ satisfying $J_1 = \phi \circ J$ is easy, once one notices that each such map $\psi : \mathsf{Y} \to \mathsf{Y}_1$ fulfills $J - J = (\phi - \psi) \circ J$ by linearity, hence $(\phi - \psi) \restriction_{J(\mathsf{X})} = 0$. The uniqueness of the extension of $(\phi - \psi) \restriction_{J(\mathsf{X})}$, continuous and with dense domain $J(\mathsf{X})$, to $\overline{J(\mathsf{X})} = \mathsf{Y}$, eventually warrants that $\phi = \psi$.ll ☐

The next proposition is a useful criterion to check if a normed space is Banach.

Proposition 2.16. *Let* $(\mathsf{X}, ||\ ||)$ *be a normed space, and assume every absolutely convergent series* $\sum_{n=0}^{+\infty} x_n$ *of elements of* X *(i.e.* $\sum_{n=0}^{+\infty} ||x_n|| < +\infty$*) converges in* X. *Then* $(\mathsf{X}, ||\ ||)$ *is a Banach space.*

Proof. Take a Cauchy sequence $\{v_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ and let us show that if the above property holds, the sequence converges in X. Since the sequence is Cauchy, for any $k =$

$0,1,2,\dots$ there is N_k such that $||v_n - v_m|| < 2^{-k}$ whenever $n,m \geq N_k$. Choose N_k so that $N_{k+1} > N_k$ and extract the subsequence $\{v_{N_k}\}_{k\in\mathbb{N}}$. Now define vectors $z_0 := v_{N_1}$, $z_k := v_{N_{k+1}} - v_{N_k}$ and consider the series $\sum_{k=0}^{+\infty} z_k$. Notice $v_{N_k} = \sum_{k'=0}^{k} z_{k'}$. By construction $||z_k|| < 2^{-k}$, so the series converges absolutely. Under the assumptions made, there will be a $v \in \mathsf{X}$ such that:

$$\lim_{k\to+\infty} v_{N_k} = \lim_{k\to+\infty} \sum_{k'=0}^{k} z_{k'} = v\,.$$

Hence the subsequence $\{v_{N_k}\}_{k\in\mathbb{N}}$ of the Cauchy sequence $\{v_n\}_{n\in\mathbb{N}}$ converges to $v \in \mathsf{X}$. To finish it suffices to show that the whole $\{v_n\}_{n\in\mathbb{N}}$ converges to v. As

$$||v_n - v|| \leq ||v_n - v_{N_k}|| + ||v_{N_k} - v||\,,$$

for a given $\varepsilon > 0$ we can find N_ε such that $||v_n - v_{N_k}|| < \varepsilon/2$ whenever $n, N_k > N_\varepsilon$, because $\{v_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. On the other hand we can find M_ε such that $||v_{N_k} - v|| < \varepsilon/2$ whenever $k > M_\varepsilon$, since $v_{N_k} \to v$. Therefore taking $k > M_\varepsilon$ large enough, so that $N_k > N_\varepsilon$, we have $||v_n - v|| < \varepsilon$ for $n > N_\varepsilon$. As $\varepsilon > 0$ was arbitrary, we have $v_n \to v$ for $n \to +\infty$. □

2.1.3 Example: the Banach space $C(\mathsf{K};\mathbb{K}^n)$, the theorems of Dini and Arzelà–Ascoli

One of the simplest examples of a non-trivial (and generically, infinite-dimensional) Banach space is $C(\mathsf{K};\mathbb{K}^n)$, the space of continuous maps from a compact space K to $\mathbb{K}^n$, with $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. The chosen norm is the supremum norm $||f||_\infty := \sup_{x\in\mathsf{K}} ||f(x)||$. This is always finite for $f \in C(\mathsf{K};\mathbb{K}^n)$ (Proposition 1.21).

Proposition 2.17. *Let $\mathbb{K} = \mathbb{C}$ (or $\mathbb{R}$) and consider the normed space $(\mathbb{K}^n, ||\ ||)$ with norm 1.1. If K is compact the vector space $C(\mathsf{K};\mathbb{K}^n)$ of continuous functions from K to $\mathbb{K}^n$, equipped with the norm:*

$$||f||_\infty := \sup_{x\in\mathsf{X}} ||f(x)||$$

is a complex (or real) Banach space.

Proof. Let $\{f_n\}_{n\in\mathbb{N}} \subset C(\mathsf{K};\mathbb{K})$ be a Cauchy sequence. We want to show there is $f \in C(\mathsf{K};\mathbb{K})$ such that $||f_n - f||_\infty \to 0$ for $n \to +\infty$. Since $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy, for any given $x \in \mathsf{K}$, also n-tuples $f_n(x) \in \mathbb{K}^n$ are a Cauchy sequence. Thus, since $\mathbb{K}^n$ is complete, we have a pointwise-defined map:

$$f(x) := \lim_{n\to+\infty} f_n(x)\,.$$

The claim is that $f \in C(\mathsf{K};\mathbb{K})$ and $||f_n - f||_\infty \to 0$ as $n \to +\infty$. Since $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy, for any $\varepsilon > 0$ there is N_ε such that, if $n,m > N_\varepsilon$,

$$||f_n(x) - f_m(x)|| < \varepsilon\,, \quad \text{for } every\ x \in \mathsf{K}.$$

By definition of f, on the other hand, for a *given* $x \in \mathsf{K}$ and any $\varepsilon'_x > 0$, there is N_{x,ε'_x} such that $||f_m(x) - f(x)|| < \varepsilon'_x$ whenever $m > N_{x,\varepsilon'_x}$. Using these two facts we have

$$||f_n(x) - f(x)|| \leq ||f_n(x) - f_m(x)|| + ||f_m(x) - f(x)|| < \varepsilon + \varepsilon'_x$$

provided $n > N_\varepsilon$ and chosing $m > \max(N_\varepsilon, N_{x,\varepsilon'_x})$. Overall, if $n > N_\varepsilon$, then

$$||f_n(x) - f(x)|| < \varepsilon + \varepsilon'_x\,, \quad \text{for any } \varepsilon'_x > 0.$$

Since $\varepsilon'_x > 0$ is arbitrary, the inequality holds when $\varepsilon'_x = 0$, possibly becoming an equality. Thus the dependency on x disappears, and in conclusion, for any $\varepsilon > 0$ we have found $N_\varepsilon \in \mathbb{N}$ such that

$$||f_n(x) - f(x)|| \leq \varepsilon\,, \quad \text{for } all\ x \in \mathsf{K} \tag{2.5}$$

when $n > N_\varepsilon$. Hence $\{f_n\}$ converges to f *uniformly*. Since (2.5) holds for any $x \in \mathsf{K}$, it holds for the supremum on K: for any $\varepsilon > 0$ there is $N_\varepsilon \in \mathbb{N}$ such that $\sup_{x\in\mathsf{K}} ||f_n(x) - f(x)|| < \varepsilon$, whenever $n > N_\varepsilon$. Put differently,

$$||f_n - f||_\infty \to 0\,, \quad \text{as } n \to +\infty.$$

To finish we must prove f is continuous. Given $x \in \mathsf{K}$, for any $\varepsilon > 0$ we will find $\delta > 0$ such that $||f(x') - f(x)|| < \varepsilon$ when $||x' - x|| < \delta$. For that we exploit uniform convergence, and choose n such that $||f(z) - f_n(z)|| < \varepsilon/3$ for the given ε and any $z \in \mathsf{K}$. Furthermore, as f_n is continuous, there is $\delta > 0$ such that $||f_n(x') - f_n(x)|| < \varepsilon/3$ whenever $||x' - x|| < \delta$. Putting everything together and using the triangle inequality allows to conclude the following: if $||x' - x|| < \delta$,

$$||f(x') - f(x)|| \leq ||f(x') - f_n(x')|| + ||f_n(x') - f_n(x)|| + ||f_n(x) - f(x)||$$

$$< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

as claimed, so $f \in C(\mathsf{K};\mathbb{K})$. □

Notation 2.18. From now on we will write $C(\mathsf{K}) := C(\mathsf{K};\mathbb{C})$. ■

A useful analytical result about the uniform convergence of monotone sequences of real functions on compact sets is a classical result of Dini.

Theorem 2.19 (Dini's theorem on uniform convergence). *Let* K *be a compact space and* $\{f_n\}_{n\in\mathbb{N}} \subset C(\mathsf{K};\mathbb{R})$ *such that:*

(i) *each* f_n *is continuous;*
(ii) $f_n(x) \leq f_{n+1}(x)$ *for* $n = 1, 2, \ldots$ *and* $x \in \mathsf{K}$*;*
(iii) $f_n \to f$ *pointwise as* $n \to +\infty$.

Then, if f *is continuous,* $||f - f_n||_\infty \to 0$ *as* $n \to +\infty$. *The same is true if* (ii) *is replaced with:* $f_n(x) \geq f_{n+1}(x)$.

Proof. Fix $\varepsilon > 0$ and define $g_n := f - f_n$ for any $n \in \mathbb{N}$. Denote by B_n the set of $x \in \mathsf{K}$ for which $g_n(x) < \varepsilon$. B_n is open because g_n is continuous, and $B_{n+1} \supset B_n$ since $g_{n+1}(x) \leq g_n(x)$, by construction. Since $g_n(x) \to 0$, then necessarily $\cup_{n\in\mathbb{N}} B_n = \mathsf{K}$. But K is compact, so we can choose $B_{n_1}, B_{n_2}, \ldots B_{n_N}$ ordered so that $B_{n_{k+1}} \supset B_{n_k}$ and $B_{n_1} \cup B_{n_2} \cup \cdots \cup B_{n_N} \supset \mathsf{K}$. As $\mathsf{K} \supset B_{n_{k+1}} \supset B_{n_k}$, we have $B_{n_N} = \mathsf{K}$. Hence we have that for the given $\varepsilon > 0$, there is n_N such that $|f(x) - f_n(x)| < \varepsilon$ for $n > n_N$, $x \in \mathsf{K}$. Therefore $||f - f_n||_\infty < \varepsilon$, as claimed. The case $f_n(x) \geq f_{n+1}(x)$ is completely analogous. □

In the special case K is a compact set containing a dense and countable subset, the Banach space $C(\mathsf{K})$ has an interesting property by the *theorem of Arzelà–Ascoli*. We state below the simplest version of this result: even if it is not strictly related to the contents of this book, its relevance (especially in its more general form) and the typical argument of its proof make it worthy of attention.

Definition 2.20. *A sequence $\{f_n\}_{n\in\mathbb{N}}$ of functions $f_n : \mathsf{X} \to \mathbb{C}$ on a normed space*[1] *$(\mathsf{X}, ||\ ||)$ is* **equicontinuous** *if for any $\varepsilon > 0$ there is $\delta > 0$ such that $|f_n(x) - f_n(x')| < \varepsilon$ whenever $||x - x'|| < \delta$ for every $n \in \mathbb{N}$ and every $x, x' \in \mathsf{X}$.*

Theorem 2.21 (Arzelà–Ascoli). *Let K be a compact separable (cf. Definition 1.5) space. Suppose the sequence of functions $\{f_n\}_{n\in\mathbb{N}} \subset C(\mathsf{K})$ is:*

(a) *equicontinuous and*
(b) *bounded by some $C \in \mathbb{R}$, i.e. $||f_n||_\infty < C$ for any $n \in \mathbb{N}$.*

Then there exists a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ converging to some map $f \in C(\mathsf{K})$ in the topology induced by the norm $||\ ||_\infty$.

Proof. Consider the points q, labelled by $\mathbb{N}$, of a dense and countable set $Q \subset \mathsf{K}$. If q_1 denotes the first point, consider the values $|f_n(q_1)|$. They lie in a compact set $[0, C]$, so either there are finitely many, and $f_n(q_1) = x_1 \in \mathbb{C}$ for a single x_1 and infinitely many n, or the $f_n(q_1)$ accumulate at $x_1 \in \mathbb{C}$. In either case there is a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ such that $f_{n_k}(q_1) \to x_1 \in \mathbb{C}$ for some $x_1 \in \mathbb{C}$. Call the elements of $\{f_{n_k}\}_{k\in\mathbb{N}}$ by f_{1n}, where $n \in \mathbb{N}$. Now repeat the procedure and consider $|f_{1n}(q_2)|$, where q_2 is the second point of Q, and extract a subsequence $\{f_{2n}\}_{n\in\mathbb{N}}$ from $\{f_{1n}\}_{n\in\mathbb{N}}$. By construction, $f_{2n}(q_1) \to x_1$ and $f_{2n}(q_2) \to x_2 \in \mathbb{C}$, as $n \to +\infty$. Continuing in this way for every $k \in \mathbb{N}$ we build a subsequence $\{f_{kn}\}_{n\in\mathbb{N}}$ of $\{f_n\}_{n\in\mathbb{N}}$ that converges to $x_1, x_2, \ldots, x_k \in \mathbb{C}$ when evaluated at the points $q_1, q_2, \ldots, q_k \in Q$. Take the subsequence of $\{f_n\}_{n\in\mathbb{N}}$ formed by all diagonal terms in the various subsequences, $\{f_{nn}\}_{n\in\mathbb{N}}$. We will show that this is a Cauchy sequence for the norm $||\ ||_\infty$, and end the proof. So let us fix $\varepsilon > 0$ and find the $\delta > 0$ corresponding to $\varepsilon/3$ from equicontinuity, then cover K with balls of radius δ centred at every point of K. Using K's compactness we extract a finite covering of balls with radius δ, say $B^{(1)}_\delta, B^{(2)}_\delta, \ldots, B^{(N)}_\delta$ and choose $q^{(j)} \in B^{(j)}_\delta \cap Q$, for any $j = 1, \ldots, N$. For any $x \in B^{(j)}_\delta$ we have:

$$|f_{nn}(s) - f_{mm}(s)|$$
$$\leq |f_{nn}(s) - f_{nn}(q^{(j)})| + |f_{nn}(q^{(j)}) - f_{mm}(q^{(j)})| + |f_{mm}(q^{(j)}) - f_{mm}(s)|\,.$$

[1] The definition generalises trivially to metric spaces.

The first and third terms are smaller than $\varepsilon/3$ by construction. Since $f_{nn}(q^{(j)})$ converges in $\mathbb{C}$ as $n \to +\infty$, the second term is less than $\varepsilon/3$ provided $n,m > M_\varepsilon^{(j)}$ for some $M_\varepsilon^{(j)} \geq 0$. Hence if $M_\varepsilon = \max_{j=1,\ldots,N} M_\varepsilon^{(j)}$:

$$|f_{nn}(s) - f_{mm}(s)| < \varepsilon \quad \text{for } n,m > M_\varepsilon, \text{ and any } s \in \mathsf{K}.$$

In other words

$$||f_{nn} - f_{mm}||_\infty < \varepsilon \quad \text{if } n,m > M_\varepsilon$$

as claimed. □

Remark 2.22. **(1)** The theorem applies in particular when K is the closure of a non-empty open and bounded set $A \subset \mathbb{R}^n$, because points with rational coordinates form a countable dense subset in K. Moreover, the theorem holds, with the same proof (modulo trivial changes), in case we replace $C(\mathsf{K})$ with $C(\mathsf{K};\mathbb{K}^n)$.
(2) We will prove in Chapter 4, Proposition 4.3, that in a normed space $(\mathsf{X}, ||\ ||)$ a subset $A \subset \mathsf{X}$ is relatively compact (its closure is compact) if we can extract a convergent subsequence from any sequence of A. By virtue of this fact, the theorem of Arzelà–Ascoli says the following.

If K is a compact separable space, every equicontinuous subset of $C(\mathsf{K})$ that is bounded for $||\ ||_\infty$ is also relatively compact in $(C(\mathsf{K}), ||\ ||_\infty)$.

(3) An important result in functional analysis [Mrr01], which we will not prove, is the *Theorem of Banach–Mazur*: any complex separable Banach space is isometrically isomorphic to a closed subspace of $(C([0,1]), ||\ ||_\infty)$. ■

Several examples of Banach spaces will be presented at the end of the next section, after we have talked about normed and Banach algebras.

2.1.4 Normed algebras, Banach algebras and examples

As we shall see in a moment, in many applications there is a tight connection between *algebras* and normed spaces, which goes through linear operators on a normed space. The most important normed algebras in physics are, as a matter of fact, operator algebras.

But the notions of algebra and normed algebra are completely independent of operators. An algebra arises by enriching a vector space with an additional product that is associative, distributive over the sum and behaves associatively for the vector-scalar multiplication. A normed algebra is an algebra equipped with a norm that renders the vector space normed and behaves "well" with respect to the product. Here are the main definitions.

Definition 2.23. *A* **algebra** $\mathfrak{A}$ *over the field* $\mathbb{K} = \mathbb{C}$ *or* $\mathbb{R}$ *is a* $\mathbb{K}$*-vector space with an operation* $\circ : \mathfrak{A} \times \mathfrak{A} \to \mathfrak{A}$*, called* **product***, that is* **associative***:*

$$(a \circ b) \circ c = a \circ (b \circ c) \quad \textit{for each triple } a,b,c \in \mathfrak{A}$$

and distributes over the vector space operations:
A1. $a \circ (b+c) = a \circ b + a \circ c \quad \forall a,b,c \in \mathfrak{A}$;
A2. $(b+c) \circ a = b \circ a + c \circ a \quad \forall a,b,c \in \mathfrak{A}$;
A3. $\alpha(a \circ b) = (\alpha a) \circ b = a \circ (\alpha b) \quad \alpha \in \mathbb{K}$ *and* $\forall a,b \in \mathfrak{A}$.
The algebra $(\mathfrak{A}, \circ)$ *is called:*

commutative *(or* **Abelian***) if*
A4. $a \circ b = b \circ a$ *for any pair* $a,b \in \mathfrak{A}$;

with unit *if it contains an element* $\mathbb{I}$*, called* **unit** *of the algebra, such that:*
A5. $\mathbb{I} \circ u = u \circ \mathbb{I} = u$ *for any* $a \in \mathfrak{A}$;

normed algebra *or* **normed algebra with unit** *if it is a normed vector space with norm* $|| \; ||$ *satisfying*
A6. $||a \circ b|| \leq ||a||\,||b||$ *for* $a,b \in \mathfrak{A}$;
and in the second case also:
A7. $||\mathbb{I}|| = 1$;

Banach algebra *or* **Banach algebra with unit** *if* $\mathfrak{A}$ *is a Banach space plus a normed algebra, or normed algebra with unit, for the same norm.*

A **homomorphism** $\phi : \mathfrak{A}_1 \to \mathfrak{A}_2$ *between algebras whether with unit, normed or Banach, is a linear map preserving products, and units if present. In the obvious notations:*

$$\phi(a \circ_1 b) = \phi(a) \circ_2 \phi(b) \quad \text{if } a,b \in \mathfrak{A}_1, \qquad \phi(\mathbb{I}_1) = \mathbb{I}_2\,.$$

An algebra homomorphism is an **algebra isomorphism** *(between normed or Banach algebras, with or without unit) if bijective. If there is an isomorphism* $\phi : \mathfrak{A} \to \mathfrak{A}'$*, the algebras* $\mathfrak{A}$*,* $\mathfrak{A}'$ *(normed/Banach/with unit) are* **isomorphic**.

Remark 2.24. **(1)** If a unit exists it must be unique: if both $\mathbb{I}$ and $\mathbb{I}'$ satisfy **A5**, then $\mathbb{I}' = \mathbb{I}' \circ \mathbb{I} = \mathbb{I}$.
(2) For normed algebras, the axioms easily imply that all operations are *continuous* in the norm topologies involved. We showed in Section 2.1.1 that this is true for the sum and the multiplication scalar-vector. The product $\circ$, too, is jointly continuous in its arguments (i.e. in the product topology of $\mathfrak{A} \times \mathfrak{A}$), by **A6**. Let us see why. Dropping the symbol $\circ$, we have

$$||ab - a_0 b_0|| = ||ab - a_0 b + a_0 b - a_0 b_0|| \leq ||a - a_0||\,||b|| + ||a_0||\,||b - b_0||\,.$$

Fixing $\varepsilon > 0$, since the norm is continuous in its own-induced topology, there is $\delta_0 > 0$ such that $||b - b_0|| < \delta_0$ implies $-1 < ||b|| - ||b_0|| < 1$. Therefore:

$$||ab - a_0 b_0|| = ||ab - a_0 b + a_0 b - a_0 b_0|| \leq ||b - b_0||(1 + ||b_0|| + ||a_0||)\,.$$

Choosing now $\delta \leq \min(\delta_0, \varepsilon/(1 + ||b_0|| + ||a_0||))$ and considering elements $(a,b) \in B_\delta(a_0) \times B_\delta(b_0)$:

$$||ab - a_0 b_0|| < \varepsilon\,.$$

This proves the continuity of $\circ$ in the product topology of $\mathfrak{A} \times \mathfrak{A}$ (and hence also continuity in each argument alone).

Indicate by $G\mathfrak{A} \subset \mathfrak{A}$ the group of invertible elements and let us show that if $\mathfrak{A}$ is complete, i.e. a Banach algebra, then $G\mathfrak{A}$ is open (so it makes sense to invert elements in a neighbourhood of any $a_0 \in G\mathfrak{A}$) and the map $G\mathfrak{A} \ni a \mapsto a^{-1}$ continuous. With $a \in \mathfrak{A}$, the series

$$\sum_{n=0}^{+\infty}(-1)^n a^n$$

converges in the norm topology when $||a|| < 1$, because its partial sums are Cauchy sequences and the space is complete by hypothesis. The proof now is the same as for the convergence of the geometric series. Moreover, since the product is continuous:

$$(\mathbb{I}+a)\sum_{n=0}^{+\infty}(-1)^n a^n = \sum_{n=0}^{+\infty}(-1)^n(\mathbb{I}+a)a^n = \mathbb{I} + \lim_{n\to+\infty}(-1)^{n+1}a^n = \mathbb{I}\,.$$

Similarly:

$$\left(\sum_{n=0}^{+\infty}(-1)^n a^n\right)(\mathbb{I}+a) = \mathbb{I}\,.$$

Hence we have, if $||a|| < 1$, that $\mathbb{I}+a \in G\mathfrak{A}$ and:

$$(\mathbb{I}+a)^{-1} = \sum_{n=0}^{+\infty}(-1)^n a^n\,.$$

At this point, if $b \in G\mathfrak{A}$ we can write $c = b + c - b = b(\mathbb{I} + b^{-1}(c-b))$. Therefore $||b^{-1}(c-b)|| < 1$ implies c has an inverse:

$$c^{-1} = \sum_{n=0}^{+\infty}(-1)^n((c-b)b^{-1})^n b^{-1}\,.$$

In particular, if $b \in G\mathfrak{A}$ and we fix $0 < \delta < 1/||b^{-1}||$, then $c \in B_\delta(b)$ gives $c \in G\mathfrak{A}$, because: $||b^{-1}(c-b)|| \leq ||b^{-1}||\,||(c-b)|| < 1$. Thus we have proved $G\mathfrak{A}$ open.

Now to the continuity of $G\mathfrak{A} \ni a \mapsto a^{-1}$. Fix $a_0 \in G\mathfrak{A}$ and δ with $0 < \delta < ||a_0^{-1}||^{-1}$, and note $||a - a_0|| < \delta$ forces

$$\begin{aligned}||a^{-1} - a_0^{-1}|| &\leq ||a^{-1}(a_0 - a)a_0^{-1}|| \leq ||a^{-1}||\,||a - a_0||\,||a_0^{-1}|| \\ &\leq (||a^{-1} - a_0^{-1}|| + ||a_0^{-1}||)\delta||a_0^{-1}||\,.\end{aligned}$$

Therefore (the first factor is positive by construction)

$$(1 - \delta||a_0^{-1}||)\,||a^{-1} - a_0^{-1}|| \leq \delta||a_0^{-1}||^2\,.$$

We conclude that if $||a - a_0|| < \delta$,

$$||a^{-1} - a_0^{-1}|| \leq \frac{\delta}{1 - \delta||a_0^{-1}||}||a_0^{-1}||^2\,.$$

Defining $\varepsilon := \frac{\delta}{1-\delta||a_0^{-1}||}||a_0^{-1}||^2$ we have $\delta = \frac{\varepsilon}{\varepsilon||a_0^{-1}||+||a_0^{-1}||^2}$. The conclusion, as claimed, is that for any $\varepsilon > 0$ (satisfying the starting constraint) $||a^{-1} - a_0^{-1}|| < \varepsilon$ with $a \in B_\delta(a_0)$ and $\delta > 0$ above, so $a \mapsto a^{-1}$ is continuous.

(3) Observe that the norm does *not* show up in the definition of homomorphism and isomorphism between the various kinds of algebra.

(4) A (**normed/Banach**) **subalgebra** (**with unit**) is the obvious object: a subset $\mathfrak{A}_1 \subset \mathfrak{A}$ in a (normed/Banach) algebra $\mathfrak{A}$ (with unit) that inherits the algebra structure by *restricting* the algebra operations (if present: the same unit of $\mathfrak{A}$, the restricted norm of $\mathfrak{A}$, and completeness if $\mathfrak{A}$ is Banach). ■

Notation 2.25. In the sequel we will conventionally denote the product of two elements of an algebra by juxtaposition, as in ab, rather than by $a \circ b$; in other contexts a dot could be used: $f \cdot g$, especially when working with functions. ■

Examples 2.26. Let us see examples of Banach spaces and Banach algebras, a few of which will require some abstract measure theory.

(1) The number fields $\mathbb{C}$ and $\mathbb{R}$ are commutative Banach algebras with unit. For both the norm is the modulus/absolute value.

(2) Given any set X, and $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, let $L(\mathsf{X})$ be the set of **bounded** maps $f : \mathsf{X} \to \mathbb{K}$, i.e. $\sup_{x\in\mathsf{X}} |f(x)| < \infty$. $L(\mathsf{X})$ is naturally a $\mathbb{K}$-vector space for the usual linear combinations: if $\alpha, \beta \in \mathbb{C}$ and $f, g \in L(\mathsf{X})$,

$$(\alpha f + \beta g)(x) := \alpha f(x) + \beta g(x) \quad \text{for all } x \in \mathsf{X}.$$

We can define a product making $L(\mathsf{X})$ an algebra: for $f, g \in L(\mathsf{X})$,

$$(f \cdot g)(x) := f(x)g(x) \quad \text{for any } x \in \mathsf{X}.$$

The algebra is commutative and has a unit (the constant map 1). A norm that renders $L(\mathsf{X})$ Banach is the sup norm: $||f||_\infty := \sup_{x\in\mathsf{X}} |f(x)|$. The proof is simple (it uses the completeness of $\mathbb{C}$, and goes pointwise on X) and can be found in the exercises at the end of the chapter.

(3) Defining on the above X a σ-algebra Σ, the subalgebra of Σ-measurable functions $M_b(\mathsf{X}) \subset L(\mathsf{X})$ is closed in $L(\mathsf{X})$ in the topology of the sup norm. Thus $M_b(\mathsf{X})$ is a commutative Banach algebra. This is immdiate from the previous example, because the pointwise limit of measurable maps is measurable.

(4) The vector space of continuous maps from a topological space X to $\mathbb{C}$ is written $C(\mathsf{X})$; the symbol already appeared for X compact in Section 2.1.3.
$C_b(\mathsf{X}) \subset C(\mathsf{X})$ is the subspace of bounded continuous maps.
$C_c(\mathsf{X}) \subset C_b(\mathsf{X})$ the space of continuous maps with compact support.
They all coincide if X compact, and are clearly commutative algebras with the operations of example (2). $C(\mathsf{X})$ and $C_b(\mathsf{X})$ have a unit given by the constant map 1, whereas $C_c(\mathsf{X})$ has no unit when X is not compact. Here is list of general facts:

(a) $C_b(\mathsf{X})$ is a Banach algebra for the sup norm $|| \, ||_\infty$.

(b) If $\mathsf{X} = \mathsf{K}$ is compact, $C_c(\mathsf{K}) = C(\mathsf{K})$ is a Banach algebra with unit for the sup norm $|| \, ||_\infty$, as we saw in Section 2.1.3. An important result in the theory of Banach algebras [Rud91] states that *any commutative Banach algebra with unit over* $\mathbb{C}$ *is isomorphic to an algebra* $C(\mathsf{K})$ *for some compact* K.
(c) If the space X is

1. *Hausdorff*, and
2. *locally compact*,

then the completion of the normed space $C_c(\mathsf{X})$ is a commutative Banach algebra $C_0(\mathsf{X})$ (without unit), called algebra of continuous maps $f : \mathsf{X} \to \mathbb{C}$ that **vanish at infinity in** X [Rud82]: this means that for any $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset \mathsf{X}$ (depending on f in general) such that $|f(x)| < \varepsilon$ for any $x \in \mathsf{X} \setminus K_\varepsilon$.
(d) Irrepective of compactness, $C(\mathsf{X}), C_c(\mathsf{X}), C_0(\mathsf{X})$ are in general *not* dense in $M_b(\mathsf{X})$ for $|| \, ||_\infty$ and using on X the Borel σ-algebra. If we take $\mathsf{X} = [0,1]$ for instance, the bounded map $f : [0,1] \to \mathbb{C}$ equal to 0 except for $f(1/2) = 2$ is Borel measurable (with topologies: standard on $\mathbb{C}$ and induced one on $[0,1]$), hence $f \in M_b([0,1])$. However, there cannot exist any sequence of continuous maps $f_n : [0,1] \to \mathbb{C}$ converging to f uniformly. The same can be said if $\mathsf{X} \subset \mathbb{R}^n$ is a compact set with non-empty interior and we take $f : \mathsf{X} \to \mathbb{C}$ to be $f(q) = 0$ for $q \in \mathsf{X} \setminus \{p\}$, $p \in Int(\mathsf{X})$, and $f(p) = 1$.

(5) If X is Hausdorff and compact, consider in $C(\mathsf{X})$ a subalgebra A as follows. It must contain the unit (the function 1) and be closed under complex conjugation: $f \in A$ implies $f^* \in A$, where $f^*(x) := \overline{f(x)}$ for any $x \in \mathsf{X}$ and the bar denotes complex conjugation. Then A is said to **separate points** in X if given any $x, y \in \mathsf{X}$ with $x \neq y$, there is a map $f \in A$ with $f(x) \neq f(y)$. The *Stone–Weierstrass* theorem [Rud91] states the following.

Theorem 2.27 (Stone–Weierstrass). *Let* X *be a compact Hausdorff space and consider the Banach algebra with unit* $(C(\mathsf{X}), || \, ||_\infty)$. *Then any subalgebra* $\mathfrak{A} \subset C(\mathsf{X})$ *containing the unit, closed under complex conjugation and separating points, has* $C(\mathsf{X})$ *as closure with respect to* $|| \, ||_\infty$.

A typical example is X compact in $\mathbb{R}^n$ and $\mathfrak{A}$ the algebra of complex polynomials in n variables (the standard coordinates of $\mathbb{R}^n$) restricted to X. The theorem says that these polynomials approximate uniformly any complex, continuous function on X. This is useful to construct bases in Hilbert spaces, as we will explain.

(6) Let $(\mathsf{X}, \Sigma, \mu)$ be a positive, σ-additive *measure space*. Recall this means a set X, a σ-algebra Σ of subsets in X, a positive and σ-additive measure $\mu : \Sigma \to [0, +\infty]$. Then we have **Hölder's inequality** and **Minkowski's inequality**, respectively:

$$\int_{\mathsf{X}} |f(x)g(x)| d\mu(x) \leq \left(\int_{\mathsf{X}} |f(x)|^p d\mu(x) \right)^{1/p} \left(\int_{\mathsf{X}} |g(x)|^q d\mu(x) \right)^{1/q} \tag{2.6}$$

$$\left(\int_{\mathsf{X}} |f(x)+g(x)|^p d\mu(x) \right)^{1/p} \leq \left(\int_{\mathsf{X}} |f(x)|^p d\mu(x) \right)^{1/p} + \left(\int_{\mathsf{X}} |g(x)|^p d\mu(x) \right)^{1/p} \tag{2.7}$$

for $f, g : \mathsf{X} \to \mathbb{C}$ measurable, $p, q > 0$ subject to $1/p + 1/q = 1$ in the former, $p \geq 1$ in the latter [Rud82]. These inequalities are proved in two exercises at the end of the chapter.

Let $\mathscr{L}^p(\mathsf{X},\Sigma,\mu)$, or henceforth $\mathscr{L}^p(\mathsf{X},\mu)$ by omitting the σ-algebra, be the set of Σ-measurable maps $f:\mathsf{X}\to\mathbb{C}$ such that $\int_{\mathsf{X}}|f(x)|^p d\mu(x)<\infty$. Using Minkowski's inequality one proves easily $\mathscr{L}^p(\mathsf{X},\mu)$ is a vector space with the linear composition of functions, and

$$P_p(f) := \left(\int_{\mathsf{X}} |f(x)|^p d\mu(x)\right)^{1/p} \tag{2.8}$$

is a *seminorm*. Since $P_p(f)=0$ if and only if $f=0$ a.e. for μ, to obtain a norm (to have **N3**) we must identify the zero map and any function that differs from it by a zero-measure set. To this end we define an equivalence relation on $\mathscr{L}^p(\mathsf{X},\mu)$: $f\sim g$ iff $f-g$ is zero a.e. for μ. The quotient space $\mathscr{L}^p(\mathsf{X},\mu)/\sim$, written $L^p(\mathsf{X},\mu)$, inherits a vector-space structure over $\mathbb{C}$ from $\mathscr{L}^p(\mathsf{X},\mu)$ by setting:

$$[f]+[g] := [f+g] \quad \text{and} \quad \alpha[f] := [\alpha f] \quad \text{for any } \alpha\in\mathbb{C},\ f,g\in\mathscr{L}^p(\mathsf{X},\mu).$$

It is not hard to see the left-hand sides of these definitions are independent of the representatives chosen in the equivalence classes on the right.
It can be proved that $L^p(\mathsf{X},\mu)$ is a Banach space for:

$$||[f]||_p := \left(\int_{\mathsf{X}} |f(x)|^p d\mu(x)\right)^{1/p}, \tag{2.9}$$

where f is any representative of $[f]\in L^p(\mathsf{X},\mu)$. We shall slightly abuse the notation in the sequel, and write $||f||_p$, not $P_p(f)$, when dealing with functions and not equivalence classes.

If $(\mathsf{X},\Sigma',\mu')$ is the completion of (X,Σ,μ) (cf. Remark 1.47(1)), in general $\mathscr{L}^p(\mathsf{X},\mu')$ is larger than $\mathscr{L}^p(\mathsf{X},\mu)$. But if we pass to the quotient then $L^p(\mathsf{X},\mu') = L^p(\mathsf{X},\mu)$ by way of Proposition 1.57.

Theorem 2.28 (Fischer–Riesz). *If (X,Σ,μ) is a positive, σ-additive measure space, the associated normed space $L^p(\mathsf{X},\mu)$ is, for any $1\le p<+\infty$, a Banach space.*

Proof. Throughout this proof we will omit the square brackets in denoting coset elements of $L^p(\mathsf{X},\mu)$, and identify them with functions (up to null sets). To prove the claim, thanks to Proposition 2.16 it is sufficient to verify that if the series $\sum_{n=0}^{+\infty} f_n$ in $\mathscr{L}^p(\mathsf{X},\mu)$ converges absolutely, $\sum_{n=0}^{+\infty}||f_n||_p\le K<+\infty$, then $\sum_{n=0}^{+\infty} f_n = f$ a.e. for some $f\in\mathscr{L}^p(\mathsf{X},\mu)$ in the topology of $||\ ||_p$. We will need an auxiliary sequence: $g_N(x) := \sum_{n=1}^N |f_n(x)|$, $N=1,2,\ldots$. By construction $||g_N||_p\le\sum_{n=1}^N||f_n||_p\le K$ for any $N=1,2,\ldots$. We show the limit $lim_{N\to+\infty} g_N(x)$ is finite for almost all $x\in\mathsf{X}$. The sequence of integrable functions g_N^p is non-negative and non-decreasing by construction, and $\int_{\mathsf{X}} g_N(x)^p d\mu(x) < K^p$ for any N. By *monotone convergence* the limit g^p of g_N^p exists, as a map in $[0,+\infty]$, because the sequence of the given $g_N^p\ge 0$ is non-decreasing, and must have finite integral; thus $g^p\ge 0$ is finite up to possible zero-measure sets. As $p\ge 0$, at points $x\in\mathsf{X}$ where $g(x)^p<+\infty$ we have $\lim_{N\to+\infty} g_N(x) = g(x) < +\infty$. By construction, where $g(x)$ is finite the series $\sum_{n=0}^{+\infty} f_n(x)$ converges absolutely. Therefore it converges to certain values $f(x)\in\mathbb{C}$. Defining $f(x)=0$ where the series of f_n does not converge, we obtain a series $\sum_{n=0}^{+\infty} f_n$ that converges a.e. to a map $f : \mathsf{X} \to \mathbb{C}$ (measurable since limit, a.e., of measurable functions, and, say, null on the zero-measure set where the series does not converge). The map f belongs to $\mathscr{L}^p(\mathsf{X},\mu)$: if $f_N(x) := \sum_n^N f_n(x)$, the sequence of $|f_N|^p$ is non-negative and $\int_\mathsf{X} |f_N(x)|^p d\mu(x) < K^p$ for any N. By *Fatou's lemma* $f \in \mathscr{L}^p(\mathsf{X},\mu)$. Now we prove $\int_\mathsf{X} |f_N(x) - f(x)|^p d\mu(x) \to 0$ as $n \to +\infty$. Easily (see the footnote in Exercise 2.14) $|f_N(x) - f(x)|^p \leq 2^p(|f_N(x)|^p + |f(x)|^p)$. Since, by construction, $|f_N|^p + |f|^p \leq |g|^p + |f|^p \in \mathscr{L}^1(\mathsf{X},\mu)$, we can invoke the *dominated convergence theorem* for the sequence of $|f_N - f|^p$, known to converge a.e. to 0, and obtain $\int_\mathsf{X} |f_N(x) - f(x)|^p d\mu(x) \to 0$ as $n \to +\infty$. We have thus proved that the initial series $\sum_{n=0}^{+\infty} f_n$, assumed absolutely convergent in $\mathscr{L}^p(\mathsf{X},\mu)$, has $\sum_{n=0}^{+\infty} f_n = f$ a.e. for the above $f \in \mathscr{L}^p(\mathsf{X},\mu)$ in norm $|| \;||_p$. This ends the proof. □

This argument implies a technical fact, extremely useful in the applications, that deserves separate mentioning.

Proposition 2.29. *Take* $1 \leq p < +\infty$ *and let* (X,Σ,μ) *be a* σ*-additive, positive measure space. If* $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{L}^p(\mathsf{X},\mu)$ *converges to* f *in* $|| \;||_p$ *as* $n \to +\infty$*, there exists a subsequence* $\{f_{n_k}\}_{k\in\mathbb{N}}$ *such that* $f_{n_k} \to f$ *a.e. for* μ.

Proof. The sequence $\{f_n\}_{n\in\mathbb{N}}$ is convergent hence Cauchy, and we can extract a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ such that $||f_{n_{k+1}} - f_{n_k}|| \leq 2^{-k}$. Define a telescopic sequence $s_k := f_{n_{k+1}} - f_{n_k}$. The series $f_{n_0} + \sum_{k=1}^{+\infty} s_k$ is absolutely convergent, for $\sum_{k=1}^{+\infty} ||s_k||_p < \sum_{k=1}^{+\infty} 2^{-k} < +\infty$. As in the proof of Theorem 2.28, we conclude (a) the sum $s \in \mathscr{L}^p(\mathsf{X},\mu)$ of the series exists, in the sense of $|| \;||_p$ convergence, and (b) the series converges pointwise to s a.e.: $f_{n_0}(x) + \sum_{k\in\mathbb{N}} s_k(x) = s(x)$. Since $f_{n_0}(x) + \sum_{k=0}^{+\infty} s_k(x) = f_{n_k}(x)$, what we have found is that $f_{n_k} \to s \in \mathscr{L}^p(\mathsf{X},\mu)$ both pointwise μ-almost everywhere, and with respect to $|| \;||_p$. But by assumption $f_{n_k} \to f \in \mathscr{L}^p(\mathsf{X},\mu)$ with respect to $|| \;||_p$, so $||f - s||_p = 0$ and hence $f(x) = s(x)$ a.e. for μ. Eventually, then, $f_{n_k}(x) \to f(x)$ a.e. for μ. □

To conclude the example notice the Banach space $L^p(\mathsf{X},\mu)$ is, in general, not an algebra (for the usual pointwise product of functions) as the pointwise product in $\mathscr{L}^p(\mathsf{X},\mu)$ does not normally belong to the space.

(7) In reference to example (6), consider the special case where X is not countable, Σ the power set of X and μ the **counting measure**: if $S \subset \mathsf{X}$,

$$\mu(S) = \text{number of elements of } S, \text{ with } \mu(S) = \infty \text{ if } S \text{ is infinite.}$$

In this case, given a measurable space Y, any map $f : \mathsf{X} \to \mathsf{Y}$ is measurable, and $L^p(\mathsf{X},\mu)$ is simply denoted $\ell^p(\mathsf{X})$. Its elements are "sequences" $\{z_x\}_{x\in\mathsf{X}}$ of complex numbers, labelled by X, such that:

$$\sum_{x\in\mathsf{X}} |z_x|^p < \infty\,,$$

where the sum is given by:

$$\sup\left\{ \sum_{x\in\mathsf{X}_0} |z_x|^p \;\middle|\; \mathsf{X}_0 \subset \mathsf{X},\; \mathsf{X}_0 \text{ finite} \right\}.$$

If X is countable, $\mathsf{X} = \mathbb{N}$ or $\mathbb{Z}$ in particular, the above definition of sum of *positive* numbers indexed by X is the usual sum of a series. For example, $\ell^p(\mathbb{N})$ is the space of sequences $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ with:

$$\sum_{n=0}^{+\infty} |c_n|^p < +\infty .$$

(8) With $(\mathsf{X}, \Sigma, \mu)$ a measure space, consider the class $\mathscr{L}^\infty(\mathsf{X}, \mu)$ of complex measurable maps $f : \mathsf{X} \to \mathbb{C}$ such that $|f(x)| < M_f$ a.e. for μ, for some $M_f \in \mathbb{R}$ (depending on f). $\mathscr{L}^\infty(\mathsf{X}, \mu)$ has a natural structure of vector space and of commutative algebra with unit (the function 1) if we use the ordinary product and linear combinations as in example (2). We can give $\mathscr{L}^\infty(\mathsf{X}, \mu)$ a seminorm:

$$P_\infty(f) := \text{ess sup}|f|$$

where the **essential supremum of** $f \in \mathscr{L}^\infty(\mathsf{X}, \mu)$ is:

$$\text{ess sup}|f| := \inf\{r \in \mathbb{R} \mid \mu\left(\{x \in \mathsf{X} \,|\, |f(x)| > r\}\right) = 0\} . \tag{2.10}$$

Naïvely speaking, the latter is the “smallest” upper bound of $|f|$ when ignoring what happens on zero-masure sets.

In particular (exercise):

$$P_\infty(f \cdot g) \le P_\infty(f) P_\infty(g) \quad \text{if } f, g \in \mathscr{L}^\infty(\mathsf{X}, \mu) .$$

As for $\mathscr{L}^p$, if we identify maps that differ by zero-measure sets, we can form the complex quotient space $L^\infty(\mathsf{X}, \mu)$, where the product:

$$[f] \cdot [g] := [f \cdot g] \quad \text{se } f, g \in \mathscr{L}^\infty(\mathsf{X}, \mu)$$

is well defined. Exactly as for the L^p spaces, the seminorm P_∞ is (clearly) a norm on $L^\infty(\mathsf{X}, \mu)$:

$$||[f]||_\infty := \text{ess sup}|f| .$$

As for the L^p, $L^\infty(\mathsf{X}, \mu)$ is again a Banach space. Being closed under products it is a Banach algebra as well.

Theorem 2.30 (Fischer–Riesz, L^∞ case). *If $(\mathsf{X}, \Sigma, \mu)$ is a σ-additive, positive measure space, the associated normed space $L^\infty(\mathsf{X}, \mu)$ is a Banach space.*

Proof. As customary, we indicate with f (no brackets) the generic element of $L^\infty(\mathsf{X}, \mu)$, and identify it when necessary with a function (up to null sets). Let $\{f_n\}_{n\in\mathbb{N}} \subset L^\infty(\mathsf{X}, \mu)$ be a Cauchy sequence for $||\ ||_\infty$. Define, for $k, m, n \in \mathbb{N}$, sets $A_k := \{x \in \mathsf{X} \mid |f_k(x)| > ||f_k||_\infty\}$ and $B_{n,m} := \{x \in \mathsf{X} \mid |f_n(x) - f_m(x)| > ||f_n - f_m||_\infty\}$. By construction $E := \cup_{k\in\mathbb{N}} \cup_{n,m\in\mathbb{N}} A_k \cup B_{n,m}$ must have zero measure, and the sequence of f_n converges uniformly in $\mathsf{X} \setminus E$ to some f, that is therefore bounded. Extend f to the entire X by setting it to zero on $\mathsf{X} \setminus E$. Thus $f \in L^\infty(\mathsf{X}, \mu)$ and $||f_n - f||_\infty \to 0$ as $n \to +\infty$. □

(9) Going back to example (8), in case Σ is the power set of X and μ the counting measure, the space $L^\infty(\mathsf{X},\mu)$ is written simply $\ell^\infty(\mathsf{X})$. Its points are "sequences" $\{z_x\}_{x\in\mathsf{X}}$ of complex numbers indexed by X such that $\sup_{x\in\mathsf{X}}|z_x| < +\infty$. With the notation of example (2), $\ell^\infty(\mathsf{X}) = L(\mathsf{X})$. ■

Notation 2.31. The literature prefers using the naked letter f to indicate the equivalence class $[f] \in L^p(\mathsf{X},\mu)$, $1 \le p \le \infty$. We shall stick to this convention when no confusion arises. ■

2.2 Operators, spaces of operators, operator norms

With the next definition we introduce *linear operators* and *linear functionals*, whose importance is paramount in the whole book. We shall assume from now on familiarity with the elementary theory of linear operators (matrices) on finite-dimensional vector spaces, and freely use results without explicit mention.

Definition 2.32. *Let* X, Y *be vector spaces on the same field* $\mathbb{K} := \mathbb{R}, \mathbb{C}$.

(a) $T : \mathsf{X} \to \mathsf{Y}$ *is a* **linear operator** *(simply, an* **operator***)* **from** X **to** Y *if it is linear:*

$$T(\alpha f + \beta g) = \alpha T(f) + \beta T(g) \quad \textit{for any } \alpha,\beta \in \mathbb{K},\ f,g \in \mathsf{X}.$$

$\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *denotes the set of linear operators from* X *to* Y.
If X *and* Y *are normed,* $\mathfrak{B}(\mathsf{X},\mathsf{Y}) \subset \mathfrak{L}(\mathsf{X},\mathsf{Y})$ *is the subset of continuous operators. In particular* $\mathfrak{L}(\mathsf{X}) := \mathfrak{L}(\mathsf{X},\mathsf{X})$ *and* $\mathfrak{B}(\mathsf{X}) := \mathfrak{B}(\mathsf{X},\mathsf{X})$.
(b) $T : \mathsf{X} \to \mathbb{K}$ *is a* **linear functional** *(a* **functional***)* **on** X *if it is linear.*
(c) *We call the space* $\mathsf{X}^* := \mathfrak{L}(\mathsf{X},\mathbb{K})$ *the* **algebraic dual** *of* X*, whereas* $\mathsf{X}' := \mathfrak{B}(\mathsf{X},\mathbb{K})$ *is the* **topological dual** *(the* **dual***)* **of** X*, with* $\mathbb{K}$ *normed by the absolute value.*

Notation 2.33. Linear algebra textbooks usually write Tu for $T(u)$ when $T : \mathsf{X} \to \mathsf{Y}$ is a linear operator and $u \in \mathsf{X}$, and we will adhere to this. ■

If $T,S \in \mathfrak{L}(\mathsf{X},\mathsf{Y})$ and $\alpha,\beta \in \mathbb{K}$, the linear combination $\alpha T + \beta S$ is the usual map: $(\alpha T + \beta S)(u) := \alpha(Tu) + \beta(Su)$ for any $u \in \mathsf{X}$.

Thus $\alpha T + \beta S$ is still in $\mathfrak{L}(\mathsf{X},\mathsf{Y})$. As linear combinations preserve continuity, we have the following.

Proposition 2.34. *Let* X, Y *be vector spaces on the same* $\mathbb{K} := \mathbb{R}$*, or* $\mathbb{C}$.
$\mathfrak{L}(\mathsf{X},\mathsf{Y})$, $\mathfrak{L}(\mathsf{X})$, X^*, $\mathfrak{B}(\mathsf{X},\mathsf{Y})$, $\mathfrak{B}(\mathsf{X})$ *and* X' *are* $\mathbb{K}$*-vector spaces.*

Another fundamental notion we introduce is that of *bounded* operator (or functional). We begin with an elementary, yet important, fact.

Theorem 2.35. *Let* $(\mathsf{X},|| \ ||_\mathsf{X}),(\mathsf{Y},|| \ ||_\mathsf{Y})$ *be normed spaces on the same* $\mathbb{K} = \mathbb{C}$ *or* $\mathbb{R}$ *and take* $T \in \mathfrak{L}(\mathsf{X},\mathsf{Y})$.

(a) *The following conditions are equivalent:*

(i) *there exists* $K \in \mathbb{R}$ *such that* $||Tu||_{\mathsf{Y}} \leq K||u||_{\mathsf{X}}$ *for all* $u \in \mathsf{X}$;

(ii) $\sup_{u \in \mathsf{X}\setminus\{0\}} \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}} < +\infty$.

(b) *If either of* (i), (ii) *hold:*

$$\sup\left\{ \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}} \;\middle|\; u \in \mathsf{X}\setminus\{0\} \right\} = \inf\{K \in \mathbb{R} \mid ||Tu||_{\mathsf{Y}} \leq K||u||_{\mathsf{X}} \;\textit{for any}\; u \in \mathsf{X}\}\ .$$

Proof. (a) Under (i), $\sup_{u \in \mathsf{X}\setminus\{0\}} \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}} \leq K < +\infty$ by construction. If (ii) holds, set $A := \sup_{u \in \mathsf{X}\setminus\{0\}} \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}}$, and then $K := A$ satisfies (i).
(b) Calling I the greatest lower bound of the numbers K fulfilling (i), since

$$\sup_{u \in \mathsf{X}\setminus\{0\}} \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}} \leq K\,,$$

we have $\sup_{u \in \mathsf{X}\setminus\{0\}} \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}} \leq I$. If the two sides are different then there is K_0 with $\sup_{u \in \mathsf{X}\setminus\{0\}} \frac{||Tu||_{\mathsf{Y}}}{||u||_{\mathsf{X}}} < K_0 < I$, whence $||Tu||_{\mathsf{Y}} < K_0||u||_{\mathsf{X}}$ for any $u \neq 0$, so $||Tu||_{\mathsf{Y}} \leq K_0||u||_{\mathsf{X}}$ for all $u \in \mathsf{X}$. Therefore K_0 satisfies (i), and $I \leq K_0$ by definition of I, in contradiction to $K_0 < I$. □

Notation 2.36. We will start omitting indices in norms when the corresponding spaces will be clear from the context. ■

Definition 2.37. *Let* X, Y *be normed spaces on the field* $\mathbb{C}$ *or* $\mathbb{R}$. $T \in \mathfrak{L}(\mathsf{X}, \mathsf{Y})$ *is* **bounded** *if either condition in Theorem 2.35(a) holds. The number*

$$||T|| := \sup_{||u|| \neq 0} \frac{||Tu||}{||u||} \tag{2.11}$$

is called **(operator) norm** *of* T.

Remark 2.38. **(1)** From the definition of $||T||$, if $T : \mathsf{X} \to \mathsf{Y}$ is bounded then:

$$||Tu|| \leq ||T||\,||u||\,, \quad \text{for any } u \in \mathsf{X}\,. \tag{2.12}$$

(2) The notion of bounded linear operator cannot clearly correspond to that of a bounded function. That is because the image of a linear map, in a vector space, cannot be bounded precisely because of linearity. Proposition 2.39 shows, though, that it still makes sense to view "boundedness" in terms of the bounded image of an operator, provided one restricts the domain to a bounded set. ■

The operator norm can be computed in alternative ways, at times useful for proofs. In this respect,

Proposition 2.39. *Let* X, Y *be normed spaces on* $\mathbb{C}$ *or* $\mathbb{R}$.
$T \in \mathfrak{L}(\mathsf{X}, \mathsf{Y})$ *is bounded if and only if the right-hand side of any of the identities below exists and is finite, in which case:*

$$||T|| = \sup_{||u||=1} ||Tu|| \,, \tag{2.13}$$

$$||T|| = \sup_{||u||\leq 1} ||Tu|| \,, \tag{2.14}$$

$$||T|| = \inf\{K \in \mathbb{R} \mid ||Tu|| \leq K||u|| \text{ for any } u \in \mathsf{X}\} \,. \tag{2.15}$$

Proof. That T is bounded if and only if the right-hand side of (2.13) is finite, and the validity of (2.13) too, follow from the linearity of T and **N1**. T is bounded iff the right-hand side of (2.14) is finite, and (2.14) holds, by the following argument. The set of u with $||u|| \leq 1$ includes $||u|| = 1$, so $\sup_{||u||\leq 1} ||Tu|| \geq \sup_{||u||=1} ||Tu||$. On the other hand, $||u|| \leq 1$ implies $||Tu|| \leq ||Tv||$ for some v with $||v|| = 1$ (any such if $u = 0$, and $v = u/||u||$ otherwise). Hence $\sup_{||u||\leq 1} ||Tu|| \leq \sup_{||u||=1} ||Tu||$, from which $\sup_{||u||\leq 1} ||Tu|| = \sup_{||u||=1} ||Tu||$, as claimed. That T is bounded if and only if the right-hand side of (2.15) is finite, and property (2.15), are consequences of Theorem 2.35(b). □

There is a relationship between *continuity* and *boundedness* of linear operators and functionals, which makes boundedness very important. The following simple theorem shows that linear operators are precisely the continuous ones.

Theorem 2.40. *Consider* $T \in \mathfrak{L}(\mathsf{X}, \mathsf{Y})$ *with* X, Y *normed over the same field* $\mathbb{R}$ *or* $\mathbb{C}$. *The following are equivalent facts:*

(i) *T is continuous at* 0*;*
(ii) *T is continuous;*
(iii) *T is bounded.*

Proof. (i) ⇔ (ii). Continuity trivially implies continuity at 0. We show the latter forces continuity. As $(Tu) - (Tv) = T(u - v)$ we have $(\lim_{u\to v} Tu) - Tv = \lim_{u\to v}(Tu - Tv) = \lim_{(u-v)\to 0} T(u - v) = 0$ by continuity at 0.
(i) ⇒ (iii). From continuity at 0 there is $\delta > 0$ such that $||u|| < \delta$ implies $||Tu|| < 1$. Fixing $\delta' > 0$ with $\delta' < \delta$, if $v \in \mathsf{X} \setminus \{0\}$, then $u = \delta' v/||v||$ has norm smaller than δ, so $||Tu|| < 1$, i.e. $||Tv|| < (1/\delta')||v||$. Therefore Theorem 2.35(a) holds with $K = 1/\delta'$, and by Definition 2.37 T is bounded.
(iii) ⇒ (i). This is obvious: if T is bounded then $||Tu|| \leq ||T||||u||$, hence continuity at 0. □

The name "norm" for $||T||$ is not accidental; namely, the operator norm renders $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$, *hence also* $\mathfrak{B}(\mathsf{X})$ *and* X', normed spaces, as we shall shortly see. More precisely, $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$ is a Banach space if Y is Banach, so in particular X' is always a Banach space.

The next result is about the algebra structure. Let us start by saying that the vector spaces $\mathfrak{L}(\mathsf{X})$ and $\mathfrak{B}(\mathsf{X})$ are closed under composites (since composing preserves continuity). Furthermore, it is immediate that $\mathfrak{L}(\mathsf{X})$, $\mathfrak{B}(\mathsf{X})$ satisfy axioms **A1**, **A2**, **A3** for an algebra whenever the product of two operators is the *composite*. Thus $\mathfrak{L}(\mathsf{X})$ and $\mathfrak{B}(\mathsf{X})$ possess a natural structure of *algebra with unit*, where the unit is the identity map $I : \mathsf{X} \to \mathsf{X}$, and $\mathfrak{B}(\mathsf{X})$ is a subalgebra in $\mathfrak{L}(\mathsf{X})$.

The final part of the theorem is a stronger statement than this, for it says $\mathfrak{B}(\mathsf{X})$ is a *normed algebra with unit* for the operator norm, and a *Banach algebra* if X is Banach.

Theorem 2.41. *Let* X, Y *be normed spaces on* $\mathbb{C}$*, or* $\mathbb{R}$*.*

(a) *The map* $|| \; || : T \mapsto ||T||$*, where* $||T||$ *is as in (2.11), is a norm on* $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$*.*
(b) *On the algebra with unit* $\mathfrak{B}(\mathsf{X})$ *the following properties hold, which turn it into a normed algebra with unit:*

(i) $||TS|| \leq ||T||||S||$, $T, S \in \mathfrak{B}(\mathsf{X})$;
(ii) $||I|| = 1$.

(c) *If* Y *is complete* $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$ *is a Banach space.*
In particular:

(i) *if* X *is a Banach space,* $\mathfrak{B}(\mathsf{X})$ *is a Banach algebra with unit (the identity operator);*
(ii) X' *is always a Banach space with the functionals' norm, even if* X *is not complete.*

Proof. (a) is a direct consequence of the definition of operator norm: properties **N0**, **N1**, **N2**, **N3** can be checked for the operator norm by using them on the norm of Y, together with formula (2.13) and the definition of supremum.
(b) Part (i) is immediate from (2.12) and (2.13). (ii) is straighforward if we use (2.13). Let us see to **(c)**. We claim Y complete implies $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$ Banach. Take a Cauchy sequence $\{T_n\} \subset \mathfrak{B}(\mathsf{X}, \mathsf{Y})$ for the operator norm. By (2.12) we have

$$||T_n u - T_m u|| \leq ||T_n - T_m||||u|| \,.$$

As $\{T_n\}$ is Cauchy, $\{T_n u\}$ is too. Since Y is complete, for any given $u \in \mathsf{X}$ there is a vector in Y:

$$Tu := \lim_{n\to\infty} T_n u \,.$$

$\mathsf{X} \ni u \mapsto Tu$ is a linear operator, because every T_n is. The remains to show $T \in \mathfrak{B}(\mathsf{X}, \mathsf{Y})$ and $||T - T_n|| \to 0$ as $n \to \infty$.

$\{T_n\}$ is a Cauchy sequence, so if $\varepsilon > 0$, then $||T_n - T_m|| \leq \varepsilon$ for n, m sufficiently large, hence $||T_n u - T_m u|| \leq ||T_n - T_m||||u|| \leq \varepsilon ||u||$. Therefore:

$$||Tu - T_m u|| = || \lim_{n\to+\infty} T_n u - T_m u|| = \lim_{n\to+\infty} ||T_n u - T_m u|| \leq \varepsilon ||u||$$

if m is big enough. From this estimate, since $||Tu|| \leq ||Tu - T_m u|| + ||T_m u||$ and by (2.12), we have

$$||Tu|| \leq (\varepsilon + ||T_m||)||u|| \,.$$

This proves T is bounded, so $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$ by Theorem 2.40. Now, since $||Tu - T_m u|| \le \varepsilon ||u||$ we also have $||T - T_m|| \le \varepsilon$ where ε can be arbitrarily small so long as m is large enough. That is to say, $||T - T_n|| \to 0$ as $n \to \infty$.

The proof of subcases (i), (ii) is quick: (i) follows from $\mathfrak{B}(\mathsf{X}) = \mathfrak{B}(\mathsf{X},\mathsf{X})$, while (ii) because $\mathsf{X}' := \mathfrak{B}(\mathsf{X},\mathbb{K})$ and the scalar field $\mathbb{K}$ is a complete normed space. □

As final notion we define *conjugate* or *adjoint operators* in normed spaces. Beware there is a different notion of conjugate operator specific to Hilbert spaces, which we will see in the next chapter.

Take $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$, with X, Y normed. We can build an operator $T' \in \mathfrak{L}(\mathsf{Y}',\mathsf{X}')$ between the dual spaces (swapped), by imposing:

$$(T'f)(x) = f(T(x)) \quad \text{for any } x \in \mathsf{X},\ f \in \mathsf{Y}'.$$

This is well defined, and for every $f \in \mathsf{Y}'$ it produces a function $T'f : \mathsf{X} \to \mathbb{C}$ that is linear by construction, because it coincides with the composite of linear maps f and T. Furthermore, $T' : \mathsf{Y}' \ni f \to T'f \in \mathsf{X}'$ is also linear:

$$\begin{aligned}(T'(af+bf))(x) &= (af+bg)(T(x)) = af(T(x)) + bg(T(x)) \\ &= a(T'f)(x) + b(T'g)(x) \qquad \text{for any } x \in \mathsf{X}.\end{aligned}$$

Eventually, T' is bounded, in the obvious sense:

$$|(T'f)(x)| = |f(T(x))| \le ||f||\,||T||\,||x||\,,$$

and so:

$$||T'f|| = \sup_{||x||=1} |T'f(x)| \le ||f||\,||T||\,.$$

Taking, on the left, the supremum over the collection of $f \in \mathsf{Y}'$ with $||f|| = 1$ gives:

$$||T'|| \le ||T||\,. \tag{2.16}$$

After proving the *Hahn–Banach theorem*, we will show that $||T'|| = ||T||$ if X, Y are Banach spaces.

Definition 2.42. *Let* X,Y *be normed spaces on the same field* $\mathbb{C}$, *or* $\mathbb{R}$, *and* $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$. *The* **conjugate**, *or* **adjoint operator** *to* T, *in the sense of normed spaces, is the operator* $T' \in \mathfrak{B}(\mathsf{Y}',\mathsf{X}')$ *defined by:*

$$(T'f)(x) = f(T(x)) \quad \textit{for any } x \in \mathsf{X},\ f\mathsf{Y}'. \tag{2.17}$$

Remarks 2.43. The map $\mathfrak{B}(\mathsf{X},\mathsf{Y}) \ni T \mapsto T' \in \mathfrak{B}(\mathsf{Y}',\mathsf{X}')$ is linear:

$$(aT+bS)' = aT' + bS' \quad \text{for any } a,b \in \mathbb{C},\ T,S \in \mathfrak{B}(\mathsf{X},\mathsf{Y}).$$ ■

Before passing to the examples, we state an elementary result, very important in the applications, about the uniqueness of extensions of bounded operators and functionals defined, to begin with, on dense subsets.

Proposition 2.44 (Extension of bounded operators). *Let* X,Y *be normed spaces on* $\mathbb{C}$*, or* $\mathbb{R}$*, with* Y *Banach. If* $S \subset \mathsf{X}$ *is a dense subspace of* X *and* $T : S \to \mathsf{Y}$ *is a bounded linear operator on S,*

(a) *there is a unique bounded linear operator* $\tilde{T} : \mathsf{X} \to \mathsf{Y}$ *such that* $\tilde{T}\restriction_S = T$.
(b) $||\tilde{T}|| = ||T||$.

Proof. (a) Given $x \in \mathsf{X}$, there is a sequence $\{x_n\}$ in S converging to x. By hypothesis $||Tx_n - Tx_m|| \leq K||x_n - x_m||$ for $K < +\infty$. Since $x_n \to x$, the sequence of the x_n is Cauchy, and so is Tx_n. Y is complete so there is $\tilde{T}x := \lim_{n\to\infty} Tx_n \in \mathsf{Y}$. The limit depends only on x and not upon the sequence in S used to approximate: if $S \ni z_n \to x$ then by the norms' continuity

$$||\lim_{n\to+\infty} Tx_n - \lim_{n\to+\infty} Tz_n|| = \lim_{n\to+\infty} ||Tx_n - Tz_n|| \leq \lim_{n\to+\infty} K||x_n - z_n|| = K||x-x|| = 0\,.$$

Clearly $\tilde{T}\restriction_S = T$, i.e. $\tilde{T}$ extends T, by choosing for any $x \in S$ the constant sequence $x_n := x$, that tends to x trivially. The linearity of $\tilde{T}$ is straighforward from the definition. Eventually, taking the limit for $n \to +\infty$ of $||Tx_n|| \leq K||x_n||$ gives $||\tilde{T}x|| \leq K||x||$, so $\tilde{T}$ is bounded. Uniqueness: if U is a second bounded extension T on X, then for any $x \in \mathsf{X}$, $\tilde{T}x - Ux = \lim_{n\to+\infty}(\tilde{T}x_n - Ux_n)$ by continuity, where the x_n belong to S (dense in X). As $\tilde{T}\restriction_S = T = U\restriction_S$, the limit is trivial and gives $\tilde{T}x = Ux$ for all $x \in \mathsf{X}$, i.e. $\tilde{T} = U$.
(b) Let $x \in \mathsf{X}$ and $\{x_n\} \subset S$ converge to x: then

$$||\tilde{T}x|| = \lim_{n\to+\infty} ||Tx_n|| \leq \lim_{n\to+\infty} ||T||||x_n|| = ||T||||x||\,,$$

so $||\tilde{T}|| \leq ||T||$. But since $S \subset \mathsf{X}$ and $\tilde{T}\restriction_S = T$,

$$||\tilde{T}|| = \sup\left\{ \frac{||\tilde{T}x||}{||x||} \,\middle|\, 0 \neq x \in \mathsf{X} \right\} \geq \sup\left\{ \frac{||\tilde{T}x||}{||x||} \,\middle|\, 0 \neq x \in S \right\}$$
$$= \sup\left\{ \frac{||Tx||}{||x||} \,\middle|\, 0 \neq x \in S \right\}\,.$$

The last limit in the chain is $||T||$, hence $||\tilde{T}|| \geq ||T||$, and thus $||\tilde{T}|| = ||T||$. □

Examples 2.45. (1) *Complex measures* (see Section 1.4.8) allow to construct every bounded linear functional on $C_0(\mathsf{X})$, where X is locally compact and Hausdorff. Consider a locally compact Hausdorff space X equipped with a complex measure μ defined on the Borel σ-algebra of X. We know that the normed algebra $(C_c(\mathsf{X}), ||\ ||_\infty)$ completes to the Banach algebra $(C_0(\mathsf{X}), ||\ ||_\infty)$ of maps that vanish at infinity (Example 2.26(4)). Under the assumptions made $||\mu|| := |\mu|(\mathsf{X})$, where the positive, σ-additive and finite measure $|\mu|$ is the *total variation* of μ (cf. Section 1.4.8). Easily then, $||\ ||$ is a norm on the space of complex Borel measures on X. Moreover, if $f \in C_0(\mathsf{X})$,

$$|\Lambda_\mu f| \leq ||\mu||||f||_\infty \quad \text{where } \Lambda_\mu f := \int_{\mathsf{X}} f d\mu\,,$$

and, as usual, $||f||_\infty = \sup_{x\in\mathsf{X}} |f(x)|$.

Consequently, every complex Borel measure μ defines an element Λ_μ in the (topological) dual of $C_0(\mathsf{X})$. *Riesz's theorem for complex measures* [Rud82] proves this is a general fact, and even more.

To state it, we recall that a complex Borel measure μ is called *regular* if the finite positive Borel measure given by the total variation $|\mu|$ is regular (Definition 1.65).

Theorem 2.46 (Riesz's theorem for complex measures). *Let X be a locally compact Hausdorff space, and $\Lambda : C_0(\mathsf{X}) \to \mathbb{C}$ a continuous linear functional. Then there exists a unique regular complex Borel measure μ_Λ such that, for every $f \in C_0(\mathsf{X})$:*

$$\Lambda(f) = \int_{\mathsf{X}} f d\mu_\Lambda \,.$$

Moreover, $||\Lambda|| = ||\mu_\Lambda||$.

Since every regular complex Borel measure determines a bounded functional on $C_0(\mathsf{X})$ by integration, the theorem has a corollary.

Corollary 2.47. *If X is locally compact and Hausdorff, the topological dual $C_0(\mathsf{X})'$ of the Banach space $(C_0(\mathsf{X}), ||\ ||_\infty)$ is identified with the real vector space of regular complex Borel measures μ on X, endowed with norm $||\mu|| := |\mu|(\mathsf{X})$. The function mapping μ to the functional $\Lambda_\mu : C_0(\mathsf{X}) \to \mathbb{R}$, with $\Lambda_\mu f := \int_{\mathsf{X}} f d\mu$, is an isomorphism of normed spaces.*

Also note, $C_c(\mathsf{X})$ being dense in $C_0(\mathsf{X})$, that a continuous functional on the former space determines a unique functional on the latter, so the theorem characterises as well continuous functionals on $C_c(\mathsf{X})$ for the sup norm.

Further, if in X every open set is the countable union of compact sets (as in $\mathbb{R}^n$, where each open set is the union of countably many closed balls of finite radius), the word *regular* can be dropped in statement 2.46, by way of Proposition 1.60, because compact sets have finite measure in the finite $|\mu|$. In particular we have:

Theorem 2.48 (Riesz's theorem for complex measures on $\mathbb{R}^n$). *Let $K \subset \mathbb{R}^n$, or $K \subset \mathbb{C}$, be a compact set and $\Lambda : C_0(K) \to \mathbb{C}$ a continuous linear functional. Then there is a unique complex Borel measure μ_Λ on K such that*

$$\Lambda(f) = \int_K f d\mu_\Lambda$$

for any $f \in C_0(K)$. Additionally, μ_Λ is regular.

(2) Another nice class of duals to Banach spaces is that of L^p spaces, cf. Example 2.26(6). In this respect [Rud82],

Proposition 2.49. *Let (X, Σ, μ) be a positive measure space. If $1 \le p < +\infty$ the dual to the Banach space $L^p(\mathsf{X}, \mu)$ is $L^q(\mathsf{X}, \mu)$, with $1/p + 1/q = 1$, in the sense that the linear map:*

$$L^q(\mathsf{X}, \mu) \ni [g] \mapsto \Lambda_g \quad \text{where } \Lambda_g(f) := \textstyle\int_{\mathsf{X}} fg\, d\mu \ , \ f \in L^p(\mathsf{X}, \mu)$$

is an isomorphism of the normed spaces $L^q(\mathsf{X}, \mu)$, $(L^p(\mathsf{X}, \mu))'$.

In the same way the dual to $L^1(\mathsf{X},\mu)$ *is identified with* $L^\infty(\mathsf{X},\mu)$*, because the linear map*

$$L^\infty(\mathsf{X},\mu) \ni [g] \mapsto \Lambda'_g \quad \textit{where } \Lambda'_g(f) := \int_{\mathsf{X}} fg\,d\mu\ ,\ f \in L^1(\mathsf{X},\mu)$$

is an isomorphism of the normed spaces $L^\infty(\mathsf{X},\mu)$, $(L^1(\mathsf{X},\mu))'$. ∎

2.3 The fundamental theorems of Banach spaces

This section is devoted the foremost theorems on normed and Banach spaces in their most elementary versions, and we will study their main consequences. These are the theorems of Hahn–Banach, Banach–Steinhaus and the open mapping theorem.

The applications of the theorem of Banach–Steinhaus call forth several kinds of topologies, which play a major role in QM when the domain space is the Hilbert space of the theory, bounded operators are (certain) observables, and the basic features of the quantum system associated to measurement processes are a subclass of orthogonal projectors. In order to pass with continuity from the algebra of observables to that of projectors we need weaker topologies than the standard one. This sort of issues, that we shall discuss later, lead to the notion of *von Neumann algebra (of operators)*.

2.3.1 The Hahn–Banach theorem and its immediate consequences

The first result we present is the celebrated Hahn–Banach theorem, that deals with extending a continuous linear functional from a subspace to the ambient space in a continuous and norm-preserving way. More elaborated and stronger versions can be found in [Rud91]. We shall restrict to the simplest situation possible.

First of all, we remark that if X is normed and $\mathsf{M} \subset \mathsf{X}$ is a subspace, the norm of X restricted to M defines a normed space. In this sense we can talk of continuous operators and functionals on M, meaning they are bounded for the induced normed.

Theorem 2.50 (Hahn–Banach theorem for normed spaces). *Let* M *be a subspace (not necessarily closed) in a normed space* X *over* $\mathbb{K} = \mathbb{C}$ *or* $\mathbb{R}$.
If $g : \mathsf{M} \to \mathbb{K}$ *is a continuous linear functional, there exists a continuous linear functional* $f : \mathsf{X} \to \mathbb{K}$ *such that* $f \restriction_M = g$ *and* $||f||_{\mathsf{X}} = ||g||_{\mathsf{M}}$.

Proof. We shall follow the proof of [Rud82]. Start with $\mathbb{K} = \mathbb{R}$. If $g = 0$, an extension as required is $f = 0$. So let us suppose $g \neq 0$ and without loss of generality set $||g|| = 1$. Let us build the extension f. Take $x_0 \in \mathsf{X} \setminus \mathsf{M}$ and call

$$\mathsf{M}_1 := \{x + \lambda x_0 \,|\, x \in \mathsf{M}\,, \lambda \in \mathbb{R}\}\,.$$

If we set $g_1 : \mathsf{M}_1 \to \mathbb{R}$ to be

$$g_1(x + \lambda x_0) = g(x) + \lambda v$$

for any given $v \in \mathbb{R}$, we obtain an extension of g to M_1. We claim v can be taken so that $||g_1|| = 1$. For this it suffices to have v such that:

$$|g(x) + \lambda v| \leq ||x + \lambda x_0|| , \quad \text{for any } x \in \mathsf{M} \text{ and } \lambda \in \mathbb{R} \setminus \{0\}. \tag{2.18}$$

Substitute $-\lambda x$ to x and divide (2.18) by $|\lambda|$, obtaining the equivalent relation:

$$|g(x) - v| \leq ||x - x_0|| , \quad \text{for any } x \in \mathsf{M}. \tag{2.19}$$

Set

$$a_x := g(x) - ||x - x_0|| \quad \text{and} \quad b_x := g(x) + ||x - x_0||. \tag{2.20}$$

Inequality (2.19), hence $||g_1|| = 1$, holds if v satisfies $a_x \leq v \leq b_x$ for all $x \in \mathsf{M}$. So it is enough to prove that the intervals $[a_x, b_x]$, $x \in \mathsf{M}$, have a common point; in other words, that for all $x, y \in \mathsf{M}$:

$$a_x \leq b_y \,. \tag{2.21}$$

But:

$$g(x) - g(y) = g(x - y) \leq ||x - y|| \leq ||x - x_0|| + ||y - x_0||$$

and (2.21) follows from (2.20). Hence, we managed to fix v so that $||g_1|| = 1$.

Now consider the family $\mathscr{P}$ of pairs (M', g') where $\mathsf{M}' \supset \mathsf{M}$ is a subspace in X and $g' : \mathsf{M}' \to \mathbb{R}$ is a linear extension of g with $||g'|| = 1$. $\mathscr{P}$ is not empty since (M_1, g_1) belongs in it. We can define a partial order on $\mathscr{P}$ (see Appendix A, also for the sequel) by setting $(\mathsf{M}', g') \leq (\mathsf{M}'', g'')$ if $\mathsf{M}'' \supset \mathsf{M}'$, g'' extends g' and $||g'|| = ||g''|| = 1$. It is easy to show that any totally ordered subset of $\mathscr{P}$ admits an upper bound in $\mathscr{P}$. Then Zorn's lemma detects a maximal element in $\mathscr{P}$, say (M^1, f^1). Now we must have $\mathsf{M}^1 = \mathsf{X}$, for otherwise there would be $x_0 \in \mathsf{X} \setminus \mathsf{M}^1$, and using the initial argument we could construct a non-trivial, norm-preserving extension f^1 to the subspace generated by x_0 and M^1, but this would contradict maximality. Therefore $f := f^1$ is the required extension.

Before passing to the case $\mathbb{K} = \mathbb{C}$ we need a lemma.

Lemma 2.51. *On a complex vector space* Y*:*

(a) *if $u(x) = Re\, g(x)$ for all $x \in \mathsf{Y}$ for some complex linear functional $g : \mathsf{Y} \to \mathbb{C}$, the map $u : \mathsf{Y} \to \mathbb{R}$ is a real linear functional on Y, and:*

$$g(x) := u(x) - iu(ix) \quad \textit{for any } x \in \mathsf{Y}. \tag{2.22}$$

(b) *If $u : \mathsf{Y} \to \mathbb{R}$ is a real linear functional on Y and g is defined by (2.22), then g is a complex linear functional on Y.*
(c) *If Y is normed and g, u are related by (2.22), then $||g|| = ||u||$.*

Proof. (a) are (b) are proved simultaneously by direct computation. As for (c), under the assumptions made: $|u(x)| \leq |g(x)| = \sqrt{|u(x)|^2 + |u(ix)|^2}$, so $||u|| \leq ||g||$. On the other hand taking $x \in \mathsf{Y}$, there is $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that $\alpha g(x) = |g(x)|$. Consequently $|g(x)| = g(\alpha x) = u(\alpha x) \leq ||u|| \, ||\alpha x|| = ||u|| \, ||x||$ and $||g|| \leq ||u||$. □

Now back to the main proof. If $u : \mathsf{M} \to \mathbb{R}$ is the real part of g, then $g(x) = u(x) - iu(ix)$ and $||g|| = ||u||$ by the lemma. From the real case seen before we know there exists a linear extension $U : \mathsf{X} \to \mathbb{R}$ of u with $||U|| = ||u|| = ||g||$. Therefore if we put

$$f(x) := U(x) - iU(ix)\,, \quad \text{for any } x \in \mathsf{X},$$

$f : \mathsf{X} \to \mathbb{C}$ extends g to X, and $||f|| = ||U|| = ||g||$. □

Here is one of the most useful corollaries to Hahn-Banach. We remind that the topological dual $\mathfrak{B}(\mathsf{X},\mathbb{C})$ of a normed space is indicated by X'.

Corollary 2.52 (to the Hahn–Banach theorem). *Let* X *be a normed space over* $\mathbb{K} = \mathbb{C}$*, or* $\mathbb{R}$*, and fix* $x_0 \in \mathsf{X}$*,* $x_0 \neq 0$*. Then there is* $f \in \mathsf{X}'$*,* $||f|| = 1$*, such that* $f(x_0) = ||x_0||$*.*

Proof. Choose $\mathsf{M} := \{\lambda x_0 \mid \lambda \in \mathbb{K}\}$ and $g : \lambda x_0 \to \lambda ||x_0||$. Let $f \in \mathsf{X}'$ denote the bounded functional extending g according to Hahn-Banach. By construction $f(x_0) = g(x_0) = ||x_0||$ and $||f||_{\mathsf{X}} = ||g||_{\mathsf{M}} = 1$. □

An immediate consequence of this is a statement about the norm of the conjugate operator $T' \in \mathfrak{B}(\mathsf{Y}',\mathsf{X}')$ to $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$ (cf. Definition 2.42).

Proposition 2.53. *If* $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$*, with* X*,* Y *normed over* $\mathbb{C}$ *or* $\mathbb{R}$*, then:*

$$||T'|| = ||T||\,.$$

Proof. In general we have (cf. (2.16)) $||T|| \geq ||T'||$, so we need only prove $||T|| \leq ||T'||$. Take $x \in \mathsf{X}$ and $Tx \neq 0$, define $y_0 := \frac{Tx}{||Tx||} \in \mathsf{Y}$. Clearly $||y_0|| = 1$, and by Corollary 2.52 there is $g \in \mathsf{Y}'$ such that $||g|| = 1$, $g(y_0) = 1$ hence $g(Tx) = ||Tx||$. But:

$$||Tx|| = g(Tx) = |(T'g)(x))| \leq ||T'g||\,||x|| \leq ||T'||\,||g||\,||x|| = ||T'||\,||x||\,,$$

so eventually $||T|| \leq ||T'||$ as required. □

Another fact, with important consequences for Banach algebras, is this.

Corollary 2.54 (to the Hahn-Banach theorem). *Let* $\mathsf{X} \neq \{0\}$ *be a normed space over* $\mathbb{C}$ *or* $\mathbb{R}$*.*
Then the elements of X' *separate* X*, i.e. for any* $x_1 \neq x_2$ *in* X *there is* $f \in \mathsf{X}'$ *for which* $f(x_1) \neq f(x_2)$*.*

Proof. It suffices to have $x_0 := x_1 - x_2$ in Corollary 2.52, for then $f(x_1) - f(x_2) = f(x_1 - x_2) = ||x_1 - x_2|| \neq 0$. □

Take $x \in \mathsf{X}$ and $f \in \mathsf{X}'$, $||f|| = 1$; then $|f(x)| \leq 1||x||$ and

$$\sup\{|f(x)| \mid f \in \mathsf{X}'\,,\ ||f|| = 1\} \leq ||x||\,.$$

Corollary 2.52 allows to strengthen this fact by showing

$$\sup\{|f(x)| \mid f \in \mathsf{X}'\,,\ ||f|| = 1\} = \max\{|f(x)| \mid f \in \mathsf{X}'\,,\ ||f|| = 1\} = ||x||$$

directly. This seems not so striking, but has a certain weight when comparing infinite-dimensional to finite-dimensional normed spaces.

From the elementary theory of *finite-dimensional* vector spaces X, the algebraic dual of the algebraic dual $(\mathsf{X}^*)^*$ has the nice property of being *naturally isomorphic* to X. The isomorphism is the linear function mapping $x \in \mathsf{X}$ to the linear functional $\mathfrak{J}(x)$ on X^* defined by $(\mathfrak{J}(x))(f) := f(x)$ for all $f \in \mathsf{X}^*$.

In infinite dimensions $\mathfrak{J}$ identifies X to a subspace of $(\mathsf{X}^*)^*$ only, not the whole $(\mathsf{X}^*)^*$ in general. Is there an alike general statement about *topological duals* to infinite-dimensional normed spaces?

Note $(\mathsf{X}')'$ is the dual to a normed space (X' with the operator norm). Consequently $(\mathsf{X}')'$ is a normed space, still with operator norm.

Go back to the natural linear transformation $\mathfrak{J} : \mathsf{X} \to (\mathsf{X}')^*$ mapping $x \in \mathsf{X}$ to $\mathfrak{J}(x) \in (\mathsf{X}')^*$, the linear function $\mathfrak{J}(x) : \mathsf{X}' \to \mathbb{K}$ defined by

$$(\mathfrak{J}(x))(f) := f(x) \quad \text{for any } f \in \mathsf{X}' \text{ and } x \in \mathsf{X}\,.$$

This is well defined, for $\mathfrak{J}(x)$ is a linear functional on X' for which $\mathfrak{J}(x) \in (\mathsf{X}')^*$. Now

$$\sup\{|f(x)| \mid f \in \mathsf{X}' \,,\ ||f|| = 1\} = ||x||$$

implies: (1) $\mathfrak{J}(x)$ is a *bounded* functional, so it belongs to $(\mathsf{X}')'$, and (2) $||\mathfrak{J}(x)|| = ||x||$. Therefore the linear mapping $\mathfrak{J} : \mathsf{X} \to (\mathsf{X}')'$ is an *isometry*, in particular *injective*. This gives an isometric inclusion $\mathsf{X} \subset (\mathsf{X}')'$ under the linear map $\mathfrak{J} : \mathsf{X} \to (\mathsf{X}')'$. Overall we have proved:

Corollary 2.55 (to the Hahn–Banach theorem). *Let X be a normed space over $\mathbb{C}$ or $\mathbb{R}$. The linear map $\mathfrak{J} : \mathsf{X} \to (\mathsf{X}')'$:*

$$(\mathfrak{J}(x))(f) := f(x) \quad \textit{for any } x \in \mathsf{X} \textit{ and } f \in \mathsf{X}', \tag{2.23}$$

is an isometry, and X is thus identified isometrically to a subspace of $(\mathsf{X}')'$.

There are infinite-dimensional examples where X does not fill $(\mathsf{X}')'$, and these justify the next notion.

Definition 2.56. *A normed space X on $\mathbb{C}$ or $\mathbb{R}$ is* **reflexive** *if the isometry (2.23) is onto (an isomorphism of normed spaces).*

Otherwise said, X is reflexive when X and $(\mathsf{X}')'$ are isometrically isomorphic under the natual map $\mathfrak{J}$. In Chapter 3 we will show that Hilbert spaces are reflexive.

Example 2.57. The Banach spaces $L^p(\mathsf{X},\mu)$ of Examples 2.26 are reflexive for $1 < p < \infty$. The proof is straightforward: $L^p(\mathsf{X},\mu)' = L^q(\mathsf{X},\mu)$ for $1/p + 1/q = 1$, and swapping q with p gives $L^q(\mathsf{X},\mu)' = L^p(\mathsf{X},\mu)$. Hence: $(L^p(\mathsf{X},\mu)')' = L^p(\mathsf{X},\mu)$. ■

2.3.2 The Banach–Steinhaus theorem or uniform boundedness principle

Let us go to the second core result, the theorem of Banach–Steinhaus, and present the simplest formulation and consequences. It is also known as *uniform boundedness principle*, because it essentially and remarkably states that pointwise equi-boundedness implies uniform boundedness for families of operators on a Banach space.

Theorem 2.58 (Banach–Steinhaus). *Let* X *be a Banach space,* Y *a normed space defined on the same field* $\mathbb{C}$ *or* $\mathbb{R}$*. If* $\{T_\alpha\}_{\alpha\in A} \subset \mathfrak{B}(\mathsf{X},\mathsf{Y})$ *is a family of operators such that:*

$$\sup_{\alpha\in A} ||T_\alpha x|| < +\infty \quad \textit{for any } x \in \mathsf{X}\,,$$

then there is $K \geq 0$ *that bounds uniformly the family:*

$$||T_\alpha|| \leq K \ \textit{ for any } \alpha \in A\,.$$

Proof. The proof relies on finding an open ball $B_\rho(z) \subset \mathsf{X}$ for which there is $M \geq 0$ with $||T_\alpha(x)|| \leq M$ for all $\alpha \in A$ and any $x \in B_\rho(z)$. In fact, since $x = (x+z) - z$, we would have:

$$||T_\alpha(x)|| \leq ||T_\alpha(x+z)|| + ||T_\alpha(z)|| \leq 2M\,, \quad \text{for any } \alpha \in A,\, x \in B_\rho(0)\,,$$

so $||T_\alpha|| \leq 2M/\rho$ for all $\alpha \in A$, and the claim would follow.

We shall prove that $B_\rho(z)$ and M exist by contradiction. If such a ball did not exist, for some arbitrary open $B_{r_0}(x_0)$, there would be $x_1 \in B_{r_0}(x_0)$ for which $||T_{\alpha_1}(x_1)|| > 1$, with $\alpha_1 \in A$. As T_{α_1} is continuous, we could find a second open ball $B_{r_1}(x_1)$ with $\overline{B_{r_1}(x_1)} \subset B_{r_0}(x_0)$ and $0 < r_1 < r_0$ such that $||T_{\alpha_1}(x)|| \geq 1$, provided $x \in \overline{B_{\alpha_1}(x_1)}$. Now this recipe could be iterated to give rise to a sequence of open balls in X, $\{B_{r_k}(x_k)\}_{k\in\mathbb{N}}$, satisfying:

(i) $B_{r_k}(x_k) \supset \overline{B_{r_{k+1}}(x_{k+1})}$;
(ii) $r_k \longrightarrow 0$ as $k \longrightarrow +\infty$;
(iii) for every $k \in \mathbb{N}$ there is $\alpha_k \in A$ such that $||T_{\alpha_k}(x)|| \geq k$ if $x \in \overline{B_{r_k}(x_k)}$.

(i) and (ii) imply the sequence $\{x_k\}_{k\in\mathbb{N}}$ is Cauchy, so it converges to some $y \in \mathsf{X}$ by completeness, and by construction $y \in \cap_{k\in\mathbb{N}}\overline{B_{r_k}(x_k)}$. But (iii) tells $||T_{\alpha_k}(y)|| \geq k$ for all $k \in \mathbb{N}$, contradicting the assumption that $\sup_{n\in\mathbb{N}} ||T_\alpha x|| < +\infty$ for any $x \in \mathsf{X}$. □

Here is a straightforward and useful corollary.

Corollary 2.59 (to the Banach–Steinhaus theorem). *Under the assumptions of the of Banach–Steinhaus theorem the family of operators* $\{T_\alpha\}_{\alpha\in A}$ *is* **equicontinuous***: given any* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that* $||x - x'|| < \delta$ *for* $x, x' \in \mathsf{X}$ *implies* $||T_\alpha x - T_\alpha x'|| < \varepsilon$ *for any* $\alpha \in A$.

Proof. Set $C_\gamma := \{x \in \mathsf{X} \mid ||x|| \leq \gamma\}$ for all $\gamma > 0$. Fix $\varepsilon > 0$, so we must find the $\delta > 0$ of the conclusion. By Banach–Steinhaus and Proposition 2.39, $||T_\alpha x|| \leq K < +\infty$ for any $\alpha \in A$ and $x \in C_1$. If $K = 0$ there is nothing to prove, so assume $K > 0$. Choose $\delta > 0$ for which $C_\delta \subset C_{\varepsilon/K}$. Then if $||x - x'|| < \delta$, we have $K(x-x')/\varepsilon \in C_{K\delta/\varepsilon} \subset C_1$ and so:

$$||T_\alpha x - T_\alpha x'|| = ||T_\alpha(x-x')|| = \frac{\varepsilon}{K}\left|\left|T_\alpha \frac{K(x-x')}{\varepsilon}\right|\right| < \frac{\varepsilon}{K}K = \varepsilon \quad \text{for any } \alpha \in A. \ \square$$

And here is another consequence about topological duals.

Corollary 2.60 (to the Banach–Steinhaus theorem). *Let* X *be a normed space over* $\mathbb{C}$ *or* $\mathbb{R}$*. If* $S \subset \mathsf{X}$ *is weakly bounded, i.e.*

$$\textit{for any } f \in \mathsf{X}' \textit{ there exists } c_f \geq 0 \textit{ such that } |f(x)| \leq c_f \textit{ for all } x \in S,$$

then S *is bounded for the norm of* X.

Proof. Consider the elements $x \in S \subset \mathsf{X}$ as functionals on the dual $(\mathsf{X}')'$ to X' (using the isometry $\mathfrak{I} : \mathsf{X} \to (\mathsf{X}')'$ of Corollary 2.55). The family $S \subset (\mathsf{X}')'$ of functionals on X' is bounded on every $f \in \mathsf{X}'$, since by assumption $|x(f)| = |f(x)| \leq c_f$ (we have written x for $\mathfrak{I}(x)$). Since X' is complete the theorem of Banach–Steinhaus guarantees $\sup\{|x(f)| \mid ||f|| = 1\} \leq K < +\infty$ for all $x \in S$, i.e. ($\mathfrak{I}$ is an isometry) $||x|| \leq K < +\infty$ for all $x \in S$. □

2.3.3 Weak topologies. *-weak completeness of X'

To state the last corollary to Banach–Steinhaus we need to introduce a new section on general topology and apply it to the operator spaces encountered so far. This will allow to see notions, useful for the applications, on types of convergence for sequences of operators, which in turn will help us prove a simple and useful result known as the *Banach–Alaoglu theorem.*

We begin with basic facts about *convexity*.

Definition 2.61. *A subset* $\varnothing \neq K \subset \mathsf{X}$ *in a vector space is* **convex** *when:*

$$\lambda x + (1-\lambda)y \in K \quad \textit{for any } \lambda \in [0,1] \textit{ and } x, y \in K.$$

A point $e \in K$ *is* **extremal** *if it cannot be written as:*

$$e = \lambda x + (1-\lambda)y \quad \textit{for some } \lambda \in (0,1) \textit{ and } x, y \in K \setminus \{e\}.$$

It should be clear that the intersection of convex sets is convex, because the segment joining two points in the intersection belongs to all. So we are lead to the notion of *convex hull.*

Definition 2.62. *The* **convex hull** *of a subset* E *in a vector space* X *is the convex set*

$$co(E) := \bigcap \{K \supset E \,|\, K \subset \mathsf{X}, K \textit{ convex}\}.$$

Let us go back to the definition of an open ball via a seminorm.

Notation 2.63. Take $\delta > 0$ and a seminorm p on the vector space X over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, and a point $x \in \mathsf{X}$. We denote by $B_{p,\delta}(x)$ the **open ball associated to the seminorm** p, centred at x and of radius δ:

$$B_{p,\delta}(x) := \{z \in \mathsf{X} \mid p(z-x) < \delta\}.$$

If $x = 0$ we will just write $B_{p,\delta}$, not $B_{p,\delta}(0)$.

If $A \subset \mathsf{X}$, $B \subset \mathsf{X}$, $x \in \mathsf{X}$ and $\alpha, \beta \in \mathbb{K}$, we will also abbreviate:

$$x + \beta A := \{x + \beta u \mid u \in A\} \quad \text{and} \quad \alpha A + \beta B := \{\alpha u + \beta v \mid u \in A, v \in B\}. \quad ■$$

Immediately, then, the balls $B_{p,\delta}$, $\delta > 0$, are:

(i) **convex**, since $x, y \in B_{p,\delta}$ implies trivially $(1-\lambda)x + \lambda y \in B_{p,\delta}$ with $\lambda \in [0,1]$;
(ii) **balanced**, i.e. $\lambda x \in B_{p,\delta}$ if $x \in B_{p,\delta}$ and $0 \leq \lambda \leq 1$;
(iii) **absorbing**, i.e. $x \in \mathsf{X}$ implies $\lambda^{-1}x \in B_{p,\delta}$ for some $\lambda > 0$.

All these properties are patently invariant under intersection; hence also intersections of balls *centred at the origin* but defined with different seminorms enjoy the property.

Definition 2.64. *Let $\mathscr{P} := \{p_i\}_{i\in I}$ be a family of seminorms on the vector space X over $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$. The* **topology** $\mathscr{T}(\mathscr{P})$ **on X generated, or induced, by** $\mathscr{P}$*, is the unique one admitting as* **basis** *(Definition 1.1) the collection:*

$$x + \left(B_{p_{i_1},\delta_1} \cap \cdots \cap B_{p_{i_n},\delta_n}\right) \tag{2.24}$$

for any choice of: centres $x \in \mathsf{X}$, numbers $n = 1, 2, \ldots$, indices $i_1, \ldots, i_n \in I$ and radii $\delta_1 > 0, \ldots \delta_n > 0$. The pair $(\mathsf{X}, \mathscr{P})$, where X is simultaneously a vector space with topology induced by the seminorms $\mathscr{P}$ and a topological space, is called **locally convex space**.

Let us put it differently: the topology on $\mathscr{T}(\mathscr{P})$ has as open sets $\varnothing$ and all possible unions of sets (2.24), with any centre $x \in \mathsf{X}$, for any $n = 1, 2, \ldots$, any index $i_1, \ldots, i_n \in I$ and any $\delta_1 > 0, \ldots \delta_n > 0$.

Remarks 2.65. If $\mathscr{P}$ reduces to one single *norm*, the induced topology is the usual one induced by a norm discussed at the beginning of the chapter. If this sole element is a seminorm, we still have a topology, with the crucial difference that the Hausdorff property might be no longer valid. ■

Since adding vectors and multiplying a vector by a scalar are continuous operations in any seminorm (the proof is the same we gave for a norm), they are continuous in the topology generated by a family $\mathscr{P}$ of seminorms as well. This means the vector space structure is **compatible** with the topology generated by $\mathscr{P}$. A vector space with a compatible topology as above is a **topological vector space**. A locally convex space is thus a topological vector space.

Keeping Definition 1.13 in mind we can prove the next fact without effort.

Proposition 2.66. *A sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ converges to $x_0 \in \mathsf{X}$ in the topology $\mathscr{T}(\mathscr{P})$ if and only if $p_i(x_n - x_0) \to 0$, for all $p_i \in \mathscr{P}$, as $n \to +\infty$.*

Our first example of topology induced by seminorms arises from the dual X' of a normed space.

Definition 2.67. *If X is a normed space, the* **weak topology** *on X is the topology induced by the collection of seminorms p_f on X:*

$$p_f(x) := |f(x)| \quad \textit{with } x \in \mathsf{X}$$

for $f \in \mathsf{X}'$.

Consider pairs of normed spaces and the corresponding sets of operators between them; using the topology induced by seminorms we can define certain "standard" topologies on the vector spaces $\mathfrak{L}(\mathsf{X},\mathsf{Y})$, $\mathfrak{B}(\mathsf{X},\mathsf{Y})$ and the dual X', thus making them locally convex topological vector spaces. One such topology (and the corresponding dual one) is already known to us, namely the topology induced by the operator norm.

Definition 2.68. *Let* X,Y *be normed spaces over* $\mathbb{K}=\mathbb{C}$ *or* $\mathbb{R}$.

(a) *Define on* $\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(respectively* $\mathfrak{B}(\mathsf{X},\mathsf{Y})$*) the following operator topologies.*

(i) *The topology induced on* $\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(*$\mathfrak{B}(\mathsf{X},\mathsf{Y})$*) by the family of seminorms* $p_{x,f}$*:*

$$p_{x,f}(T) := |f(T(x))| \quad \textit{with } T \in \mathfrak{L}(\mathsf{X},\mathsf{Y})\ (\mathfrak{B}(\mathsf{X},\mathsf{Y}),$$

for given $x \in \mathsf{X}$ *and* $f \in \mathsf{Y}'$*, is called* **weak topology** *on* $\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(*$\mathfrak{B}(\mathsf{X},\mathsf{Y})$*);*

(ii) *The topology induced on* $\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(*$\mathfrak{B}(\mathsf{X},\mathsf{Y})$*) by the seminorms* p_x*:*

$$p_x(T) := ||T(x)||_{\mathsf{Y}} \quad \textit{with } T \in \mathfrak{L}(\mathsf{X},\mathsf{Y})\ (\mathfrak{B}(\mathsf{X},\mathsf{Y})),$$

for given $x \in \mathsf{X}$*, is the* **strong topology** *on* $\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(*$\mathfrak{B}(\mathsf{X},\mathsf{Y})$*);*

(iii) *The topology induced on* $\mathfrak{B}(\mathsf{X},\mathsf{Y})$ *by the operator norm (2.11) is the* **uniform topology** *on* $\mathfrak{B}(\mathsf{X},\mathsf{Y})$.

(b) *In case* $\mathsf{Y}=\mathbb{K}$ *(we are talking about* X'*) the uniform topology of (iii) goes under the name of* **(dual) strong topology** *of* X'*, and the topologies of (i) and (ii), now coinciding, are called* ***-weak topology** *of* X'*. The* **-weak topology on* X' *is thus induced by the seminorms* p^*_x*:*

$$p^*_x(f) := |f(x)| \quad \textit{with } f \in \mathsf{X}'$$

for a given $x \in \mathsf{X}$.

Remark 2.69. (1) It is not hard to see that the open sets, in a normed space, of the weak topology are also open for the standard topology, not the opposite. Likewise, in $\mathfrak{L}(\mathsf{X},\mathsf{Y})$, open sets in the weak topology are open for the strong topology but not conversely. We can rephrase this better by saying that *the standard topology on* X *and the strong topology on* $\mathfrak{L}(\mathsf{X},\mathsf{Y})$ *are* **finer** *than the corresponding weak topologies.*

In the same way, when talking of operator spaces it is not hard to show that *the uniform topology is finer than the strong topology.*

For dual spaces an analogous property obviously holds: *the strong topology is finer than the* **-weak topology.*

(2) From Proposition 2.66 these consequences descend immediately.

Proposition 2.70. *Take* $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ *with* X *normed. Then* $x_n \to x \in \mathsf{X}$*,* $n \to +\infty$*, in the weak topology if and only if:*

$$f(x_n) \to f(x)\,, \quad \textit{as } n \to +\infty, \textit{for any } f \in \mathsf{X}'.$$

Proposition 2.71. *If* $\{T_n\}_{n\in\mathbb{N}} \subset \mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(o* $\mathfrak{B}(\mathsf{X},\mathsf{Y})$*) and* $T \in \mathfrak{L}(\mathsf{X},\mathsf{Y})$ *(resp.* $\mathfrak{B}(\mathsf{X},\mathsf{Y})$*), then* $T_n \to T$*,* $n \to +\infty$*, in weak topology if and only if:*

$$f(T_n(x)) \to f(T(x))\,, \quad \textit{as } n \to +\infty, \textit{for any } x \in \mathsf{X},\ f \in \mathsf{Y}'.$$

Proposition 2.72. *$T_n \to T$, $n \to +\infty$, in the strong topology if and only if:*

$$||T_n(x) - T(x)||_{\mathsf{Y}} \to 0\,, \quad \text{as } n \to +\infty, \text{ for any } x \in \mathsf{X}.$$

Now it is clear that:

(a) *Convergence of a sequence in a normed space X in the standard sense (for the norm topology) implies weak convergence (convergence in the weak topology).*
(b) *Uniform convergence of a sequence of operators in $\mathfrak{B}(\mathsf{X},\mathsf{Y})$ (in the uniform topology) implies strong convergence (for the strong topology).*
(c) *Strong convergence of a sequence of operators in $\mathfrak{L}(\mathsf{X},,\mathsf{Y})$ or $\mathfrak{B}(\mathsf{X},\mathsf{Y})$ implies weak convergence.*

(3) Proposition 2.66 also gives:

Proposition 2.73. *Let $\{f_n\}_{n\in\mathbb{N}} \subset \mathsf{X}'$ be a sequence of functionals and take a functional $f \in \mathsf{X}'$. Then $f_n \to f$, $n \to +\infty$, in the *-weak topology if and only if:*

$$f_n(x) \to f(x)\,, \quad \text{as } n \to +\infty \text{ for any chosen } x \in \mathsf{X}.$$

Now we know that *the strong convergence of a sequence of functionals of X' (for the dual strong topology) implies *-weak convergence.*

(4) We can put on X' yet a further weak topology, by viewing X' as acted upon by $(\mathsf{X}')'$. The seminorms inducing the topology are

$$p_s(f) := |s(f)|$$

for any $s \in (\mathsf{X}')'$. If X is not reflexive, this weak topology does *not* coincide, in general, with the *-weak topology seen above, because X is identified to a proper subspace in $(\mathsf{X}')'$, and so the seminorms of the *-weak topology are fewer than the weak topology ones. The weak topology is finer than the *-weak one: a weakly open set is *-weakly open, but the converse may not hold. Analogously, weak convergence of sequences in X' implies *-weak convergence, not the opposite. ■

Notation 2.74. To distinguish strong limits from weak limits in operator spaces, it is customary to use these symbols:

$$T = s\text{-}\lim T_n$$

means T is the limit of the sequence of operators $\{T_n\}_{n\in\mathbb{N}}$ in the strong topology; the same notation goes if the operators are functionals and the topology is the dual strong one. Similarly,

$$T = w\text{-}\lim T_n$$

denotes the limit in the weak topology of the sequence of operators $\{T_n\}_{n\in\mathbb{N}}$, and one writes

$$f = w^*\text{-}\lim f_n$$

if f is the limit of the sequence $\{f_n\}_{n\in\mathbb{N}}$ in the *-weak topology. ■

All the theory learnt so far eventually enables us to prove the last corollary to Banach–Steinhaus. If X is normed we know X' is complete in the strong topology, see Theorem 2.41(c)(ii). We can also prove completeness, as explained below, for the *-weak topology too, as long as X is a Banach space.

Corollary 2.75 (to the Banach–Steinhaus theorem). *If X is a Banach space on $\mathbb{K} = \mathbb{C}$, or $\mathbb{R}$, then X' is *-weak complete: if $\{f_n\}_{n\in\mathbb{N}} \subset \mathsf{X}'$ is such that $\{f_n(x)\}_{n\in\mathbb{N}}$ is a Cauchy sequence for any $x \in \mathsf{X}$, then there exists $f = w^*\text{-}\lim f_n \in \mathsf{X}'$.*

Proof. The field over which X is defined is complete by assumption, so for any $x \in \mathsf{X}$ there is $f(x) \in \mathbb{K}$ with $f_n(x) \to f(x)$. Immediately $f : \mathsf{X} \ni x \mapsto f(x)$ defines a linear functional. To end the proof we have to prove f is continuous. For any $x \in \mathsf{X}$ the sequence $f_n(x)$ is bounded (as Cauchy), so Banach–Steinhaus implies $|f_n(x)| \leq K < +\infty$ for all $x \in \mathsf{X}$ with $||x|| \leq 1$. Taking the limit for $n \to +\infty$ gives $|f(x)| \leq K$ if $||x|| = 1$, hence $||f|| \leq K < +\infty$. Therefore, by Theorem 2.40 f is continuous. □

As last topic of this section, related to the topological facts just seen, we state and prove a useful technical tool, the *Theorem of Banach–Alaoglu*: according to it, the unit ball in X', defined via the natural norm of X', is compact (Definition 1.19) in the *-weak topology of X'.

Theorem 2.76 (Banach–Alaoglu). *Let X be a normed space over $\mathbb{C}$. The closed unit ball $B := \{f \in X' \mid ||f|| \leq 1\}$ in the dual X' is compact in the *-weak topology.*

Proof. For any $x \in X$ define $B_x := \{c \in \mathbb{C} \mid |c| \leq ||x||\} \subset \mathbb{C}$. As B_x is obviously compact, Tychonoff's Theorem 1.25 forces $P := \times_{x\in X} B_x$ to be compact in the product topology. A point p in P is, for each $x \in X$, just a real number $p(x)$ with $|p(x)| \leq ||x||$. Elements in P are therefore functions (not necessarily linear!) $p : X \to \mathbb{C}$ such that $|p(x)| \leq ||x||$ for any $x \in X$. By construction $B \subset P$, and the topology induced by P on B is precisely the *-weak topology, as the definitions confirm. To end the proof we need to prove B is closed, because closed subsets in a compact space are compact. Suppose, then, $B \ni p_n \to p \in P$ as $n \to +\infty$, in the topology of P. Since $|p(x)| \leq ||x||$, to prove that $p \in B$ it suffices to see that p is linear. This is evident by arguing pointwise: if $a, b \in \mathbb{C}$ and $x, y \in X$, then

$$p(ax+by) = \lim_{n\to+\infty} p_n(ax+by) = a \lim_{n\to+\infty} p_n(x) + b \lim_{n\to+\infty} p_n(y) = ap(x) + bp(y)\,,$$

and the proof is over. □

We will see in Chapter 4 that B is never compact for the natural norm of X' if the space X' is infinite-dimensional. The same holds for any infinite-dimensional normed space.

2.3.4 Excursus: the theorem of Krein–Milman, locally convex metrisable spaces and Fréchet spaces

With this part we take a short break to digress on important properties of locally convex spaces in relationship to the issue of metrisability.

Let X be a locally convex space. In general, the topology induced by a seminorm or a family of seminorms $\mathscr{P} = \{p_i\}_{i\in I}$ on X will not be Hausdorff. It is easy to see the Hausdorff property holds if and only if $\cap_{i\in I} p_i^{-1}(0)$ is the null vector in X. This happens in particular if at least one p_i is a norm.

Locally convex Hausdorff spaces have this very relevant feature: not only extremal elements always exist in convex and compact subsets, but a convex subset is characterised by its extremal points. All this is the content of the known *Krein–Milman theorem*, which we only state [Rud91].

Theorem 2.77 (Krein–Milman). *Let X be a locally convex Hausdorff space and $K \subset \mathsf{X}$ a compact convex set. Then:*

(a) *the set E_K of extremal elements of K is not empty.*
(b) *$K = \overline{co(E_K)}$, where the bar denotes the closure in the ambient topology of X.*

And now to metrisable spaces. Let us recall a notion that should be familiar from basic courses.

Definition 2.78. *A* **metric space** *is a set M equipped with a function $d : \mathsf{M}\times\mathsf{M} \to \mathbb{R}$, called* **distance** *or* **metric***, such that, for every $x,y,z \in \mathsf{M}$:*
D1. *$d(x,y) = d(y,x)$;*
D2. *$d(x,y) \geq 0$, and $d(x,y) = 0$ iff $x = y$;*
D3. *$d(x,z) \leq d(x,y) + d(y,x)$.*

Remark 2.79. (1) Property **D1** is known as *symmetry* of the metric, **D2** is called *positive definiteness* and **D3** is the *triangle inequality*.
(2) Any normed space $(\mathsf{X}, ||\ ||)$ (hence also $\mathbb{R}^n$ and $\mathbb{C}^n$) admits a natural metric structure (X, d) by setting $d(x,x') := ||x - x'||$, $x,x' \in \mathsf{X}$. Then clearly

$$d(x+z, y+z) = d(x,y) \text{for any } x,y,z \in \mathsf{X};$$

in this sense the distance d is **translation-invariant**. ■

Generally speaking the structure of a metric space is much simpler than that of a normed space, because the former lacks the vector space operations. We have, nevertheless, the following notion, in complete analogy to normed spaces.

Definition 2.80. *Given a metric space (M, d), an* **open (metric) ball** *centred at x of radius $r > 0$ is the set:*

$$B_\delta(x) := \{y \in \mathsf{M} \mid d(x,y) < \delta\}\,. \tag{2.25}$$

Like normed spaces, metric spaces have a *natural topology* whose open sets are the empty set $\varnothing$ and the unions of *open metric balls* with any centre and radius.

Definition 2.81. *Let (M, d) be a metric space.*

(a) *$A \subset \mathsf{M}$ is* **open** *if $A = \varnothing$ or A is the union of open balls.*
(b) *The* **metric topology** *of M is the norm topology of open sets as in (a).*

Remark 2.82. **(1)** Exactly as in normed spaces, by checking the axioms we see that the metric topology is an honest topology, and open metric balls form a *basis* for it. The metric topology is trivially *Hausdorff*, as for normed spaces.
(2) If the metric space (X,d) is separable, i.e. it has a dense countable subset $S \subset \mathsf{X}$, then it is **second countable**: it has a countable basis $\mathscr{B}$ for the topology. The latter is the family of open balls centred on S with rational radii. One can prove the converse holds too [KoFo99]:

Proposition 2.83. *A metric space is second countable if and only if it is separable.*

(3) In a normed space $(\mathsf{X},||\ ||)$ the open balls defined by $||\ ||$ coincide with the open balls of the norm distance $d(x,x') := ||x-x'||$. Thus the two topologies of X, viewed as a normed or metric space, coincide.
(4) The previous remark applies in particular to $\mathbb{R}^n$ and $\mathbb{C}^n$, both metric spaces if we use the **Euclidean** or **standard distance**:

$$d((x_1,\ldots,x_n),(y_1,\ldots,y_n)) := \sqrt{\sum_{k=1}^{n} |x_k - y_k|^2}\,.$$

As observed above, the balls defined with the standard distance on $\mathbb{R}^n$ and $\mathbb{C}^n$ are precisely those associated to the standard norm (1.2) generating the standard topology. Therefore the topology defined by the Euclidean distance on $\mathbb{R}^n$ and $\mathbb{C}^n$ is just the standard topology.
(5) The metric spaces $\mathbb{R}^n$ and $\mathbb{C}^n$ are complete, for they are complete as normed spaces and the metric is the norm distance. ■

As for normed spaces, also metric spaces admit a characterisation of continuity equivalent to 1.16.

Definition 2.84. *Given metric spaces* $(\mathsf{M},d_{\mathsf{M}})$, $(\mathsf{N},d_{\mathsf{N}})$, *a map* $f:\mathsf{M}\to\mathsf{N}$ *is* **continuous at** $x_0 \in \mathsf{M}$ *if for any* $\varepsilon > 0$ *there is* $\delta > 0$ *such that* $d_{\mathsf{N}}(f(x),f(x_0)) < \varepsilon$ *whenever* $d_{\mathsf{M}}(x,x_0) < \delta$. *A function* $f:\mathsf{M}\to\mathsf{N}$ *is* **continuous** *if continuous at every point in* M.

Convergent sequences (Definition 1.13) specialise in metric spaces as they do in normed spaces.

Definition 2.85. *In a metric space* (M,d) *a sequence* $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{M}$ **converges** *to* $x \in \mathsf{M}$, *the* **limit of the sequence***:*

$$x_n \to x \quad \textit{as} \quad n \to +\infty \quad \textit{or} \quad \lim_{n\to+\infty} x_n = x\,,$$

if, for any $\varepsilon > 0$ *there is* $N_\varepsilon \in \mathbb{R}$ *such that* $d(x_n,x) < \varepsilon$ *for* $n > N_\varepsilon$, *i.e.*

$$\lim_{n\to+\infty} d(x_n,x) = 0\,.$$

It turns out, here too, that convergent sequences in the metric topology are Cauchy sequences (see below), but not conversely.

Definition 2.86. *Let* (M, d) *be a metric space.*

(a) *A sequence* $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{M}$ *is a* **Cauchy sequence** *if for any* $\varepsilon > 0$ *there is* $N_\varepsilon \in \mathbb{R}$ *such that* $d(x_n, x_m) < \varepsilon$ *when* $n, m > N_\varepsilon$.
(a) (M, d) *is* **complete** *if every Cauchy sequence converges somewhere in the space.*

A technically relevant problem is to tell whether a topological space, esp. a locally convex space, admits a distance function whose metric topology coincides with the pre-existing one (note that in general distances do not exist if the topology is induced by seminorms). But when that happens the space is called **metrisable**.

Going back to topological vector spaces, one can prove that any locally convex space $(\mathsf{X}, \mathscr{P})$ satisfying:

(a) $\mathscr{P} = \{p_n\}_{n=1,2,\ldots}$, i.e. $\mathscr{P}$ is *countable*;
(b) $\cap_{n=1,2,\ldots} p_n^{-1}(0) = 0$

is not just Hausdorff but even *metrisable*: the seminorm topology coincides with the metric topology of (X, d), provided we pick $d : \mathsf{X} \times \mathsf{X} \to \mathbb{R}_+$ suitably. In particular, the distance can be chosen to be:

$$d(x,y) := \sum_{n=1}^{+\infty} \frac{1}{2^n} \frac{p_n(x-y)}{1+p_n(x-y)}$$

(thus becoming invariant under translations). This is not the only possible distance that recovers the seminorm topology of X. Multiplying d by a given positive constant, for instance, will give a distance yielding the same topology as d.

A **Fréchet space** is a locally convex space X whose topology is Hausdorff, induced by a *finite or countable* number of seminorms, and *complete* when seen as metric space (X, d). A sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ is Cauchy for a distance d in a locally convex metrisable space X if and only if it is Cauchy for every seminorm p generating the topology: for every $\varepsilon > 0$ there is $N_\varepsilon^{(p)} \in \mathbb{R}$ such that $p(x_n - x_m) < \varepsilon$ whenever $n, m > N_\varepsilon^{(p)}$. Consequently, completeness does not actually depend on the distance used to generate the locally convex topology.

Fréchet spaces, which we will not treat in this book, are of highest interest in theoretical and mathematical physics as far as quantum field theories are concerned. Banach spaces are elementary instances of Fréchet spaces, of course.

Example 2.87. A good example of a Fréchet space is *Schwartz's space* . To define it we need some notation, which will come in handy at the end of Chapter 3 as well. Points in $\mathbb{R}^n$ will be denoed by letters and components by added subscripts, e.g. $x = (x_1, \ldots, x_n)$. A **multi-index** is an n-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$, $\alpha_i = 0, 1, 2, \ldots$, and $|\alpha|$ is conventionally the sum $|\alpha| := \sum_{i=1}^n \alpha_i$. Moreover,

$$\partial_x^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} .$$

Let $C^\infty(\mathbb{R}^n)$ denote the complex vector space of smooth complex functions on $\mathbb{R}^n$ (differentiable with continuity infinitely many times). **Schwartz's space** $\mathscr{S}(\mathbb{R}^n)$, seen as complex vector space, is the subspace in $C^\infty(\mathbb{R}^n)$ of functions f that *vanish at infinity, together with every derivative, faster than any inverse power of* $|x| := \sqrt{\sum_{i=1}^n x_i^2}$. Define

$$p_N(f) := \sup_{|\alpha|\le N} \sup_{x\in\mathbb{R}^n} (1+|x|^2)^N |(\partial_x^\alpha f)(x)| < +\infty \quad N = 0,1,2,\ldots$$

The above $p_N : \mathscr{S}(\mathbb{R}^n) \to \mathbb{R}_+$ are seminorms, and cleary satisfy $\cap_{N\in\mathbb{N}} p_N^{-1}(0) = 0$ because $p_0 = ||\ ||_\infty$ is a norm. Thus $\mathscr{S}(\mathbb{R}^n)$, with the topology induced by the seminorms $\{p_N\}_{N\in\mathbb{N}}$, becomes a locally convex space. It is easy to show that $\mathscr{S}(\mathbb{R}^n)$ is a Fréchet space [Rud91]. The points in the dual $\mathscr{S}(\mathbb{R})'$ of $\mathscr{S}(\mathbb{R})$, i.e. linear functionals from $\mathscr{S}(\mathbb{R}^n)$ to $\mathbb{C}$ that are continuous for the topology generated by the seminorms $\{p_N\}_{N\in\mathbb{N}}$, are the famous Schwartz distributions. ■

2.3.5 Baire's category theorem and its consequences: the open mapping theorem and the inverse operator theorem

We wish to discuss a general theorem about Banach spaces, the open mapping theorem, which counts among its consequences the continuity of inverse operators.

To prove these facts we introduce the minimum possible on *Baire spaces*.

Definition 2.88. *Let* $(\mathsf{X}, \mathscr{T})$ *be a topological space and* $S \subset \mathsf{X}$.

(a) *The* **interior** $Int(S)$ *of* S *is the set:*

$$Int(S) := \{x \in \mathsf{X} \mid \exists A \subset \mathsf{X}, A \text{ open and } x \in A \subset S\}.$$

(b) S *is* **nowhere dense** *if* $Int(\overline{S}) = \varnothing$.
(c) S *is a set of the* **first category**, *or* **meagre set**, *if it is the countable union of nowhere dense sets.*
(d) S *is a set of the* **second category**, *or* **non-meagre**, *if not of the first category.*

The following are immediate to prove.

(1) Countable unions of sets of the first category are of the first category.
(2) If $h : \mathsf{X} \to \mathsf{X}'$ is a homeomorphism, $S \subset \mathsf{X}$ is of the first/second category if and only if $h(S)$ is of the first/second category respectively.
(3) If $A \subset B \subset \mathsf{X}$ and B is of the first category in X, then A is of the first category.
(4) If $B \subset \mathsf{X}$ is closed and $Int(B) = \varnothing$, then B is of the first category in X.

We have the following important result.

Theorem 2.89 (Baire). *Let* (X, d) *be a complete metric space.*

(a) *If* $\{U_n\}_{n\in\mathbb{N}}$ *is a countable family of open dense sets in* X, *then* $\cap_{n\in\mathbb{N}} U_n$ *is dense in* X.
(b) X *is of the second category.*

Proof. (a) Let $A \subset \mathsf{X}$ be open. If $U_0 \cap A = \varnothing$, then $z \in A$ would admit an open neighbourhood disjoint from U_0, and hence not dense in X. Therefore $U_0 \cap A$ is open (intersection of open sets) and non-empty. Then there is an open metric ball $B_{r_0}(x_0)$ of radius $r_0 > 0$ and centre $x_0 \in \mathsf{X}$ (2.25) such that $\overline{B_{r_0}(x_0)} \subset U_0 \cap A$. We may repeat the procedure with $B_{r_0}(x_0)$ replacing A, U_1 replacing U_0, to find an open ball $B_{r_1}(x_1)$ with $\overline{B_{r_1}(x_1)} \subset U_1 \cap B_{r_0}(x_0)$. Iterating, we construct a countable collection of open balls $B_{r_n}(x_n)$ with $0 < r_n < 1/n$, such that $\overline{B_{r_n}(x_n)} \subset U_n \cap B_{r_{n+1}}(x_{n-1})$. Since $x_n \in \overline{B_{r_m}(x_m)}$ for $n \geq m$, the sequence $\{x_n\}_{n\in\mathbb{N}}$ is Cauchy. And since X is complete, $x_n \to x \in \mathsf{X}$ as $n \to +\infty$. By construction $x \in \overline{B_{r_{n-1}}(x_{n-1})} \subset B_{r_n}(x_n) \subset \cdots \subset U_0 \cap A \subset A$ for any $n \in \mathbb{N}$. Hence $x \in A \cap U_n$ for every $n \in \mathbb{N}$, and so $(\cap_{n\in\mathbb{N}} U_n) \cap A \neq \varnothing$ for every open subset $A \subset \mathsf{X}$. This implies $\cap_{n\in\mathbb{N}} U_n$ is dense in X, for it meets every open neighbourhood of any point in X.
(b) Assume $\{E_k\}_{k\in\mathbb{N}}$ is a collection of nowhere dense sets $E_k \subset \mathsf{X}$. If V_k is the complement of $\overline{E_k}$, it is open (its complement is closed) and dense in X (it is open and the complement's interior is empty). Part (a) then tells $\cap_{k\in\mathbb{N}} V_n \neq \varnothing$, so $\mathsf{X} \neq \cup_{k\in\mathbb{N}} \overline{E_k}$ by taking complements. *A fortiori* then $\mathsf{X} \neq \cup_{k\in\mathbb{N}} E_k$, so X is not of the first category and the claim is proved. □

Remark 2.90. **(1)** Baire's category theorem states, among other things, that *any collection, finite or countable, of dense open sets in a complete metric space always has non-empty (dense) intersection*. In the finite case it suffices to adapt the statement to $U_n = U_m$ for some $N \leq n, m$.
(2) Baire's theorem holds when X is a locally compact Hausdorff space. The first part is proved in analogy to the previous situation [Rud91], the second is identical.
(3) Baire's theorem applies, obviously, to Banach spaces, using the norm distance. ■

We can pass to the *open mapping theorem*. Remember a map $f : \mathsf{X} \to \mathsf{Y}$ between normed spaces (or topological spaces) is **open** if $f(A)$ is open in Y whenever $A \subset \mathsf{X}$ is open. As usual, $B_r^{(\mathsf{Z})}(z)$ denotes the open ball of radius r and centre z in the normed space $(\mathsf{Z}, || \, ||_{\mathsf{Z}})$.

Theorem 2.91 (Open mapping theorem of Banach–Schauder). *Let* X *and* Y *be Banach spaces over* $\mathbb{C}$ *or* $\mathbb{R}$*. If the operator* $T \in \mathfrak{B}(\mathsf{X}, \mathsf{Y})$ *is onto, then*

$$T(B_1^{(\mathsf{X})}(0)) \supset B_\delta^{(\mathsf{Y})}(0) \quad \text{for } \delta > 0 \text{ small enough}. \tag{2.26}$$

As a consequence, T is an open map.

Proof. Define in X the open ball $B_n := B_{2^{-n}}^{(\mathsf{X})}(0)$ at the origin, of radius 2^{-n}. We will show there is an open neighbourhood W_0 of the origin $0 \in \mathsf{Y}$ with:

$$W_0 \subset \overline{T(B_1)} \subset T(B_0), \tag{2.27}$$

and that will imply (2.26). To prove (2.27), note $B_1 \supset B_2 - B_2$ (from now on we use Notations 2.63), so

$$T(B_1) \supset T(B_2) - T(B_2)$$

and $\overline{T(B_1)} \supset \overline{T(B_2) - T(B_2)}$. On the other hand, since $\overline{A+B} \supset \overline{A} + \overline{B}$, $A, B \subset \mathsf{Y}$ with Y normed (prove it as exercise), we have:

$$\overline{T(B_1)} \supset \overline{T(B_2) - T(B_2)} \supset \overline{T(B_2)} - \overline{T(B_2)}. \tag{2.28}$$

The first inclusion of (2.27) is thus true if $\overline{T(B_2)}$ has non-empty interior: if $z \in Int(\overline{T(B_2)})$ then $z \in A \subset \overline{T(B_2)})$ with A open, so that $0 \in W_0 := A - A \subset \overline{T(B_2)} - \overline{T(B_2)} \subset \overline{T(B_1)}$ with W_0 open. To show $Int(\overline{T(B_2)}) \neq \varnothing$, notice that the assumptions imply

$$\mathsf{Y} = T(\mathsf{X}) = \bigcup_{k=1}^{+\infty} kT(B_2), \tag{2.29}$$

because B_2 is an open neighbourhood of 0. But Y is of the second category, so at least one $kT(B_2)$ is of the second category (otherwise Y would be of the first category, which is impossible by the second statement in Baire's category Theorem 2.89, for Y is complete). Since $y \mapsto ky$ is a homeomorphism of Y, $T(B_2)$ is of the second category in Y. Hence the closure of $T(B_2)$ has non-empty interior, proving one inclusion of (2.27). For the other inclusion (the second from the left), we use a sequence of elements $y_n \in \mathsf{Y}$ built inductively. Fix $y_1 \in \overline{T(B_1)}$ and suppose, for $n \geq 1$, that y_n is in $\overline{T(B_n)}$ and let us define y_{n+1} as follows. What was proved for $\overline{T(B_1)}$ holds for $\overline{T(B_{n+1})}$ too, so $\overline{T(B_{n+1})}$ contains an open neighbourhood of the origin. Now:

$$\left(y_n - \overline{T(B_{n+1})}\right) \cap T(B_n) \neq \varnothing, \tag{2.30}$$

implying there exists $x_n \in B_n$ such that:

$$T(x_n) \in y_n - \overline{T(B_{n+1})}. \tag{2.31}$$

Define: $y_{n+1} := y_n - Tx_n$ and note it belongs to $\overline{T(B_{n+1})}$. This is the inductive step. Since $||x_n|| < 2^{-n}$, $n = 1, 2, \ldots$, the sum $x_1 + \cdots + x_n$ gives a Cauchy sequence converging to some $x \in \mathsf{X}$ by completeness of X, and $||x|| < 1$. Hence $x \in B_0$. Since:

$$\sum_{n=1}^{m} Tx_n = \sum_{n=1}^{m} (y_n - y_{n+1}) = y_1 - y_{m+1}, \tag{2.32}$$

and because $y_{m+1} \to 0$ as $m \to +\infty$ (by continuity of T), we conclude $y_1 = Tx \in T(B_0)$. Now as y_1 was generic in $\overline{T(B_1)}$, that proves the second inclusion of (2.27) and ends the first proof. As for the second statement, (2.26) and the linearity of T imply that the image under T of any open ball $B_\varepsilon^{(\mathsf{X})}(x) = x + \varepsilon B_1(0)$, centred at any $x \in \mathsf{X}$, contains the open ball in Y centred at Tx: $B_{\delta_\varepsilon}^{(\mathsf{Y})}(0) := Tx + \varepsilon B_\delta^{(\mathsf{Y})}(0)$ ($\delta > 0$ sufficiently small). Therefore the image under T of an open set $A = \cup_{x\in A} B_{\varepsilon_x}^{(\mathsf{X})}(x)$ is open in Y: $T(A) = \cup_{x\in A} B_{\delta_{\varepsilon_x}}^{(\mathsf{Y})}(Tx)$. This means T is *open*. □

The most important elementary corollary of this theorem is without doubt *Banach's inverse operator theorem* for Banach spaces (there is a version for complete metric vector spaces as well).

Theorem 2.92 (Inverse operator theorem of Banach). *Let* X *and* Y *be Banach spaces over* $\mathbb{C}$ *or* $\mathbb{R}$*, and suppose* $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$ *is injective and surjective. Then*
(a) $T^{-1} : \mathsf{Y} \to \mathsf{X}$ *is bounded, i.e.* $T^{-1} \in \mathfrak{B}(\mathsf{Y},\mathsf{X})$.
(b) *There exists* $c > 0$ *such that:*

$$||Tx|| \geq c||x|| \,, \quad \text{for any } x \in \mathsf{X}. \tag{2.33}$$

Proof. (a) That T^{-1} is linear is straightforward, for we need only prove it is continuous. As T is open, the pre-image under T^{-1} of an open set in X is open, making T^{-1} continuous. (b) Since T^{-1} is bounded, there is $K \geq 0$ with $||T^{-1}y|| \leq K||y||$, for any $y \in \mathsf{Y}$. Notice that $K > 0$, for otherwise $T^{-1} = 0$ and T^{-1} could not be invertible. For any $x \in \mathsf{X}$ we set $y = Tx$. Substituting in $||T^{-1}y|| \leq K||y||$ gives back, for $c = 1/K$, equation (2.33). □

2.3.6 The closed graph theorem

Now we discuss a very useful theorem in operator theory, known as the *closed graph theorem*.

Notation 2.93. (1) If X is a vector space and $\varnothing \neq X_1, \ldots, X_n \subset \mathsf{X}$, then:

$$< X_1, \cdots, X_n >$$

will denote the linear **span** of the sets X_i, i.e. the vector subspace of X containing all *finite* linear combinations of elements of any X_i.
(2) Take $\varnothing \neq \mathsf{X}_1, \ldots, \mathsf{X}_n$ *subspaces* of a vector space X. Then

$$\mathsf{Y} = \mathsf{X}_1 \oplus \cdots \oplus \mathsf{X}_n$$

denotes the **direct sum** $\mathsf{Y} \subset \mathsf{X}$ of the X_i, i.e.:

(i) $\mathsf{Y} = < \mathsf{X}_1, \cdots, \mathsf{X}_n >$ (so Y is a subspace in X) and
(ii) $\mathsf{X}_i \cap \mathsf{X}_j = \{0\}$ for any pair $i, j = 1, \ldots, n$, $i \neq j$.

As is well known, (i) and (ii) are equivalent to demanding

$$x \in \mathsf{Y} \Rightarrow x = x_1 + \cdots + x_n \quad \text{with } x_k \in \mathsf{Y}_k \text{ determined } \textit{uniquely} \text{ by } x,\ k = 1, \ldots, n.$$

(3) If $\mathsf{X}_1, \ldots, \mathsf{X}_n$ are *vector spaces* over the same field $\mathbb{K} = \mathbb{C}$ or $\mathbb{R}$, we may furnish $\mathsf{X}_1 \times \cdots \times \mathsf{X}_n$ with the structure of a $\mathbb{K}$-vector space by:

$$\alpha(x_1, \ldots, x_n) := (\alpha x_1, \ldots, \alpha x_n) \quad \text{and}$$
$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

for any $\alpha \in \mathbb{K}$, $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathsf{X}_1 \times \cdots \times \mathsf{X}_n$. Calling

$$\Pi_{\mathsf{X}_k} : (x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n) \mapsto (0, \ldots, 0, x_k, 0, \ldots, 0)$$

the kth **canonical projection**, the vector space built on $\mathsf{X}_1 \times \cdots \times \mathsf{X}_n$ coincides with $Ran(\Pi_{\mathsf{X}_1}) \oplus \cdots \oplus Ran(\Pi_{\mathsf{X}_n})$. As each X_k is naturally identified with the corresponding $Ran(\Pi_{\mathsf{X}_k})$, we will write $\mathsf{X}_1 \oplus \cdots \oplus \mathsf{X}_n$ to denote the natural vector space $\mathsf{X}_1 \times \cdots \times \mathsf{X}_n$ above, even when the X_k are not all contained in one common space. ■

To prove the theorem we need some preliminaries. First, if $(\mathsf{X}, N_{\mathsf{X}})$ and $(\mathsf{Y}, N_{\mathsf{Y}})$ are normed spaces over $\mathbb{K} = \mathbb{C}$, or $\mathbb{R}$, we can consider $\mathsf{X} \oplus \mathsf{Y}$, in the Notation 2.93(3). The space $\mathsf{X} \oplus \mathsf{Y}$ has the *product topology* of X and Y, seen in Definition 1.10. The operations of the vector space $\mathsf{X} \oplus \mathsf{Y}$ are continuous in the product topology, as one proves with ease (the proof is the same as the one used for the operations on a normed space). And the canonical projections $\Pi_{\mathsf{X}} : \mathsf{X} \oplus \mathsf{Y} \to \mathsf{X}$, $\Pi_{\mathsf{Y}} : \mathsf{X} \oplus \mathsf{Y} \to \mathsf{Y}$ are continuous in the product topology on the domain and the topologies of X and Y on the codomains, another easy fact.

The product topology of $\mathsf{X} \oplus \mathsf{Y}$ admits **compatible norms**: there exist norms on $\mathsf{X} \oplus \mathsf{Y}$ inducing the product topology. One possibility is:

$$||(x,y)|| := \max\{N_{\mathsf{X}}(x), N_{\mathsf{Y}}(y)\} \quad \text{for any } (x,y) \in \mathsf{X} \oplus \mathsf{Y}. \tag{2.34}$$

That this norm generates the product topology, i.e. open sets are unions of products of open balls in X and Y, is proved as follows. Take the open neighbourhood of (x_0, y_0) product of two open balls $B_\delta^{(\mathsf{X})}(x_0) \times B_\mu^{(\mathsf{Y})}(y_0)$ in X and Y respectively. The open ball in $\mathsf{X} \oplus \mathsf{Y}$

$$\{(x,y) \in \mathsf{X} \times \mathsf{Y} \mid ||(x,y) - (x_0,y_0)|| < \min\{\delta, \mu\}/2\}$$

centred at (x_0, y_0) is contained in $B_\delta^{(\mathsf{X})}(x_0) \times B_\mu^{(\mathsf{Y})}(y_0)$. *Vice versa*, the product $B_\delta^{(\mathsf{X})}(x_0) \times B_\delta^{(\mathsf{Y})}(y_0)$, to which (x_0, y_0) belongs, is contained in the open ball

$$\{(x,y) \in \mathsf{X} \times \mathsf{Y} \mid ||(x,y) - (x_0,y_0)|| < \varepsilon\}$$

centred in (x_0, y_0), if $\varepsilon > \delta$. This implies that unions of products of metric balls in X and Y is also union of metric balls in norm (2.34), and conversely too. Hence the two topologies coincide and the proof ends.

Immediately we can prove $(\mathsf{X} \oplus \mathsf{Y}, || \, ||)$ *is a Banach space if such are* $(\mathsf{X}, N_{\mathsf{X}})$ *and* $(\mathsf{Y}, N_{\mathsf{Y}})$. (By Proposition 2.101, proved later, this fact will guarantee that any norm generating the product topology makes $\mathsf{X} \oplus \mathsf{Y}$ a Banach space.) In fact let $\{(x_n, y_n)\}$ be a Cauchy sequence in $\mathsf{X} \oplus \mathsf{Y}$. Then $\{x_n\}$ and $\{y_n\}$ are both Cauchy in X and Y respectively, by the above definition of norm on $\mathsf{X} \oplus \mathsf{Y}$. Call $x \in \mathsf{X}$ and $y \in \mathsf{Y}$ the limits of those sequences, which exist for X and Y are Banach spaces. If $\varepsilon > 0$, there are positive integers N_x and N_y satisfying

$$||(x,y) - (x_n, y_n)|| < \varepsilon$$

if $n > \max\{N_x, N_y\}$, Therefore $(x_n, y_n) \to (x,y)$ as $n \to +\infty$ in the norm topology of $\mathsf{X} \oplus \mathsf{Y}$, and the latter is a Banach space.

Definition 2.94. *Let* X, Y *be normed spaces on* $\mathbb{C}$ *or* $\mathbb{R}$. $T \in \mathfrak{L}(\mathsf{X}, \mathsf{Y})$ *is said* **closed** *if the* **graph** *of the operator* T*, i.e. the subspace*

$$G(T) := \{(x, Tx) \in \mathsf{X} \oplus \mathsf{Y} \mid x \in \mathsf{X}\},$$

is closed in the product topology. Equivalently, T *is closed iff for any converging sequence* $\{x_n\}_{n \in \mathbb{N}} \subset \mathsf{X}$ *such that* $\{Tx_n\}_{n \in \mathbb{N}}$ *converges in* Y*, we have:*

$$\lim_{n \to \infty} T(x_n) = T(\lim_{n \to \infty} x_n).$$

The last equivalence relies on a general fact: a set ($G(T)$ in our case) is closed if and only if it coincides with its closure, if and only if it contains its limit points; making this explicit in the product topology gives our proof. We are ready for the *closed graph theorem.*

Theorem 2.95 (Closed graph theorem). *Let* $(\mathsf{X}, || \; ||_{\mathsf{X}})$ *and* $(\mathsf{Y}, || \; ||_{\mathsf{Y}})$ *be Banach spaces over* $\mathbb{K} = \mathbb{C}$.
Then $T \in \mathfrak{L}(\mathsf{X}, \mathsf{Y})$ *is bounded if and only if it is closed.*

Proof. Suppose T is bounded. Then it is banally closed by the definition of closed operator. Assume T is closed. Consider the linear bijective map $M : G(T) \ni (x, Tx) \mapsto x \in \mathsf{X}$. By hypothesis $G(T)$ is a closed subspace in the Banach space $\mathsf{X} \oplus \mathsf{Y}$, hence it becomes Banach for the restricted norm $|| \; ||$ of (2.34). The latter's definition implies $||M(x, Tx)||_{\mathsf{X}} = ||x||_{\mathsf{X}} \leq ||(x, Tx)||$, so M is bounded. Banach's bounded inverse theorem tells $M^{-1} : \mathsf{X} \to G(T) \subset \mathsf{X} \oplus \mathsf{Y}$ is bounded. As the canonical projection $\Pi_{\mathsf{Y}} : \mathsf{X} \oplus \mathsf{Y} \to \mathsf{Y}$ is continuous, we infer that the linear map $\Pi_{\mathsf{Y}} \circ M^{-1} : x \mapsto Tx$ is continuous, hence bounded. □

2.4 Projectors

Using the closed graph theorem we define a class of continuous operators, called *projectors*. This notion plays the leading role in formulating QM when the normed space is a Hilbert space.

Definition 2.96. *Let* $(\mathsf{X}, || \; ||)$ *be a normed space over* $\mathbb{C}$ *or* $\mathbb{R}$*. The operator* $P \in \mathfrak{B}(\mathsf{X})$ *is a* **projector** *if it is* **idempotent**, *i.e.*

$$PP = P\,. \tag{2.35}$$

The target $\mathsf{M} := P(\mathsf{X})$ *is called* **projection space** *of* P*, and we say* P **projects** *on* M.

Projectors are naturally associated to a direct sum decomposition of X into a pair of closed subspaces.

Proposition 2.97. *Let* $P \in \mathfrak{B}(\mathsf{X})$ *be a projector on the normed space* $(\mathsf{X}, || \; ||)$.
(a) *If* $Q : \mathsf{X} \to \mathsf{X}$ *is the linear map such that*

$$Q + P = I\,, \tag{2.36}$$

then Q *is a projector and:*

$$PQ = QP = 0\,, \tag{2.37}$$

where 0 *is the null operator (transforming any vector into the null vector* $0 \in \mathsf{X}$*).*
(b) *The projection spaces* $\mathsf{M} := P(\mathsf{X})$ *and* $\mathsf{N} := Q(\mathsf{X})$ *are closed subspaces satisfying:*

$$\mathsf{X} = \mathsf{M} \oplus \mathsf{N}\,. \tag{2.38}$$

Proof. (a) Q is continuous as sum of continuous operators, $QQ = (I-P)(I-P) = I - 2P + PP = I - 2P + P = I - P = Q$. $PQ = P(I-P) = P - PP = P - P = 0$, $(I-P)P = P - PP = P - P = 0$.
(b) If $P(x_n) \to y$ as $n \to +\infty$, by continuity of P we have $PP(x_n) \to P(y)$. Using equation (2.35) we rephrase this as $P(x_n) \to P(y)$, whence $y = P(y)$ by uniqueness of the limit (X is Hausdorff). So, $y \in \overline{\mathsf{M}}$ implies $y \in \mathsf{M}(\subset \overline{\mathsf{M}})$, and $\mathsf{M} = \overline{\mathsf{M}}$ is closed. The same argument proves N is closed. That M, N are subspaces is immediate from the linearity of P and Q. If we take $x \in \mathsf{X}$, then

$$x = P(x) + Q(x)\,,$$

and $\mathsf{X} =< \mathsf{M}, \mathsf{N} >$. To finish we need to have $\mathsf{M} \cap \mathsf{N} = \{0\}$. Pick $x \in \mathsf{M} \cap \mathsf{N}$. Then $x = P(x)$, so $x = Q(x)$ by (2.35) ($x \in \mathsf{M}$ implies $x = Pz$ for some $z \in \mathsf{X}$, but then $Px = PPz = Pz = x$). Using Q on $x = Px$, and recalling $x = Qx$, gives $x = Q(x) = QP(x) = 0$ by (2.37), i.e. $x = 0$. □

The closed graph theorem explains that Proposition 2.97 can be reversed, provided we further suppose the ambient space is Banach.

Proposition 2.98. *Let* $(\mathsf{X}, ||\ ||)$ *be a Banach space,* $\mathsf{M}, \mathsf{N} \subset \mathsf{X}$ *closed subspaces such that* $\mathsf{X} = \mathsf{M} \oplus \mathsf{N}$. *Consider the functions* $P : \mathsf{X} \to \mathsf{M}$ *and* $Q : \mathsf{X} \to \mathsf{N}$ *that map* $x \in \mathsf{X}$ *to the respective elements in* M *and* N *according to* $\mathsf{X} = \mathsf{M} \oplus \mathsf{N}$. *Then:*

(a) *P and Q are projectors on* M *and* N *respectively.*
(b) *Properties (2.36) and (2.37) hold.*

Proof. By assumption $x \in \mathsf{X}$ decomposes as $x = u_{\mathsf{M}} + u_{\mathsf{N}}$ for certain $u_{\mathsf{M}} \in \mathsf{M}$, $u_{\mathsf{N}} \in \mathsf{N}$, and the sum is unique once the subspaces are fixed. Uniqueness, and the fact that M and N are closed under linear combinations, imply that $P : x \mapsto u_{\mathsf{M}}$ and $Q : x \mapsto u_{\mathsf{N}}$ are linear, $PP = P$ and $QQ = Q$. Note that $P(\mathsf{X}) = \mathsf{M}$ and $Q(\mathsf{X}) = \mathsf{N}$ by construction; moreover (2.36) holds since $\mathsf{X} =< \mathsf{M}, \mathsf{N} >$, while (2.37) is true by $\mathsf{M} \cap \mathsf{N} = \{0\}$. To finish we need to see P and Q are continuous. Let us show P is closed, and the closed graph theorem will force continuity. The strategy for Q is analogous. So let $\{x_n\} \subset \mathsf{X}$ be a sequence converging to $x \in \mathsf{X}$, and such that also $\{Px_n\}$ converges in X. We claim that

$$Px = \lim_{n \to +\infty} Px_n\,.$$

As N is closed,

$$\mathsf{N} \ni Qx_n = x_n - Px_n \to x - \lim_{n \to +\infty} Px_n = z \in \mathsf{N}\,.$$

So we have

$$x = \lim_{n \to +\infty} Px_n + z\,,$$

with $z \in \mathsf{N}$, but $\lim_{n \to +\infty} Px_n \in \mathsf{M}$ as well, because M is closed and $Px_n \in \mathsf{M}$ for all n. On the other hand we know that

$$x = Px + Qx.$$

Since the decomposition is unique, necessarily

$$Px = \lim_{n\to+\infty} Px_n$$

and $z = Qx$. Therefore P is closed and so continuous. □

2.5 Equivalent norms

One interesting consequence of Banach's inverse operator theorem is a criterion to establish when two norms on a vector space, complete for both, induce the same topology. Before stating the criterion (Proposition 2.101), let us begin with preliminaries.

The section ends with the proof that all norms on a vector space of finite dimension are equivalent, and make the space Banach.

Definition 2.99. *Two norms N_1, N_2 defined on one vector space X (over $\mathbb{C}$ or $\mathbb{R}$) are* **equivalent** *if there are constants $c, c' > 0$ such that:*

$$cN_2(x) \le N_1(x) \le c'N_2(x)\,, \quad \textit{for any } x \in \mathsf{X}. \tag{2.39}$$

Remark 2.100. (1) Note how (2.39) is equivalent to the similar inequality obtained by swapping N_1, N_2, and writing $1/c', 1/c$ in place of c, c' respectively.
(2) By this observation, it is straightforward that *if a given normed space is complete, then it is complete for any equivalent norm.*
(3) Two equivalent norms on a vector space generate the same topology, as is easy to prove. The next proposition discusses the converse.
(4) Equivalent norms define an equivalence relation on the space of norms on a given vector space. The proof is immediate from the definitions. ■

Proposition 2.101. *Let X be a vector space on $\mathbb{C}$ or $\mathbb{R}$. The norms N_1 and N_2 on X are equivalent if and only if the indentity map $I : (\mathsf{X}, N_2) \ni x \mapsto x \in (\mathsf{X}, N_1)$ is a homeomorphism (which is to say, the metric topologies generated by the norms are the same).*

Proof. It suffices to prove the 'if' part, for the sufficient condition is trivial by definition of equivalent norms. If I is a homeomorphism it is continuous at the origin, and in particular the unit open ball (for N_1) centred at 0 must contain an entire open ball (for N_2) at 0 of small enough radius $\delta > 0$. That is to say, $N_2(x) \le \delta \Rightarrow N_1(x) < 1$. In particular, for $x \neq 0$, $N_2(\delta x/N_2(x)) \le \delta$, so $N_1(\delta x/N_2(x)) < 1$, i.e. $\delta N_1(x) \le N_2(x)$. For $x = 0$ the equality is trivial. So we proved there is $c' = 1/\delta > 0$ for which $N_1(x) \le c'N_2(x)$, for any $x \in \mathsf{X}$. The other half of (2.39) is similar if we swap spaces. □

Proposition 2.102. *Let X be a vector space on $\mathbb{C}$ or $\mathbb{R}$ and suppose N_1, N_2 both make X Banach. If there is a constant $c > 0$ such that:*

$$cN_2(x) \le N_1(x)$$

for any $x \in \mathsf{X}$, the norms are equivalent.

Proof. Consider the identity $I : x \mapsto x$, a linear and continuous map when thought of as $I : (\mathsf{X}, N_1) \to (\mathsf{X}, N_2)$, since $N_2(x) \leq (1/c)N_1(x)$ for all $x \in \mathsf{X}$. Banach's inverse function theorem, part (b), guarantees the existence of $c' > 0$ such that $N_1(x) \leq c' N_2(x)$ for all $x \in \mathsf{X}$. Then N_1 and N_2 satisfy (2.39). □

The important, and final, proposition in this section holds also on real vector spaces (writing $\mathbb{R}$ instead of $\mathbb{C}$ in the proof).

Proposition 2.103. *Let* X *be a* $\mathbb{C}$*-vector space of finite dimension. Then all norms are equivalent, and any one defines a Banach structure on* X.

Proof. We can simply study $\mathbb{C}^n$, given that any complex vector space of finite dimension n is isomorphic to $\mathbb{C}^n$. Owing to remarks (2) and (4) above, it is sufficient to prove that any norm on $\mathbb{C}^n$ is equivalent to the standard norm. Keep in mind the fact, known from elementary analysis, that the standard $\mathbb{C}^n$ is complete, so any other equivalent norm makes it a Banach space, by Remark 2.100(2).

Let N be a norm on $\mathbb{C}^n$ and $e_1, \cdots, e_n$ the canonical basis. If $x = \sum_i x_i e_i$ and $y = \sum_i y_i e_i$ are generic points in $\mathbb{C}^n$, from properties (N0), (N2) and (N1) (cf. the definition of norm) we have

$$0 \leq N(x-y) \leq \sum_{i=1}^{n} |x_i - y_i| N(e_i) \leq Q \sum_{i=1}^{n} |x_i - y_i| ,$$

where $Q := \sum_i N(e_i)$. At the same time, trivially, if $||\cdot||$ is the standard norm then

$$|x_j - y_j| \leq Max\{|x_i - y_i| \,|\, i = 1, 2, \cdots, n\} \leq \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} = ||x - y|| ,$$

whence

$$0 \leq N(x-y) \leq nQ||x-y|| .$$

This shows N is continuous in the standard topology. If $\mathbb{S} := \{x \in \mathbb{C}^n \,|\, ||x|| = 1\}$, and N' is a second norm on $\mathbb{C}^n$ continuous in the standard topology, the map

$$\mathbb{S} \ni x \mapsto f(x) := \frac{N(x)}{N'(x)}$$

is continuous as ratio of continuous maps with non-zero denominator. But $\mathbb{S}$ is compact in the standard topology, so f has a minimum m and a maximum M. In particular, $M \geq m > 0$ because N, N' are strictly positive on $\mathbb{S}$ and m, M are attained at suitable points x_m, x_M in $\mathbb{S}$. By construction

$$mN'(x) \leq N(x) \leq MN'(x) , \quad \text{for any } x \in \mathbb{S} .$$

We claim that this inequality actually holds for any $x \in \mathbb{C}^n$: write $x = \lambda x_0$ with $x_0 \in \mathbb{S}$ and $\lambda \geq 0$. Multiplying by $\lambda \geq 0$ the inequality, evaluating it at x_0 and using property **N1** gives precisely:

$$mN'(x) \leq N(x) \leq MN'(x) \quad \text{for any } x \in \mathbb{C}^n.$$

Now choosing $N' := ||\cdot||$ we conclude that any norm on $\mathbb{C}^n$ is equivalent to the standard one. □

2.6 The fixed-point theorem and applications

In this last section of the chapter, we present an elementary theorem with crucial consequences in analysis, especially in the theory of differential equations: the *fixed-point theorem*. We will state it for complete metric spaces and then examine it on Banach spaces.

2.6.1 The fixed-point theorem of Banach-Caccioppoli

Start with a definition about metric spaces, cf. Definition 2.78.

Definition 2.104. *Let* (M, d) *be a metric space. A map* $G : \mathsf{M} \to \mathsf{M}$ *is a* **contraction (map)** *in case there is a real number* $\lambda \in [0,1)$ *for which:*

$$d(G(x), G(y)) \leq \lambda d(x,y) \quad \text{for any } x, y \in \mathsf{M}. \tag{2.40}$$

Remember normed spaces $(\mathsf{X}, || \ ||)$ are metric spaces once we specify the norm distance $d(x,y) := ||x-y||$ (and the metric topology induced by d coincides with the topology induced by $|| \ ||$, as seen in Section 2.3.4). Hence we can specialise the definition to normed spaces.

Definition 2.105. *Let* $(\mathsf{Y}, || \ ||)$ *be a normed space and* $\mathsf{X} \subset \mathsf{Y}$ *a subset (possibly the whole* Y*). A function* $G : \mathsf{X} \to \mathsf{X}$ *is a* **contraction** *if there exists a real number* $\lambda \in [0,1)$ *for which:*

$$||G(x) - G(y)|| \leq \lambda ||x-y|| \quad \text{for all} \quad x, y \in \mathsf{X}. \tag{2.41}$$

Remark 2.106. (1) Note that the value $\lambda = 1$ is explicitly *excluded*.
(2) The demand of (2.40) implies immediately that any contraction is continuous in the metric topology of (M, d).
Similarly, (2.41) tells that any contraction on the set X is continuous in the induced norm topology of $(\mathsf{Y}, || \ ||)$.
(3) We stress that, in Definition 2.105, (a) the function G is *not* requested to be linear, and (b) X is not necessarily a subspace of Y, but only a subset. Linear structures play no interesting role. ■

Let us state and prove the *fixed-point theorem* (of Banach and Caccioppoli) for metric spaces.

Theorem 2.107 (Fixed-point theorem for metric spaces). *Let* $G : \mathsf{M} \to \mathsf{M}$ *be a contraction on the complete metric space* (M, d)*. Then there exists a unique element* $z \in \mathsf{M}$*, called* **fixed point***:*

$$G(z) = z. \tag{2.42}$$

A weaker version states that if $G : \mathsf{M} \to \mathsf{M}$ *is not a contraction, but the n-fold composite* $G^n = G \circ \cdots \circ G$ *is for some given* $n = 1, 2, \ldots$*, then G admits a unique fixed point.*

Proof. Let us begin by proving the existence of z. Consider, for $x_0 \in \mathsf{M}$ arbitrary, the sequence defined recursively by $x_{n+1} = G(x_n)$. The claim is this is a Cauchy sequence, and that its limit is a fixed point of G. Without loss of generality suppose $m \geq n$. If $m = n$, trivially $d(x_m, x_n) = 0$; if $m > n$ we employ the triangle inequality repeatedly to get:

$$d(x_m, x_n) \leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n)\,. \tag{2.43}$$

The generic summand on the right equals

$$\begin{aligned} d(x_{p+1}, x_p) = d(G(x_p), G(x_{p-1})) \leq \lambda d(x_p, x_{p-1}) &= \lambda d(G(x_{p-1}), G(x_{p-2})) \\ &\leq \lambda^2 d(x_{p-1}, x_{p-2}) \\ &\leq \cdots \leq \lambda^p d(x_1, x_0)\,. \end{aligned}$$

Hence, for $p = 1, 2, \ldots$ we have $d(x_{p+1}, x_p) \leq \lambda^p d(x_1, x_0)$. Inserting the latter inequality in the right-hand side of (2.43) produces the estimate:

$$\begin{aligned} d(x_m, x_n) &\leq d(x_1, x_0) \sum_{p=n}^{m-1} \lambda^p = d(x_1, x_0) \lambda^n \sum_{p=0}^{m-n-1} \lambda^p \\ &\leq \lambda^n d(x_1, x_0) \sum_{p=0}^{+\infty} \lambda^p \leq d(x_1, x_0) \frac{\lambda^n}{1-\lambda} \end{aligned}$$

where we used the fact that $\sum_{p=0}^{+\infty} \lambda^p = (1-\lambda)^{-1}$ if $0 \leq \lambda < 1$. In conclusion:

$$d(x_m, x_n) \leq d(x_1, x_0) \frac{\lambda^n}{1-\lambda}\,. \tag{2.44}$$

For us $|\lambda| < 1$, so $d(x_1, x_0)\lambda^n/(1-\lambda) \to 0$ as $n \to +\infty$. Hence $d(x_m, x_n)$ can be rendered as small as we like by picking the minimum between m and n to be arbitrarily large. Therefore the sequence $\{x_n\}_{n \in \mathbb{N}}$ is Cauchy. But M is complete, so $\lim_{n \to +\infty} x_n = x \in \mathsf{M}$ for a certain x. Moreover, G is a contraction, so continuous, and

$$G(x) = G\left(\lim_{n \to +\infty} x_n\right) = \lim_{n \to +\infty} G(x_n) = \lim_{n \to +\infty} x_{n+1} = x\,,$$

as claimed.

Let us see to uniqueness. For that assume x and x' satisfy $G(x) = x$ and $G(x') = x'$. Then

$$d(x, x') = d(G(x), G(x')) \leq \lambda d(x, x')\,.$$

If $d(x, x') \neq 0$, dividing by $d(x, x')$ would give $1 \leq \lambda$, absurd by assumption. Hence $d(x, x') = 0$, so $x = x'$, because d is positive definite.

Now let us prove the theorem if $B := G^n$ is a contraction. By the previous part B has a unique fixed point z. Clearly, if G admits a fixed point, this must be z. There remains to show that z is fixed under G as well. As B is a contraction, the sequence

$B(z_0), B^2(z_0), B^3(z_0), \dots$ converges to z, irrespective of $z_0 \in \mathsf{M}$, as we saw earlier in the proof. Therefore

$$G(z) = G(B^k(z)) = B^k(G(z)) = B^k(z_0) \to z \quad \text{as } k \to +\infty,$$

and $G(z) = z$. □

Moving to normed spaces, the theorem has as corollary the next fact, obtained using the norm distance $d(x,y) := ||x-y||$.

Theorem 2.108 (Fixed-point theorem for normed spaces). *Let $G : \mathsf{X} \to \mathsf{X}$ be a contraction on the closed set $\mathsf{X} \subset \mathsf{B}$, with B a Banach space over $\mathbb{R}$ or $\mathbb{C}$. There exists a unique element $z \in \mathsf{X}$, called* **fixed point***:*

$$G(z) = z\,. \tag{2.45}$$

A weaker version states that if $G : \mathsf{X} \to \mathsf{X}$ is not a contraction, but the n-fold composite $G^n = G \circ \cdots \circ G$ is for some $n = 1, 2, \dots$, then G admits a unique fixed point.

Proof. Define $\mathsf{M} := \mathsf{X}$ and $d(x,y) := ||x-y||$, $x, y \in \mathsf{X}$. Thus X is a metric space. (X, d) is complete as well. In fact, a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ for d is such for $||\ ||$ too, as is easy to verify. $(\mathsf{B}, ||\ ||)$ is complete so the limit $x \in \mathsf{B}$ of $\{x_n\}_{n\in\mathbb{N}}$ exists. And since X is closed inside B, $x \in \mathsf{X}$. Hence any Cauchy sequence of (X, d) converges in X, making (X, d) complete. At this point we invoke the previous theorem for the metric space (X, d) and conclude. □

The significance of the fixed-point theorem, by the way, depends on its role in proving existence and uniqueness theorems for equations of all sorts, especially integral and differential ones; the gist is to show that the equation to which we seek the solution z can be written as a fixed-point relation $G(z) = z$ in a suitable Banach space (or complete metric space). Example (1) below is a relatively simple case (G is linear), while the ensuing (2) typically pertains nonlinear contractions.

Examples 2.109. Let us present two elementary instances of how the fixed-point theory is used. A more important situation will be treated in the following section.

(1) The **homogeneous Volterra equation on** $C([a,b])$ in the unknown $f \in C([a,b])$ reads:

$$f(x) = \int_a^x K(x,y) f(y) dy\,, \tag{2.46}$$

where K is a continuous function bounded by $M \geq 0$. We equip the Banach space $C([a,b])$ with sup norm $||\ ||_\infty$. The equation may be written in the form $f = Af$, where $A : C([a,b]) \to C([a,b])$ is the integral operator determined by the integral kernel K:

$$(Af)(x) := \int_a^x K(x,y) f(y) dy\,, \quad f \in C([a,b]). \tag{2.47}$$

If a solution exists, then clearly it is the fixed point of A. Not only this: the solution is also fixed under every operator A^n whichever power $n = 1, 2, \dots$ we take. Let us show that we can fix n so to make A^n a contraction. By virtue of Theorem 2.108 this would guarantee the homogeneous Volterra equation admits one, and one only, solution, and the latter cannot be zero (the zero map always solves, because A is linear). A direct computation shows:

$$|(Af)(x)| = \left|\int_a^x K(x,y)f(y)dy\right| \leq M(x-a)||f||_\infty .$$

The first iteration gives

$$|(A^2 f)(x)| \leq M^2 \frac{(x-a)^2}{2}||f||_\infty ,$$

and, after $n-1$ steps,

$$|(A^n f)(x)| \leq M^n \frac{(x-a)^n}{n!}||f||_\infty .$$

Hence:

$$||A^n f||_\infty \leq M^n \frac{(b-a)^n}{n!}||f||_\infty ,$$

and so:

$$||A^n|| \leq M^n \frac{(b-a)^n}{n!} .$$

For n large enough then, whatever a, b, M, are, we have:

$$M^n \frac{(b-a)^n}{n!} < 1 .$$

Thus for some positive $\lambda < 1$:

$$||A^n f - A^n f'||_\infty \leq \lambda ||f - f'||_\infty ,$$

and, by the fixed-point theorem, the homogeneous Volterra equation on $C([a,b])$ only admits the trivial solution.

Consequently *the operator A of (2.47) cannot admit eigenvalues different from zero*. In fact the eigenvalue equation for A,

$$A\psi = \lambda\psi \quad \text{for some } \lambda \in \mathbb{C} \text{ and some } \psi \neq 0, \tag{2.48}$$

is equivalent to:

$$\frac{1}{\lambda}A\psi = \psi \quad \lambda \in \mathbb{C}\setminus\{0\},\ \psi \neq 0$$

if $\lambda \neq 0$. And $\lambda^{-1}A$ is a Volterra operator associated to the integral kernel $\lambda^{-1}K(x,y)$. Therefore the theorem may be used on A to give $\psi = 0$. Since the eigenvalue was not allowed to vanish, (2.48) has no solution.

This result will be generalised in Chapter 4 to Hilbert spaces. It will bear an important consequence in the study of Volterra's *inhomogeneous* equation, once we have proved *Fredholm's theorem* on integral equations.

(2) Consider the existence and uniqueness problem for a continuous map $y = f(x)$ when we only know an implicit relation of the type $F(x, y(x)) = 0$, for some given and sufficiently regular function F. We discuss a simplified version of a result that is conventionally known as either *Dini's theorem*, *implicit function theorem* or *inverse funtion theorem* [CoFr98II, Ser94II]. The point is to see the Banach-Caccioppoli theorem in action. Suppose we are given a function $F : [a,b] \times \mathbb{R} \to \mathbb{R}$, $a < b$, that is continuous and admits partial y-derivative such that $0 < m \leq |\frac{\partial F}{\partial y}| \leq M < +\infty$, $(x,y) \in [a,b] \times \mathbb{R}$.

We want to show that there exists, unique, a continuous map $f : [a,b] \to \mathbb{R}$ such that:

$$F(x, f(x)) = 0 \quad \text{for any } x \in [a,b].$$

The idea is to define a contraction map $G : C([a,b]) \to C([a,b])$ having f as fixed point. To this end set:

$$(G(\psi))(x) := \psi(x) - \frac{1}{M} F(x, \psi(x)) \quad \text{for any } \psi \in C([a,b]), x \in [a,b].$$

This is well defined on $C([a,b])$, and if it contracts then its unique fixed point f satisfies:

$$f(x) = f(x) - \frac{1}{M} F(x, f(x)) \quad \text{for any } x \in [a,b].$$

In other words:

$$F(x, f(x)) = 0 \quad \text{for any } x \in [a,b],$$

so G is what we are after. But G is easily a contraction by the mean value theorem:

$$(G(\psi))(x) - (G(\psi'))(x) = \psi(x) - \psi'(x) - \frac{1}{M}\left(F(x, \psi(x)) - F(x, \psi'(x))\right),$$

so for some number z between $\psi(x)$ and $\psi'(x)$:

$$(G(\psi))(x) - (G(\psi'))(x) = \psi(x) - \psi'(x) - \frac{1}{M}(\psi(x) - \psi'(x)) \frac{\partial F}{\partial y}|_{(x,z)},$$

and therefore:

$$|(G(\psi))(x) - (G(\psi'))(x)| \leq |\psi(x) - \psi'(x)| \left| 1 - \frac{1}{M} \frac{\partial F}{\partial y}|_{(x,z)} \right|.$$

Because the derivative's range is inside the positive interval $[m, M]$, we have:

$$||G(\psi) - G(\psi')||_\infty \leq ||\psi - \psi'||_\infty (1 - \frac{m}{M}).$$

Now, by assumption $(1 - \frac{m}{M}) < 1$, so G is indeed a contraction. ■

2.6.2 Application of the fixed-point theorem: local existence and uniqueness for systems of differential equations

The most important application, by far, of the fixed-point theorem is certainly the theorem of local existence and uniqueness for first-order systems of differential equations in *normal form* (where the highest derivative, here the first, is written alone on one side of the equation, as in (2.50) below). This result extends easily to global solutions and higher-order systems [CoFr98I, CoFr98II].

For this we need a preliminary notion. From now on $\mathbb{K}$ will be the complete field $\mathbb{R}$, or possibly $\mathbb{C}$, and $|| \, ||_{\mathbb{K}^p}$ the standard norm on $\mathbb{K}^p$.

Definition 2.110. *Let $r \geq 0$ and $n,m > 0$ be given natural numbers, $\Omega \subset \mathbb{K}^r \times \mathbb{K}^n$ a non-empty open set. A function $F : \Omega \to \mathbb{K}^m$ is* **locally Lipschitz** *(***in the variable** *$x \in \mathbb{K}^n$ for $r > 0$), if for any $p \in \Omega$ there is a constant $L_p \geq 0$ such that:*

$$||F(z,x) - F(z,x')||_{\mathbb{K}^m} \leq L_p||x - x'||_{\mathbb{K}^n}\,, \quad \text{for any pair } (z,x),(z,x') \in O_p, \tag{2.49}$$

$O_p \ni p$ being an open set.

Any C^1 map $F : \Omega \to \mathbb{K}^m$ is locally Lipschitz in the variable x, as we shall shortly see, but first the theorem.

Theorem 2.111 (Local existence and uniqueness for systems of ODEs of order one). *Let $f : \Omega \to \mathbb{K}^n$ be a continuous and locally Lipschitz map in $x \in \mathbb{K}^n$ on the open set $\Omega \subset \mathbb{R} \times \mathbb{K}^n$. Given the first-order initial value problem (in normal form):*

$$\begin{cases} \dfrac{dx}{dt} &= f(t, x(t))\,, \\ x(t_0) &= x_0 \end{cases} \tag{2.50}$$

with $(t_0, x_0) \in \Omega$, there is an open interval $I \ni t_0$ on which (2.50) has a unique solution, necesarily belonging in $C^1(I)$.

Proof. Notice, to being with, that any solution $x = x(t)$ to (2.50) is necessarily C^1. Namely, it is continuous as the derivative exists, but directly from $\frac{dx}{dt} = f(t,x(t))$ we infer $\frac{dx}{dt}$ must be continuous, because the equation's right-hand side is a composite function of continuous maps in t. Now, suppose $x : I \to \mathbb{K}^n$ is differentiable and that (2.50) holds. By the fundamental theorem of calculus, by integrating (2.50) (the derivative of $x(t)$ is continuous) $x : I \to \mathbb{K}^n$ must satisfy

$$x(t) = x_0 + \int_{t_0}^{t} f(\tau, x(\tau))\, d\tau\,, \quad \text{for any } t \in I. \tag{2.51}$$

Conversely, if $x : I \to \mathbb{K}^n$ is continuous and satisfies (2.51), again the fundamental theorem of calculus (f is continuous) tells $x = x(t)$ is differentiable and guarantees (2.50).

Therefore the continuous maps $x = x(t)$ defined on an open interval $I \ni t_0$ that solve the integral equation (2.51) are precisely the solutions to (2.50) defined over I. So instead of solving (2.50) we can solve the equivalent integral problem (2.51).

To prove existence, fix once and for all a relatively compact open set $Q \ni (t_0, x_0)$ with $\overline{Q} \subset \Omega$. Take $\overline{Q}$ small enough to have f locally Lipschitz in x. The standard norm on $\mathbb{K}^n$ will be written $|| \ ||$, and we shall use:

(i) $0 \leq M := \max\{||f(t,x)|| \mid (t,x) \in \overline{Q}\}$;
(ii) $L \geq 0$ the constant such that $||f(t,x) - f(t,x')|| \leq L||x - x'||$, $(t,x), (t,x') \in \overline{Q}$;
(iii) $B_\varepsilon(x_0) := \{x \in \mathbb{K}^n \mid ||x - x_0|| \leq \varepsilon\}$ for $\varepsilon > 0$.

Consider the closed interval $J_\delta = [t_0 - \delta, t_0 + \delta]$, $\delta > 0$ and the Banach space (Proposition 2.17) $(C(J_\delta; \mathbb{K}^n), || \ ||_\infty)$ of continuous maps $X : J_\delta \to \mathbb{K}^n$. On this space define G, that maps to any function X a function $G(X)$:

$$G(X)(t) := x_0 + \int_{t_0}^{t} f(\tau, X(\tau))\, d\tau\,, \quad \text{for any } t \in J_\delta\,.$$

Note $G(X) \in C(J_\delta; \mathbb{K}^n)$ for $X \in C(J_\delta; \mathbb{K}^n)$ by the continuity of the integral with respect to the upper limit, when the integrand is continuous. We claim G is a contraction map on a closed subset of $C(J_\delta; \mathbb{K}^n)$: [2]

$$\mathsf{M}_\varepsilon^{(\delta)} := \{X \in C(J_\delta; \mathbb{K}^n) \mid ||X(t) - x_0|| \leq \varepsilon\,, \forall t \in J_\delta\}$$

if $0 < \delta < \min\{\varepsilon/M, 1/L\}$, and $\delta, \varepsilon > 0$ are so small that $J_\delta \times B_\varepsilon(x_0) \subset Q$. (Henceforth $\varepsilon > 0$ and $\delta > 0$ will be assumed to satisfy $J_\delta \times B_\varepsilon(x_0) \subset Q$.) With $X \in \mathsf{M}_\varepsilon^{(\delta)}$ we have:

$$||G(X)(t) - x_0|| \leq \left|\left| \int_{t_0}^{t} f(\tau, X(\tau))\, d\tau \right|\right| \leq \int_{t_0}^{t} ||f(\tau, X(\tau))||\, d\tau \leq \int_{t_0}^{t} M d\tau \leq \delta M\,.$$

Therefore $G(\mathsf{M}_\varepsilon^{(\delta)}) \subset \mathsf{M}_\varepsilon^{(\delta)}$ for $0 < \delta < \varepsilon/M$. If $X, X' \in \mathsf{M}_\varepsilon^{(\delta)}$ then for all $t \in J_\delta$:

$$\begin{aligned} G(X)(t) - G(X')(t) &= \int_{t_0}^{t} \left[f(\tau, X(\tau)) - f\left(\tau, X'(\tau)\right)\right] d\tau\,, \\ \left|\left|G(X)(t) - G(X')(t)\right|\right| &\leq \left|\left| \int_{t_0}^{t} \left[f(\tau, X(\tau)) - f\left(\tau, X'(\tau)\right)\right] d\tau \right|\right| \\ &\leq \int_{t_0}^{t} \left|\left|f(\tau, X(\tau)) - f\left(\tau, X'(\tau)\right)\right|\right| d\tau\,. \end{aligned}$$

But we have the Lipschitz bound

$$||f(t,x) - f(t,x')|| < L||x - x'||\,,$$

so:

$$\left|\left|G(X)(t) - G(X')(t)\right|\right| \leq L \int_{t_0}^{t} \left|\left|X(\tau) - X'(\tau)\right|\right| d\tau \leq \delta L ||X - X'||_\infty\,.$$

[2] $\mathsf{M}_\varepsilon^{(\delta)} = \{X \in C(J_\delta; \mathbb{K}^n) \mid ||X - X_0||_\infty \leq \varepsilon\}$, where X_0 is here the constant *map* equal to x_0 on J_δ. Thus $\mathsf{M}_\varepsilon^{(\delta)}$ is the closure of the open ball of radius ε centred at X_0 inside $C(J_\delta; \mathbb{K}^n)$.

Taking the supremum on the left:

$$\big|\big|G(X) - G(X')\big|\big|_\infty \le \delta L ||X - X'||_\infty .$$

If, additionally, $\delta < 1/L$, it follows that $G : \mathsf{M}_\varepsilon^{(\delta)} \to \mathsf{M}_\varepsilon^{(\delta)}$ is a contraction on the closed set $\mathsf{M}_\varepsilon^{(\delta)}$. By Theorem 2.108 G has a fixed point, which is a continuous map $x = x(t) \in \mathbb{K}^n$, $t \in J_\delta$, that solves (2.51) by definition of G. Restricting x to the open $I := (t_0 - \delta, t_0 + \delta)$ gives a solution to the initial value problem (2.50).

As for uniqueness, take another solution $x' = x'(t)$ to (2.51) on $I := (t_0 - \delta, t_0 + \delta)$, *a priori* distinct from $x = x(t)$. For any closed $J_{\delta'} := [t_0 - \delta', t_0 + \delta']$, $0 < \delta' < \delta$, $G : \mathsf{M}_\varepsilon^{(J_{\delta'})} \to \mathsf{M}_\varepsilon^{(J_{\delta'})}$ is by construction still a contracting map, and $x' = x'(t)$ a fixed point of it; therefore x' coincides with $x = x(t)$ restricted to $J_{\delta'}$, by uniqueness. (In particular, the restriction of $\mathbf{x}'$ to $J_{\delta'}$ belongs to the complete metric space $\mathsf{M}_\varepsilon^{(J_{\delta'})}$ because we saw $||G(\mathbf{x}') - \mathbf{x}_0||_\infty \le \delta' M < \varepsilon$, since $G(\mathbf{x}') = \mathbf{x}'$.) But since δ' is arbitrary in $(0, \delta)$, the two solutions coincide on $I = (t_0 - \delta, t_0 + \delta)$. □

Just for completeness' sake we remark that the previous theorem holds when f is C^1, because of the following elementary fact.

We adopt the usual notation $x = (x_1, \dots, x_n)$, $z = (z_1, \dots, z_l)$ and $F(z,x) = (F_1(z,x), \dots, F_m(z,x))$ for an arbitrary $F : \Omega \to \mathbb{R}^m$ with $\Omega = A \times B$, $A \subset \mathbb{R}^l$ and $B \subset \mathbb{R}^n$ non-empty open sets.

Proposition 2.112. *Consider $\Omega = A \times B$, $A \subset \mathbb{R}^l$ and $B \subset \mathbb{R}^n$ non-empty open sets. The map $F : \Omega \to \mathbb{R}^m$ is locally Lipschitz in x if, for every $z \in A$, the functions $B \ni x \mapsto F_k(z,x)$ admit first derivative, and if the partial derivatives, as (z,x) varies, are continuous on Ω.*

Proof. Take $q = (z_0, x_0) \in \Omega$ and let $B' \subset \mathbb{R}^l$, $B \subset \mathbb{R}^n$ be open balls centred at z_0, x_0, with $\overline{B'} \times \overline{B} \subset \Omega$. Then $x(t) = p + t(r - p) \in \overline{B}$, for $t \in [0,1]$ and $p, r \in \overline{B}$. Fix $z \in B'$ and invoke the mean value theorem for $[0,1] \ni t \mapsto F_k(z, x(t))$, to the effect that

$$F_k(z,r) - F_k(z,p) = F_k(z,x(1)) - F_k(z,x(0)) = \sum_{j=1}^n (r_j - p_j) \left.\frac{\partial F_k}{\partial x_j}\right|_{(z,x(\xi))} ,$$

where $(z, x(\xi)) \in \overline{B'} \times \overline{B}$. Schwarz's inequality then gives:

$$\begin{aligned} |F_k(z,r) - F_k(z,p)| &\le \sqrt{\sum_{j=1}^n |r_j - p_j|^2} \sqrt{\sum_{i=1}^n \left| \left.\frac{\partial F_k}{\partial x_i}\right|_{(z,x(\xi))} \right|^2} \\ &\le ||r - p|| \sqrt{\sum_{i=1}^n \left| \left.\frac{\partial F_k}{\partial x_i}\right|_{(z,x(\xi))} \right|^2} \\ &\le M_k ||r - p|| \qquad \text{for } (z,r), (z,p) \in \overline{B'} \times \overline{B} , \end{aligned}$$

and such $M_k < +\infty$ exists since the radicand is continuous on the compact set $\overline{B'} \times \overline{B}$. Because $B' \times B$ is an open neighbourhood of the generic point $(z_0, x_0) \in \Omega$, the map F is locally Lipschitz in x:

$$||F(z,x_1) - F(z,x_2)|| \leq \sqrt{\sum_{k=1}^{m} M_k^2}\, ||x_1 - x_2|| \quad \text{for } (z,x_1),(z,x_2) \in B' \times B. \qquad \square$$

Remarks 2.113. This particular proof of the theorem requires the local Lipschitz condition for f in (2.50) in order to use the fixed-point theorem. As a matter of fact, this is not necessary to grant *existence*. A more general existence result, due to Peano, can be proved (using Theorem 2.21 of Arzelà–Ascoli) if one only assumes the continuity of f [KoFo99]. In general, though, the absence of the Lipschitz condition undermines the solution's uniqueness, as the following classical counterexamples makes clear: consider

$$\frac{dx}{dt} = f(x(t)), \quad x(0) = 0$$

where $f : \mathbb{R} \to \mathbb{R}$, defined as $f(x) = 0$ on $x \leq 0$ and $f(x) = \sqrt{x}$ for $x > 0$, is continuous but not locally Lipschitz at $x = 0$. The Cauchy problem admits *two* solutions (both maximal):

(1) $x_1(t) = 0$, for any $t \in \mathbb{R}$;
(2) $x_2(t) = 0$ for $t \leq 0$ and $x_2(t) = t^2/4$ on $t > 0$. ■

Exercises

2.1. Prove that any seminorm p satisfies $p(x) = p(-x)$.

2.2. Let $f : K \to \mathsf{X}$ be a continuous map from the compact set K to the normed space X. Show f is bounded, i.e. there exists $M \geq 0$ such that $||f(k)|| \leq M$ for any $k \in K$.

Hint. Adapt the proof of Proposition 1.21.

2.3. Prove that if S denotes a vector space of bounded maps from X to $\mathbb{C}$ (or to $\mathbb{R}$), then

$$\mathsf{S} \ni f \mapsto ||f||_\infty := \sup_{x \in \mathsf{X}} |f(x)|$$

defines a norm on S.

2.4. Prove that the spaces $L(\mathsf{X})$, of bounded complex functions, and $M_b(\mathsf{X})$, of measurable and bounded complex functions (cf. Examples 2.26), over a topological space X are Banach spaces for the norm $||\ ||_\infty$.

Solution. We prove the claim for $M_b(\mathsf{X})$, the other one being exactly the same. We will show that an arbitrary Cauchy sequence $\{f_n\}_{n\in\mathbb{N}} \subset M_b(\mathsf{X})$ converges uniformly to some $f \in M_b(\mathsf{X})$. By assumption the numerical sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ is Cauchy, for any $x \in \mathsf{X}$. Then there exists $f : \mathsf{X} \to \mathbb{C}$ such that $f_n(x) \to f(x)$, as $n \to +\infty$, for any $x \in \mathsf{X}$. This function will be measurable because it arises as limit of measurable maps. We are left to prove f is bounded and $f_n \to f$ uniformly. Start from the latter. Since

$$|f(x) - f_m(x)| = \lim_{n\to+\infty} |f_n(x) - f_m(x)| \leq \lim_{n\to+\infty} ||f_n - f_m||_\infty ,$$

and using the fact that the initial sequence is Cauchy for $||\ ||_\infty$, we have that for any $\varepsilon > 0$ there is N_ε such that:

$$\lim_{n\to+\infty} ||f_n - f_m||_\infty < \varepsilon \quad \text{for } m > N_\varepsilon .$$

Hence:

$$|f(x) - f_m(x)| < \varepsilon \quad \text{for } m > N_\varepsilon \text{ and any } x \in \mathsf{X}.$$

I.e. $||f - f_m||_\infty \to 0$ as $m \to +\infty$, as required. Now boundedness is obvious:

$$\sup_{x\in\mathsf{X}} |f(x)| \leq \sup_{x\in\mathsf{X}} |f(x) - f_m(x)| + \sup_{x\in\mathsf{X}} |f_m(x)| < \varepsilon + ||f_m||_\infty < +\infty .$$

2.5. Show that the Banach spaces $(L(\mathsf{X}), ||\ ||_\infty)$ and $(M_b(\mathsf{X}), ||\ ||_\infty)$ (cf. Examples 2.26) are Banach algebras with unit.

Sketch. The unit is clearly the constant map 1. The property $||f \cdot g||_\infty \leq ||f||_\infty ||g||_\infty$ follows from the definition of $||\ ||_\infty$, and the remaining conditions are easy.

2.6. Prove the space $C_0(\mathsf{X})$ of continuous complex functions on X that vanish at infinity (cf. Examples 2.26) is a Banach algebra for $||\ ||_\infty$. Explain in which circumstances the algebra has a unit.

Solution. We take a Cauchy sequence $\{f_n\}_{n\in\mathbb{N}} \subset C_0(\mathsf{X})$ and prove it converges uniforly to $f \in C_0(\mathsf{X})$. By hypothesis the numerical sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ is Cauchy for any $x \in \mathsf{X}$. Therefore there exists a function $f : \mathsf{X} \to \mathbb{C}$ such that $f_n(x) \to f(x)$ for any $x \in \mathsf{X}$, as $n \to +\infty$. Th proof that $||f - f_n||_\infty \to 0$, $n \to +\infty$, goes exactly as in Exercise 2.4. Since continuity is preserved by uniform limits, there remains to show $f \in C_0(\mathsf{X})$. Given $\varepsilon > 0$, pick n such that $||f - f_n|| < \varepsilon/2$, and choose a compact set $K_\varepsilon \subset \mathsf{X}$ so that $|f_n(x)| < \varepsilon/2$ for $x \in \mathsf{X} \setminus K_\varepsilon$. By construction

$$|f(x)| \leq |f(x) - f_n(x)| + |f_n(x)| < \varepsilon, \quad x \in \mathsf{X} \setminus K_\varepsilon .$$

The Banach space thus found is a Banach algebra for the familiar operations, as one proves without difficulty.

If the unit is present, it must be the constant map 1. If X is compact, the function 1 belongs to the space. But if X is not compact, then 1 cannot be in X, because the elements of $C_0(\mathsf{X})$ can be shrunk arbitrarily outside compact subsets, and no constant map does that.

2.7. Prove the space $C_b(\mathsf{X})$ of continuous and bounded complex functions (see Examples 2.26) on X is a Banach space for $||\ ||_\infty$ and a Banach algebra with unit.

2.8. Prove that in Proposition 2.16 the converse implication holds as well. In other words the proposition may be rephrased like this:

Let $(\mathsf{X}, ||\ ||)$ *be a normed space. Every absolutely converging series* $\sum_{n=0}^{+\infty} x_n$ *in* X *(i.e.* $\sum_{n=0}^{+\infty} ||x_n|| < +\infty$*) converges in* X *iff* $(\mathsf{X}, ||\ ||)$ *is a Banach space.*

Solution. Take an absolutely convergent series $\sum_{n=0}^{+\infty} x_n$ in X. The partial sums of the norms have to be a Cauchy sequence, i.e. for any $\varepsilon > 0$ we have $M_\varepsilon > 0$ with

$$\left| \sum_{j=0}^{n} ||x_j|| - \sum_{j=0}^{m} ||x_j|| \right| < \varepsilon , \quad \text{for } n, m > M_\varepsilon .$$

Supposing $n \geq m$:

$$\left| \sum_{j=m+1}^{n} ||x_j|| \right| < \varepsilon , \quad n, m > M_\varepsilon .$$

Therefore:

$$\left|\left| \sum_{j=0}^{n} x_j - \sum_{j=0}^{m} x_j \right|\right| = \left|\left| \sum_{j=m+1}^{n} x_j \right|\right| \leq \sum_{j=m+1}^{n} ||x_j|| < \varepsilon, \quad n, m > M_\varepsilon .$$

We proved the sequence of partial sums $\sum_{j=0}^{n} x_n$ is Cauchy. But the space is complete, so the series converges to a point in X.

2.9. Prove the space $C_c(\mathsf{X})$ of complex functions with compact support (cf. Examples 2.26) on X is *not*, in general, a Banach space for $||\ ||_\infty$, and neither is it dense in $C_b(\mathsf{X})$ if X is not compact.

Outline of proof. For the first statement we need to exhibit a counterexample for $\mathsf{X} = \mathbb{R}$. Consider for instance the sequence $f_n : \mathbb{R} \to \mathbb{C}$ of continuous maps, $n = 1, 2, \ldots,$: $f_n(x) := \frac{\sin x}{x}$ for $0 < |x| < 2n\pi$, $f_n(0) = 1$ and $f_n(x) = 0$ at other points of $\mathbb{R}$. The sequence evidently converges pointwise to the continuous map defined as $\frac{\sin x}{x}$ on $\mathbb{R} \setminus \{0\}$ and set to 1 at the origin. It is easy to convince ourselves the convergence is uniform too. But the limit function has no compact support. As for the second part, note that any constant map $c \neq 0$ belongs in $C_c(\mathsf{X})$. But if X is not compact, then $||f - c||_\infty \geq |c| > 0$ for any function $f \in C_c(\mathsf{X})$ because of the values attained outside the support of f.

2.10. Given a compact space K and a Banach space B, let $C(K; \mathsf{B})$ be the space of continuous maps $f : K \to \mathsf{B}$ in the norm topologies of domain and codomain. Define

$$||f||_\infty := \sup_{x \in K} ||f(x)|| \quad f \in C(K; \mathsf{B}) ,$$

where on the norm on the right is the one on B. Show this norm is well defined, and that it turns $C(K; \mathsf{B})$ into a Banach space.

Hint. Keep in mind Exercise 2.2 and adjust the proof of Proposition 2.17.

2.11. Let $(\mathfrak{A}, \circ)$ be a Banach algebra without unit. Consider the direct sum $\mathfrak{A} \oplus \mathbb{C}$ and define the product:

$$(x,c)\cdot(y,c') := (x \circ y + cy + c'x, cc'), \quad (x',c'),(x,c) \in \mathfrak{A} \oplus \mathbb{C}$$

and the norm:

$$||(x,c)|| := ||x|| + |c|, \quad (x,c) \in \mathfrak{A} \oplus \mathbb{C}.$$

Show that the vector space $\mathfrak{A} \oplus \mathbb{C}$ with these product and norm becomes a Banach algebra with unit.

2.12. Take a Banach algebra $\mathfrak{A}$ with unit $\mathbb{I}$ and an element $a \in \mathfrak{A}$ with $||a|| < 1$. Prove that the series $\sum_{n=0}^{+\infty}(-1)^n a^{2n}$, $a^0 := \mathbb{I}$, converges in the topology of $\mathfrak{A}$. What is the sum?

Hint. Show the series of partial sums is a Cauchy series. The sum is $(\mathbb{I} + a^2)^{-1}$.

2.13. (*Hard*) Prove Hölder's inequality:

$$\int_X |f(x)g(x)|d\mu(x) \leq \left(\int_{\mathsf{X}} |f(x)|^p d\mu(x)\right)^{1/p} \left(\int_{\mathsf{X}} |g(x)|^q d\mu(x)\right)^{1/q}$$

where $p,q > 0$ satisfy $1 = \frac{1}{p} + \frac{1}{q}$, f and g are measurable and μ is a positive measure on X.

Solution. Define $I := \int_X |f(x)|\,|g(x)|d\mu(x)$, $A := (\int_{\mathsf{X}} |f(x)|^p d\mu(x))^{1/p}$ and $B := (\int_{\mathsf{X}} |g(x)|^q d\mu(x))^{1/q}$. If either A or B is zero or infinite (conventionally, $\infty \cdot 0 = 0 \cdot \infty = 0$), the inequality is trivial. So let us assume $0 < A, B < +\infty$ and define $F(x) := |f(x)|/A$, $G(x) := |g(x)|/B$. Thus

$$\ln(F(x)G(x)) = \frac{1}{p}\ln(F(x)^p) + \frac{1}{q}\ln(G(x)^q) \leq \ln\left(\frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q\right),$$

because the logarithm is a convex function. Exponentiating gives

$$F(x)G(x) \leq \frac{1}{p}F(x)^p + \frac{1}{q}G(x)^q\,.$$

Integrating the above, and noting that the right-hand-side integral is $1/p + 1/q = 1$ we recover Hölder's inequality in the form:

$$\frac{\int_X |f(x)g(x)|d\mu(x)}{(\int_{\mathsf{X}} |f(x)|^p d\mu(x))^{1/p} (\int_{\mathsf{X}} |g(x)|^q d\mu(x))^{1/q}} \leq 1\,.$$

2.14. (*Hard*) Making use of Hölder's inequality, prove Minkowski's inequality:

$$\left(\int_X |f(x)+g(x)|^p d\mu(x)\right)^{1/p} \leq \left(\int_X |f(x)|^p d\mu(x)\right)^{1/p} + \left(\int_X |g(x)|^p d\mu(x)\right)^{1/p}$$

where $p \geq 1$, f and g are measurable and μ a positive measure on X.

Solution. Define $I := \int_{\mathsf{X}} |f(x)+g(x)|^p d\mu(x)$, $A := \left(\int_{\mathsf{X}} |f(x)|^p d\mu(x)\right)^{1/p}$ and $B := \left(\int_{\mathsf{X}} |g(x)|^p d\mu(x)\right)^{1/p}$. The inequality is trivial if $p=1$ or if either of A, B are infinite. So we assume $p>1$, $A,B<+\infty$. Then I must be finite too, because $(a+b)^p \leq 2^p(a^p+b^p)$ for any $a,b\geq 0$ and $p\geq 1$[3]. Minkowski's inequality is trivial also when $I=0$, so we consider only $p>1$, $A,B<+\infty$, $0<I<+\infty$. Note $|f+g|^p = |f|\,|f+g|^{p-1}+|g|\,|f+g|^{p-1}$. Using Hölder's inequality on each summand on the right we have:

$$\int_{\mathsf{X}} |f(x)+g(x)|^p d\mu(x) \leq \left(\left(|f(x)+g(x)|^{(p-1)q} d\mu(x)\right)^{1/q}\right)$$

$$\times\left(\left(\int_{\mathsf{X}} |f(x)|^p d\mu(x)\right)^{1/p} + \left(\int_{\mathsf{X}} |g(x)|^p d\mu(x)\right)^{1/p}\right),$$

where $1=\frac{1}{p}+\frac{1}{q}$. This last inequality can be written $I\leq I^{1/q}(A+B)$, dividing which by $I^{1/q}$ produces $I^{1/p}\leq A+B$, i.e. Minkowski's inequality.

2.15. Take two finite-dimensional normed spaces X,Y and consider $T\in\mathfrak{L}(\mathsf{X},\mathsf{Y})=\mathfrak{B}(\mathsf{X},\mathsf{Y})$. Fix bases in X and Y, so to represent T by the matrix $M(T)$. Show that one can choose bases for the dual spaces X', Y' so that the operator T' is determined by the transpose matrix $M(T)^t$.

2.16. Prove Proposition 2.66.

2.17. Consider the space $\mathfrak{B}(\mathsf{X})$ for X normed, and prove the strong topology is finer than the weak topology (put loosely: weakly open sets are strongly open), and the uniform topology is finer than the strong one.

2.18. Prove Propositions 2.70, 2.71, 2.72 and 2.73.

2.19. In a normed space X prove that if $\{x_n\}_{n\in\mathbb{N}}\subset\mathsf{X}$ tends to $x\in\mathsf{X}$ *weakly* (cf. Proposition 2.70), the set $\{x_n\}_{n\in\mathbb{N}}$ is bounded.

Hint. Use Corollary 2.60.

2.20. If B is a Banach space and $T,S\in\mathfrak{B}(\mathsf{B})$ show:
(i) $(TS)'=S'T'$;
(ii) $(T')^{-1}=(T^{-1})'$ if T is bijective.

2.21. Prove that if X and Y are reflexive Banach spaces, and $T\in\mathfrak{B}(\mathsf{X},\mathsf{Y})$, then $(T')'=T$.

2.22. If X is normed, the function that maps $(T,S)\in\mathfrak{B}(\mathsf{X})\times\mathfrak{B}(\mathsf{X})$ to $TS\in\mathfrak{B}(\mathsf{X})$ is continuous in the uniform topology. What can be said regarding the strong and weak topologies?

[3] This inequality descends from $(a+b)\leq 2\max\{a,b\}$, whose pth power reads $(a+b)^p\leq 2^p\max\{a^p,b^p\}\leq 2^p(a^p+b^p)$.

Solution. For both topologies the map is separately continuous in either argument, but not continuous as a function of two variables, in general.

2.23. If we define an isomorphism of normed spaces as a continuous linear map with continuous inverse, does an isomorphism preserve completeness?

Hint. Extend Proposition 2.101 to the case of a continuous linear map with continuous inverse, between normed spaces.

2.24. Using weak equi-boundedness, prove this variant of the Banach–Steinhaus Theorem 2.58.

Let X *be a Banach space,* Y *a normed space on the same field* $\mathbb{C}$*, or* $\mathbb{R}$*. Suppose the family of operators* $\{T_\alpha\}_{\alpha\in A} \subset \mathfrak{B}(\mathsf{X},\mathsf{Y})$ *satisfies:*

$$\sup_{\alpha\in A} |f(T_\alpha x)| < +\infty \quad \textit{for any } x \in \mathsf{X},\ f \in \mathsf{Y}'\,.$$

Then there exists a uniform bound $K \geq 0$*:*

$$||T_\alpha|| \leq K \ \textit{ for any } \alpha \in A\,.$$

Solution. Referring to Corollary 2.55, for any given $x \in \mathsf{X}$ define $F_{\alpha,x} := \mathfrak{I}(T_\alpha x) \in (\mathsf{Y}')'$. Then

$$\sup_{\alpha\in A} |F_{\alpha,x}(f)| < +\infty \quad \text{for any } f \in \mathsf{Y}'\,.$$

As Y' is complete, we can use Theorem 2.58 to infer the existence, for any $x \in \mathsf{X}$, of a $K_x \geq 0$ that bounds uniformly the familiy $F_{\alpha,x} : \mathsf{Y}' \to \mathbb{C}$:

$$||F_{\alpha,x}|| \leq K_x \ \text{ for any } \alpha \in A\,.$$

But $\mathfrak{I}$ is isometric, so:

$$||T_\alpha(x)|| \leq K_x \ \text{ for any } \alpha \in A$$

and hence

$$\sup_{\alpha\in A} ||T_\alpha x|| < +\infty \quad \text{for any } x \in \mathsf{X}\,.$$

The Banach–Steinhaus Theorem 2.58 ends the proof.

2.25. Let K be compact, X a Banach space, and equip $\mathfrak{B}(\mathsf{X})$ with the strong topology. Prove that any continuous map $f : K \to \mathfrak{B}(\mathsf{X})$ belongs to $C(K;\mathfrak{B}(\mathsf{X}))$. (The latter is a Banach space, defined in Exercise 2.10, if we view $\mathfrak{B}(\mathsf{X})$ as a Banach space.)

Solution. We must prove

$$\sup_{k\in K} ||f(k)|| < +\infty$$

where on the left we used the operator norm of $\mathfrak{B}(\mathsf{X})$. As f is continuous in the strong topology, for any given $x \in \mathsf{X}$ the map $K \ni k \mapsto f(k)x \in \mathsf{X}$ is continuous. If we fix $x \in \mathsf{X}$, by Exercise 2.2 there is $M_x \geq 0$ such that:

$$\sup_{k\in K} ||f(k)x||_{\mathsf{X}} < M_x\,.$$

The Banach–Steinhaus Theorem 2.58 ends the proof.

This page intentionally left blank

3

Hilbert spaces and bounded operators

There's no such thing as a deep theorem, but only theorems we haven't understood very well.

Nicholas P. Goodman

With this section we introduce the first mathematical notions relative to Hilbert spaces that we will use to build the matematical foundations of Quantum Mechanics.

A good part of the chapter is devoted to *Hilbert bases* (also known as *complete orthonormal system*), which we treat in full generality and not assuming the Hilbert space be separable. Before that, we discuss the paramount result in the theory of Hilbert spaces: *Riesz's representation theorem*, according to which there is a natural anti-isomorphism between a Hilbert space and its dual.

The third part studies the notion of *adjoint operator* (to a bounded operator), introduced by means of Riesz's theorem, and all its fundamental consequences in the theory of bounded operators. In particular, we introduce *C^*-algebras* (and operator C^*-algebras) and the examples will serve to present *von Neumann algebras* and the famous *double commutant theorem*. Here we define *self-adjoint*, *unitary* and *normal* operators, and examine the basic properties.

Section four is entirely dedicated to *orthogonal projectors* and their main properties. We also introduce the useful notion of *partial isometry*.

The fifth section is concerned with the important *polar decomposition theorem* for bounded operators defined on the whole Hilbert space. The notion of *positive square root* of a bounded operator is used as technical tool.

The elementary theory of the *Fourier* and *Fourier-Plancherel transforms*, object of the last section, is introduced very rapidly and with, alas, no reference to Schwartz's distributions (for this see [Rud91, ReSi80, Vla81]).

3.1 Elementary notions, Riesz's theorem and reflexivity

The present section deals with the basics of Hilbert spaces, starting from the elementary definitions of *Hermitian inner product* and *Hermitian inner product space*.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_3, © Springer-Verlag Italia 2013

3.1.1 Inner product spaces and Hilbert spaces

Definition 3.1. *If* X *is a complex vector space, a map* $S : \mathsf{X} \times \mathsf{X} \to \mathbb{C}$ *si called a* **Hermitian inner product***, or* **Hermitian scalar product***, and* (X, S) *an* **inner product space***, when:*

H0. $S(u,u) \geq 0$ *for any* $u \in \mathsf{X}$*;*
H1. $S(u, \alpha v + \beta w) = \alpha S(u,v) + \beta S(u,w)$ *for any* $\alpha, \beta \in \mathbb{C}$ *and* $u,v,w \in \mathsf{X}$*;*
H2. $S(u,v) = \overline{S(v,u)}$ *for any* $u,v \in \mathsf{X}$*;*
H3. $S(u,u) = 0 \Rightarrow u = 0$*, for any* $u \in \mathsf{X}$.

If **H0***,* **H1***,* **H2** *hold and* **H3** *does not, S is a* **Hermitian semi-inner product**.
Two vectors $u, v \in \mathsf{X}$ *are* **orthogonal***, written* $u \perp v$*, if* $S(u,v) = 0$*. In this case u is said* **orthogonal** *(or* **normal***) to v and v is* **orthogonal** *(or* **normal***) to u.*
If $\varnothing \neq K \subset \mathsf{X}$*, the* **orthogonal space to** K *is:*

$$K^{\perp} := \{u \in \mathsf{X} \mid u \perp v \text{ for any } v \in K\}\,.$$

Remark 3.2. (1) H1 and **H2** imply that S is **antilinear** in the first argument:

$$S(\alpha v + \beta w, u) = \overline{\alpha} S(v,u) + \overline{\beta} S(w,u) \quad \text{for any } \alpha, \beta \in \mathbb{C},\, u,v,w \in \mathsf{X}.$$

(2) From **H2** u is orthogonal to v if and only if v is orthogonal to u.
(3) It is immediate that $K^{\perp}$ is a vector subspace in X by **H1**, so the name orthogonal *space* was not casual.
(4) In a Hermitian inner product space (X, S), by definition of orthogonality follows a useful property we will use often:

$$K \subset K_1 \quad \Rightarrow \quad K_1^{\perp} \subset K^{\perp} \quad \text{for } K, K_1 \subset \mathsf{X}.$$

(5) From now, lest we misunderstand, "(semi-)inner product" will always stand for "Hermitian (semi-)inner product". ■

Proposition 3.3. *Let* X *be a* $\mathbb{C}$*-vector space with semi-inner product S.*
(a) *The* **Cauchy-Schwarz inequality** *holds:*

$$|S(x,y)|^2 \leq S(x,x)S(y,y), \quad x,y \in \mathsf{X}\,; \tag{3.1}$$

(i) *there is equality in (3.1) if x and y are linearly dependent;*
(ii) *if S is an inner product, there is equality in (3.1) if and only if x, y are linearly dependent.*

(b) *As* $x \in \mathsf{X}$ *varies,*

$$p(x) := \sqrt{S(x,x)} \tag{3.2}$$

defines the seminorm (a norm if S is an inner product) **induced by** S.
(c) *p satisfies the* **parallelogram rule***:*

$$p(x+y)^2 + p(x-y)^2 = 2(p(x)^2 + p(y)^2), \quad x,y \in \mathsf{X}\,. \tag{3.3}$$

(d) *the* **polarisation formula** *holds:*

$$S(x,y) = \frac{1}{4}\left(p(x+y)^2 - p(x-y)^2 - ip(x+iy)^2 + ip(x-iy)^2\right), \quad x,y \in \mathsf{X}. \quad (3.4)$$

Proof. (a) If $\alpha \in \mathbb{C}$, using the properties of the semi-inner product,

$$0 \leq S(x-\alpha y, x-\alpha y) = S(x,x) - \alpha S(x,y) - \overline{\alpha} S(y,x) + |\alpha|^2 S(y,y). \quad (3.5)$$

Suppose $S(y,y) \neq 0$. Then setting $\alpha := \overline{S(x,y)}/S(y,y)$, (3.5) implies:

$$0 \leq S(x,x) - |S(x,y)|^2/S(y,y),$$

as claimed. If $S(y,y) = 0$, from (3.5) we find, for any $\alpha \in \mathbb{C}$:

$$0 \leq S(x,x) - \alpha S(x,y) - \overline{\alpha} S(y,x). \quad (3.6)$$

Choosing $\alpha \in \mathbb{R}$ large enough in absolute value we see inequality (3.6) is not satisfied unless $S(x,y) + S(y,x) = 0$. Choosing now $\alpha = i\lambda$ with $\lambda \in \mathbb{R}$ large enough in absolute value, we find that (3.5) can hold only if $S(x,y) - S(y,x) = 0$; with the previous $S(x,y) = -S(y,x)$ it gives $S(x,y) = 0$. Summing up, $S(y,y) = 0$ implies (3.1) because $S(x,y) = 0$. If x, y are linearly dependent then $x = \alpha y$ or $y = \alpha x$ for some $\alpha \in \mathbb{C}$. If so, the two sides of (3.1) are equal. Now assume S is an inner product and $|S(x,y)|^2 = S(x,x)S(y,y)$, and let us prove there are $\alpha, \beta \in \mathbb{C}$, not both zero, so that $\alpha x + \beta y = 0$. If one at least of x,y is null, the claim is true. So suppose neither vanishes, so $S(x,x) > 0 < S(y,y)$ by **H3**. Then redefining $u = x/\sqrt{S(x,x)}$, $v = y/\sqrt{S(y,y)}$, we have $|S(u,v)| = 1$ and so $S(u,v) = e^{i\eta}$ for some $\eta \in \mathbb{R}$. By H3, $\alpha' u + \beta' v = 0$ is equivalent to $S(\alpha' u + \beta' v, \alpha' u + \beta' v) = 0$, i.e.

$$|\alpha'|^2 + |\beta'|^2 + \overline{\alpha'}\beta' e^{i\eta} + \overline{\beta'}\alpha' e^{-i\eta} = 0,$$

as $S(u,v) = e^{i\eta}$. Choose $\alpha' = e^{i\mu}$, $\beta' = e^{i\nu}$, so that $-\mu + \nu + \eta = \pi$. Then the above identity holds, and setting $\alpha := e^{i\mu}\sqrt{S(y,y)}$, $\beta := e^{i\nu}\sqrt{S(x,x)}$ we have $\alpha, \beta \neq 0$ and $\alpha x + \beta y = 0$.

(b) The properties of seminorms are easy from the definition of p and the properties of the inner product, except the triangle inequality **N2** which we prove now. By the properties of the inner product

$$p(x+y)^2 = p(x)^2 + 2ReS(x,y) + p(y)^2,$$

with Re denoting the real part of a complex number. As $ReS(x,y) \leq |S(x,y)|$, by (3.1), we also have $ReS(x,y) \leq p(x)p(y)$. Substituting above gives:

$$p(x+y)^2 \leq p(x)^2 + 2p(x)p(x) + p(y)^2,$$

i.e.

$$p(x,y)^2 \leq (p(x) + p(y))^2,$$

which in turn implies **N2**. Property **N3** is immediate from **H3**, in case S is a scalar product.

Statements (c) and (d) are straighforward from the definition of p and the properties of inner product. □

Remark 3.4. **(1)** The Cauchy–Schwarz inequality immediately implies that an inner product $S : \mathsf{X} \times \mathsf{X} \to \mathbb{C}$ is a *continuous* map on $\mathsf{X} \times \mathsf{X}$ in the product topology, when X has the *topology of the norm given by the inner product*, i.e. (3.2). In particular the inner product is continuous in its arguments separately.
(2) If the field of X is $\mathbb{R}$ instead of $\mathbb{C}$, we have analogous definitions to 3.1, by setting a **real inner product** $S : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ to fulfill **H0**, **H1**, **H3** and replacing **H2** with the *symmetry property*:

H2'. $S(u,v) = S(v,u)$ for any $u, v \in \mathsf{X}$.

A **real semi-inner product** is a real inner product without **H3**, so to speak.
(3) Proposition 3.3 is still true for real (semi-)inner products, with the proviso that the new polarisation formula reads:

$$S(x,y) = \frac{1}{4}\left(p(x+y)^2 - p(x-y)^2\right) , \tag{3.7}$$

over the field $\mathbb{R}$. ■

A known result – rarely proved explicitly – is the following, due to Fréchet, von Neumann and Jordan. The proof is carried out in Exercises 3.1-3.3.

Theorem 3.5. *Let* X *be a complex vector space and* $p : \mathsf{X} \to \mathbb{R}$ *a norm (or seminorm). Then p satisfies the parallelogram rule (3.3) if and only if there exists – in such case unique – an inner product (or semi-inner product) S inducing p through (3.2).*

Proof. If the norm (seminorm) is induced by a Hermitian inner product, the parallelogram rule (3.3) is valid by Proposition 3.3(c). The proof that (3.3) implies the existence of an inner product (semi-inner product) S inducing p via (3.2) can be found in Exercises 3.1–3.3. □

Let us pass to *isomorphisms* of inner product spaces.

Definition 3.6. *Let* $(\mathsf{X}, S_\mathsf{X})$, $(\mathsf{Y}, S_\mathsf{Y})$ *be inner product spaces. A linear map* $L : \mathsf{X} \to \mathsf{Y}$ *is called* **isometry** *if:*

$$S_\mathsf{Y}(L(x), L(y)) = S_\mathsf{X}(x,y) \quad \textit{for any } x, y \in \mathsf{X}.$$

If the isometry $L : \mathsf{X} \to \mathsf{Y}$ *is onto we call it* **isomorphism of inner product spaces**. *Given an isomorphism (L) of inner product spaces from* X *to* Y *the spaces are said* **isomorphic (under** L).

Remarks 3.7. Every isometry $L : \mathsf{X} \to \mathsf{Y}$ is clearly 1-1 by **H3**, but can be *not* onto, even when $\mathsf{X} = \mathsf{Y}$, if the dimension of X is not finite. Every isometry is moreover continuous in the norm topologies induced by inner products; if surjective (isomorphism) its inverse is an isometry (isomorphism). ■

Since an inner product space is also normed, we have two notions of isometry for a linear transformation $L : \mathsf{X} \to \mathsf{Y}$. The first refers to the preservation of inner products (defined before), the second was given in Definition 2.10, and corresponsds to the requirement: $||Lx||_{\mathsf{Y}} = ||x||_{\mathsf{X}}$ for any $x \in \mathsf{X}$, with reference to the norms induced by the inner products. The former type also satisfies the second definition. Using the polarisation formula (3.4) it can actually be proved that the two notions are equivalent.

Proposition 3.8. *A linear operator $L : \mathsf{X} \to \mathsf{Y}$ between inner product spaces is an isometry in the sense of Definition 3.6 if and only if:*

$$||Lx||_{\mathsf{Y}} = ||x||_{\mathsf{X}} \quad \textit{for any } x \in \mathsf{X},$$

where the norms are associated to the corresponding Hilbert spaces' inner products.

Notation 3.9. Unless we say otherwise, from now on $(\,|\,)$ will be an inner product and $||\ ||$ the induced norm, as in 3.3. ■

Now to the truly central notion of the entire book, that of a *Hilbert space*.

Definition 3.10. *A complex vector space equipped with a Hermitian inner product is called a* **Hilbert space** *if the norm induced by the inner product makes it a Banach space. An isomorphism of inner product spaces between Hilbert spaces is said:*

(i) **isomorphism of Hilbert spaces**, *or*
(ii) **unitary transformation**, *or*
(iii) **unitary operator**.

It must be clear that under an isomorphism of inner product spaces $U : \mathsf{H} \to \mathsf{H}_1$, H_1 is a Hilbert space if and only if H is. Then U is a unitary transformation.

There is a result about completions similar to the one seen for Banach spaces.

Theorem 3.11 (Completion of Hilbert spaces). *Let X be a $\mathbb{C}$-vector space with inner product S.*

(a) *There exists a Hilbert space $(\mathsf{H}, (\ |\ ))$, called* **completion** *of X, such that X is identified to a dense subspace (for the norm associated to the inner product $(\ |\ )$) of H under a 1-1 linear map $J : \mathsf{X} \to \mathsf{H}$ that extends the inner product S to $(\ |\ )$. Equivalently, there is a 1-1 linear $J : \mathsf{X} \to \mathsf{H}$ with*

$$\overline{J(\mathsf{X})} = \mathsf{H} \quad \textit{and} \quad (J(x)|J(y)) = S(x,y) \quad \textit{for any } x, y \in \mathsf{X}.$$

(b) *If the triple $(J_1, \mathsf{H}_1, (\ |\ )_1)$, with $J_1 : \mathsf{X} \to \mathsf{H}_1$ linear isometry and $(\mathsf{H}_1, (\ |\ )_1)$ Hilbert space, is such that J_1 identifies X with a dense subspace in H_1 by extending S to $(\ |\ )_1$, then there is a unique unitary transformation $\phi : \mathsf{H} \to \mathsf{H}_1$ such that $J_1 = \phi \circ J$.*

Sketch of proof. (a) It is convenient to use the completion theorem for Banach spaces and then construct the Banach completion of the normed space (X, N), where $N(x) := \sqrt{S(x,x)}$. Since S is continuous and X dense in the completion under the linear map J, S induces a semi-inner product $(\,|\,)$ on the Banach completion H. Actually, $(\,|\,)$ is an inner product on H because, still by continuity, it induces the same norm of the Banach space. Thus H is a Hilbert space and the map J identifying X to a dense subspace in H satisfies all the requirements. (b) The proof is essentially the same as in the Banach case. □

Examples 3.12. (1) $\mathbb{C}^n$ with inner product $(u|v) := \sum_{i=1}^n \overline{u_i} v_i$, where $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$, is a Hilbert space.

(2) A crucial Hilbert space in physics arises from Example 2.26(6): we are talking about the space $L^2(\mathsf{X}, \mu)$, where X is a measure space with positive, σ-additive measure μ on a σ-algebra Σ of subsets in X. We know $L^2(\mathsf{X}, \mu)$ is a Banach space with norm $|| \, ||_2$:

$$||[f]||_2^2 := \int_{\mathsf{X}} \overline{f(x)} f(x) d\mu(x)$$

f being any representative in the equivalence class $[f] \in L^2(\mathsf{X}, \mu)$ (as usual, we shall write f instead of $[f]$).

If $f, g \in L^2(\mathsf{X}, \mu)$ then $\overline{f(x)} g(x) \in L^1(\mathsf{X}, \mu)$, for $(|f(x)| - |g(x)|)^2 \geq 0$ implies $2|f(x)||g(x)| \leq |f(x)|^2 + |g(x)|^2$. Hence:

$$(f|g) := \int_{\mathsf{X}} \overline{f(x)} g(x) d\mu(x), \quad f, g \in L^2(\mathsf{X}, \mu) \tag{3.8}$$

is well defined (which follows also from Hölder's inequality, cf. Example 2.26(6)). Elementary features of integrals guarantee the right-hand side of (3.8) is a Hermitian inner product on $L^2(\mathsf{X}, \mu)$, which clearly induces $|| \, ||_2$. Therefore $L^2(\mathsf{X}, \mu)$ is a Hilbert space with inner product (3.8).

(3) If one takes $\mathsf{X} = \mathbb{N}$ and the counting measure μ (Example 2.26(7)), as subcase of the previous situation we obtain the Hilbert space $\ell^2(\mathbb{N})$ of square-integrable complex sequences, where

$$(\{x_n\}_{n \in \mathbb{N}} | \{y_n\}_{n \in \mathbb{N}}) := \sum_{n=0}^{+\infty} \overline{x_n} y_n \,. \qquad \blacksquare$$

3.1.2 Riesz's theorem and its consequences

The aim is to prove that Hilbert spaces are *reflexive*. In order to do so we need to develop a few tools related to the notion of orthogonal spaces, and prove the celebrated *Riesz's theorem.*

Let us recall the definition of *convex set* (Definition 2.61).

Definition. *A set* $\varnothing \neq K$ *in a vector space* X *is* **convex** *if:*

$$\lambda u + (1 - \lambda) v \in K, \quad \textit{for any } \lambda \in [0, 1] \textit{ and } u, v \in K.$$

Clearly any subspace of X is convex, but not all convex subsets of X are subspaces in X: open balls (with finite radius) in normed spaces are convex as sets but not subspaces. For the next theorem we remind that $< K >$ denotes the subspace in X generated by $K \subset \mathsf{X}$, and $\overline{K}$ is the closure of K.

Theorem 3.13. *Let* $(\mathsf{H}, (\;|\;))$ *be a Hilbert space and* $K \subset \mathsf{H}$ *a non-empty subset. Then*

(a) $K^\perp$ *is a closed subspace of* H.
(b) $K^\perp = < K >^\perp = \overline{< K >}^\perp = \overline{< K >^\perp}$.

(c) *If K is closed and convex, for any $x \in \mathsf{H}$ there is a unique $P_K(x) \in K$ such that $||x - P_K(x)|| = \min\{||x-y|| \mid y \in K\}$, where $|| \, ||$ is the norm induced by $(\, | \,)$.*
(d) *If K is a closed subspace, any $x \in \mathsf{H}$ decomposes in a unique fashion as $z_x + y_x$ with $z_x \in \mathsf{K}$ and $y_x \in K^\perp$, so that:*

$$\mathsf{H} = K \oplus K^\perp \,. \tag{3.9}$$

Moreover, $z_x := P_K(x)$ as from (c).
(e) $(K^\perp)^\perp = \overline{<K>}$.

Remarks 3.14. Actually, (a) and (b) hold also on more general spaces than Hilbert spaces; it is enough to have an inner product space with inner product topology. ■

Proof of Theorem 3.13. (a) $K^\perp$ is a subspace by the (anti)linearity of the inner product. By its continuity it follows that if $\{x_n\} \subset K^\perp$ converges to $x \in \mathsf{H}$ then $(x|y) = 0$ for any $y \in K$, hence $x \in K^\perp$. So $K^\perp$, containing all its limit points, is closed.
(b) The first identity is trivial by definition of orthogonality and the linearity (antilinearity) of the inner product. The second relation follows immediately from (a). As for the third one, since $<K> \subset \overline{<K>}$ we have $<K>^\perp \supset \overline{<K>}^\perp$. But $<K>^\perp \subset \overline{<K>}^\perp$ by continuity, so $<K>^\perp = \overline{<K>}^\perp$, ending the chain of equalities, because we know $<K>^\perp = \overline{<K>^\perp}$.
(c) Let $0 \leq d = \inf_{y \in K} ||x-y||$ (this exists since the set of $||x-y||$ with $y \in K$ is lower bounded and non-empty). Define a sequence $\{y_n\} \subset K$ such that $||x - y_n|| \to d$. We will show it is a Cauchy sequence. From the parallelogram rule (3.3), where x, y are replaced by $x - y_n$ and $x - y_m$, we have

$$||y_n - y_m||^2 = 2||x-y_n||^2 + 2||x-y_m||^2 - ||2x - y_n - y_n||^2 \,.$$

Now $||2x - y_n - y_n||^2 = 4||x - (y_n + y_m)/2||^2 \geq 4d^2$, since $y_n/2 + y_m/2 \in K$ under the convexity assumption on K and because d is the infimum of the numbers $||x-y||$ when $y \in K$. Given $\varepsilon > 0$, and taking n, m big enough, we have: $||x-y_n||^2 \leq d^2 + \varepsilon$, $||x - y_m||^2 \leq d^2 + \varepsilon$, whence

$$||y_n - y_m||^2 \leq 4(d^2 + \varepsilon) - 4d^2 = 4\varepsilon \,.$$

So the sequence is Cauchy. As H is complete, y_n converges to some $y \in K$ because K is closed. The norm is continuous, so $d = ||x-y||$. We claim $y \in K$ is the unique point satisfying $d = ||x-y||$. For any other $y' \in K$ with the same property:

$$||y-y'||^2 = 2||x-y||^2 + 2||x-y'||^2 - ||2x-y-y'||^2 \leq 2d^2 + 2d^2 - 4d^2 = 0 \,,$$

by the parallelogram rule; we have used, above, the fact that $||2x - y - y'||^2 = 4||x - (y+y')/2||^2 \geq 4d^2$ (K is convex, d is the inf of the $||x - z||$ when $z \in K$, so $y/2 + y'/2 \in K$). As $||y - y'|| = 0$ we have $y = y'$. Thus $P_K(x) := y$ fulfills all requirements.
(d) Take $x \in \mathsf{H}$ and $x_1 \in K$ with minimum distance from x. Set $x_2 := x - x_1$, and we will show $x_2 \in K^\perp$. Pick $y \in K$, so the map $\mathbb{R} \ni t \mapsto f(t) := ||x - x_1 + ty||^2$ has a

minimum at $t = 0$. Notice this is true if K is a subspace, so that $-x_1 + ty \in K$ for any $t \in \mathbb{R}$ if $x_1, y \in K$. Hence its derivative vanishes at $t = 0$:

$$f'(0) = \lim_{t\to 0} \frac{||x_2 + ty||^2 - ||x_2||^2}{t} = (x_2|y) + (y|x_2) = 2Re(x_2|y)\,.$$

Therefore $Re(x_2|y) = 0$. Replacing y by iy tells that the imaginary part of $(x_2|y)$ is zero too, so $(x_2|y) = 0$ and $x_2 \in K^\perp$. We have proved $< K, K^\perp >= \mathsf{H}$. There remains $K \cap K^\perp = \{0\}$. But this is obvious because if $x \in K \cap K^\perp$, x must be orthogonal to x, so $||x||^2 = (x|x) = 0$ and $x = 0$.
(e) If $y \in K$, y is orthogonal to every element of $K^\perp$; by linearity and continuity of the inner product this is true also when $y \in \overline{< K >}$. In other words,

$$\overline{< K >} \subset (K^\perp)^\perp\,. \tag{3.10}$$

Using (d) and replacing K with the closed subspace $\overline{< K >}$ we obtain $\mathsf{H} = \overline{< K >} \oplus \overline{< K >}^\perp$. By (b) that is equivalent to

$$\mathsf{H} = \overline{< K >} \oplus K^\perp\,. \tag{3.11}$$

If $u \in (K^\perp)^\perp$, by (3.11) there is a decompostion into orthogonal ($(u_0|v) = 0$) vectors $u = u_0 + v$, with $u_0 \in \overline{< K >}$ and $v \in K^\perp$; thus $(u|v) = (v|v)$. But $(u|v) = 0$ ($u \in (K^\perp)^\perp$ and $v \in K^\perp$) so $(v|v) = 0$ and therefore $(K^\perp)^\perp \ni u = u_0 \in \overline{< K >}$. We conclude $\overline{< K >} \supset (K^\perp)^\perp$, and hence the claim, by (3.10). □

From (b) and (d) descends an immediate corollary.

Corollary 3.15 (to Theorem 3.13). *If S is a subset in a Hilbert space H, $< S >$ is dense in H if and only if $S^\perp = \{0\}$.*

We are ready to state and prove a theorem due to F. Riesz, by far the most important theorem in the theory of Hilbert spaces.

Theorem 3.16 (Riesz). *Let $(\mathsf{H}, (\ |\ ))$ be a Hilbert space. For any continuous linear functional $f : \mathsf{H} \to \mathbb{C}$ there exists a unique $y_f \in \mathsf{H}$ such that:*

$$f(x) = (y_f|x) \quad \text{for any } x \in \mathsf{H}\,.$$

The map $\mathsf{H}' \ni f \mapsto y_f \in \mathsf{H}$ is a bijection.

Proof. We will prove that for any $f \in \mathsf{H}'$ there is such $y_f \in \mathsf{H}$. The *null space* of f, $Kerf := \{x \in \mathsf{H} \mid f(x) = 0\}$, is a closed subspace since f is continuous. As H is a Hilbert space, $\mathsf{H} = Kerf \oplus (Kerf)^\perp$ by Theorem 3.13. If $Kerf = \mathsf{H}$ define $y_f = 0$ and the proof ends. If $Kerf \neq \mathsf{H}$ we can show $(Kerf)^\perp$ has dimension 1. Let $0 \neq y \in (Kerf)^\perp$. Then $f(y) \neq 0$ ($y \notin Kerf$!). For any $z \in (Kerf)^\perp$, the vector $z - \frac{f(z)}{f(y)}y$ belongs to $(Kerf)^\perp$, as linear combination of elements in $(Kerf)^\perp$, but $z - \frac{f(z)}{f(y)}y \in Kerf$ as well, by the linearity of f. Thus $z - \frac{f(z)}{f(y)}y \in Kerf \cap (Kerf)^\perp$, and $z - \frac{f(z)}{f(y)}y = 0$. This means

y is a basis for $(Ker f)^\perp$, any other vector $z \in (Ker f)^\perp$ being a linear combination of y: $z = \frac{f(z)}{f(y)} y$. If y is as above, define:

$$y_f := \frac{\overline{f(y)}}{(y|y)} y$$

and we will show y_f represents f in the sense we want. If $x \in \mathsf{H}$, we decompose x along $Ker f \oplus (Ker f)^\perp$, to get $x = n + x^\perp$, where

$$x^\perp = \frac{f(x^\perp)}{f(y)} y = \frac{f(x)}{f(y)} y,$$

because $f(x^\perp) = f(x)$ (by linearity, since $f(n) = 0$). So

$$(y_f|x) = \left(\frac{\overline{f(y)}}{(y|y)} y \middle| n + \frac{f(x)}{f(y)} y \right) = 0 + \frac{f(y)}{f(y)} \frac{(y|y)}{(y|y)} f(x) = f(x)\,.$$

The function $\mathsf{H}' \ni f \mapsto y_f \in \mathsf{H}$ is well defined, i.e. f determines y_f uniquely: if $(y|x) = (y'|x)$ for any $x \in K$ then $(y - y'|x) = 0$ for any $x \in K$; choosing $x = y - y'$ gives $||y - y'||^2 = (y - y'|y - y') = 0$, so $y = y'$. Injectivity is an easy consequence of having $f(x) = (y_f|x)$. The map $\mathsf{H}' \ni f \mapsto y_f \in \mathsf{H}$ is onto because, for any $y \in \mathsf{H}$, $\mathsf{H} \ni x \mapsto (y|x)$ is a point in H' by linearity and continuity of the inner product. □

Corollary 3.17 (to Riesz's theorem). *Every Hilbert space is reflexive.*

Proof. First of all we can endow H' with an inner product $(f|g)' := (y_g|y_f)$, where $f, g \in \mathsf{H}'$ with $f(x) = (y_f|x)$ and $g(x) = (y_g|x)$, $x \in \mathsf{H}$. The norm induced by $(\,|\,)'$ on H' coincides with the norm of H'

$$||f|| := \sup_{||x||=1} |f(x)|\,,$$

for which H' is complete (Theorem 2.41). By Theorem 3.16 we may write, in fact,

$$||f|| = \sup_{||x||=1} |(y_f|x)|\,;$$

and the Cauchy-Schwarz inequality implies $||f|| \leq ||y_f||$, and we also have $|(y_f|x)| = ||y_f||$ for $x = y_f/||y_f||$. Hence $||f|| = ||y_f||$, which is precisely the norm induced by $(\,|\,)'$. Therefore $(\mathsf{H}', (\,|\,)')$ is a Hilbert space and $(\mathsf{H}')'$ its dual. Theorem 3.16 guarantees that for any element in $(\mathsf{H}')'$, say F, there is $g_F \in \mathsf{H}'$ such that $F(f) = (g_F|f)'$ for any $f \in \mathsf{H}'$. But $(g_F|f)' = (y_f|y_{g_F}) = f(y_{g_F})$. We have obtained, for any $F \in (\mathsf{H}')'$, the existence (and uniqueness, by Corollary 2.55 to Hahn–Banach) of a vector $y_{g_F} \in \mathsf{H}$ such that:

$$F(f) = f(y_{g_F})$$

for any $f \in \mathsf{H}'$. This is the reflexivity of H. □

Remarks 3.18. From this proof we see that the topological dual H', equipped with inner product $(\,|\,)'$, $(f|g)' := (y_g|y_f)$, is a Hilbert space. The map $\mathsf{H}' \ni f \mapsto y_f \in \mathsf{H}$ is antilinear, 1-1, onto and preserves the inner product by construction. In this sense H and H' are *anti-isomorphic*. ■

3.2 Hilbert bases

Now we can introduce the mathematical arsenal relative to the notion of a *Hilbert basis*. This is a well-known generalisation, in infinite dimensions, of an orthonormal basis. We shall work in the most general setting, where Hilbert spaces are *not necessarily separable* and a basis can have any cardinality, also uncountable.

First we have to explain the meaning of infinite sums of non-negative numbers, often over uncountable sets. An **indexed set** $\{\alpha_i\}_{i\in I}$ will be a function $I \ni i \mapsto \alpha_i$. The set I is the **set of indices** and the pair (i,α_i), improperly written α_i, is the i**th** element of the indexed set. Note that it may happen $\alpha_i = \alpha_j$ for $i \neq j$.

Definition 3.19. *If $A = \{\alpha_i\}_{i\in I}$ is a non-empty set of non-negative reals indexed by I, of arbitrary cardinality, the* **sum of the indexed set** *A is the number, in $[0,+\infty)\cup\{+\infty\}$, defined by :*

$$\sum_{i\in I}\alpha_i := \sup\left\{ \sum_{j\in F}\alpha_j \;\middle|\; F \subset I\,,\; F \text{ finite} \right\}. \tag{3.12}$$

Remarks 3.20. From now a set will be called **countable** when it can be mapped bijectively to the natural numbers $\mathbb{N}$. Thus, here, a *finite* set is *not* countable. ■

Proposition 3.21. *In relation to Definition 3.19 we have:*

(a) *If I is finite or countable, the sum of the indexed set A coincides with the sum $\sum_{i\in I}\alpha_i$, or the sum of the series $\sum_{n=0}^{+\infty}\alpha_{i_n}$ respectively (which always converges, perhaps to $+\infty$, because its terms are non-negative), irrespective of the ordering, i.e. independently of the bijection $\mathbb{N}\ni n\mapsto i_n \in I$.*
(b) *If the sum of the set A is finite, the subset of I for which $\alpha_i \neq 0$ is finite or countable. In this case, by restricting to I the sum of A coincides with the sum over finitely many indices, or the sum of a series as in (a), respectively.*
(c) *If μ is the counting measure on I, defined by the σ-algebra of the power set of I ($J \subset I$ implies that $\mu(J) \leq +\infty$ is the cardinality of J by definition):*

$$\sum_{i\in I}\alpha_i = \int_A \alpha_i\, d\mu(i)\,. \tag{3.13}$$

Proof. (a) The case when I is finite is obvious, so we look at I countable and suppose to have chosen a particular ordering on I, so that we can write A as $A = \{\alpha_{i_n}\}_{n\in\mathbb{N}}$. We will show that the sum $\sum_{n=0}^{+\infty}\alpha_{i_n}$ of $\{\alpha_{i_n}\}_{n\in\mathbb{N}}$ coincides with the sum of (3.12) which does not depend on the chosen ordering by definition. Because of (3.12) we have:

$$\sum_{n=0}^{N}\alpha_{i_n} \leq \sum_{i\in I}\alpha_i\,.$$

The limit for $N\to+\infty$ exists and equals the supremum of the set of partial sums, since the latter are non-decreasing. Therefore:

$$\sum_{n=0}^{+\infty}\alpha_{i_n} \leq \sum_{i\in I}\alpha_i\,. \tag{3.14}$$

On the other hand, if $F \subset I$ is finite, we may fix N_F big enough so that $\{\alpha_i\}_{i\in F} \subset \{\alpha_{i_0}, \alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{i_{N_F}}\}$. Thus

$$\sum_{i\in F} \alpha_i \leq \sum_{n=0}^{N_F} \alpha_{i_n}.$$

Taking now the supremum over finite sets $F \subset I$, and remembering the supremum of partial sums is the sum of the series, gives

$$\sum_{i\in I} \alpha_i \leq \sum_{n=0}^{+\infty} \alpha_{i_n}. \tag{3.15}$$

Then (3.14) and (3.15) produce the claim.
(b) Suppose $S < +\infty$, $S \geq 0$, is the sum of the set A. If $S = 0$ all elements of A are zero and the proof ends. So suppose $S > 0$. Any α_i is contained in $[0,S]$ – otherwise the sum would be larger than S – and in particular $\alpha_i \neq 0$ implies $\alpha_i \in (0,S]$. Define $S_n := S/n$, $n = 1,2,\ldots$. If N_k denotes the number of $i \in I$ for which α_i belongs in $(S_{k+1}, S_k]$, then surely $S \geq S_{k+1}N_k$, hence N_k is finite for any k. But $\cup_{k=1}^{+\infty}(S_{k+1}, S_k] = (0,S]$, so every $\alpha_i \neq 0$ are accounted for as $k = 1,2,\ldots$. These values can be at most countable, since: (i) there are countably many intervals $(S_{k+1}, S_k]$ and (ii) each one contains a finite number of $\alpha_i \neq 0$.
(c) Since any function is measurable with respect to the given measure (the σ-algebra is the power set), identity (3.13) is an immediate consequence of the definition of integral of a positive function (cf. Chapter 1.4.3). □

Now we can define, step by step, *complete orthonormal systems*, also known as *Hilbert bases*.

Definition 3.22. *Let* $(\mathsf{X}, (\,|\,))$ *be an inner product space and* $\varnothing \neq N \subset \mathsf{X}$.

(a) *N is an* **orthogonal system** *(of vectors) if (i) $N \not\ni 0$ and (ii) $x \perp y$ for any $x, y \in N$, $x \neq y$.*
(b) *N is an* **orthonormal system** *if all its vectors are orthogonal and unit, $(x|x) = 1$ for any $x \in N$.*
If $(\mathsf{H}, (\,|\,))$ is a Hilbert space, $N \subset \mathsf{H}$ is a **complete orthonormal system***, or a* **Hilbert basis***, if orthonormal and:*

$$N^{\perp} = \{0\}. \tag{3.16}$$

When no confusion arises, a Hilbert basis will be simply referred to as a **basis***.*

Remarks 3.23. Any orthogonal system N is made of *linearly independent* vectors: if $F \subset N$ is finite and $0 = \sum_{x\in F} \alpha_x x$, then $0 = (\sum_{x\in F} \alpha_x x | \sum_{y\in F} \alpha_y y) = \sum_{x\in F}\sum_{y\in F} \overline{\alpha_x}\alpha_y (x|y) = \sum_{x\in F} |\alpha_x|^2 ||x||^2$. As $||x|| > 0$ and $|\alpha_x|^2 \geq 0$, necessarily $|\alpha_x| = 0$, so $\alpha_x = 0$, for any $x \in F$. ■

Theorem 3.24 (Bessel's inequality). *For any orthonormal system $N \subset \mathsf{X}$ in an inner product space $(\mathsf{X}, (\ |\ ))$,*

$$\sum_{z\in N} |(x|z)|^2 \leq ||x||^2 \quad \textit{for any } x \in \mathsf{X}\,. \tag{3.17}$$

In particular: only a countable number of products $(x|z)$, at most, are non-zero.

Proof. By Definition 3.19 and Proposition 3.21(b) the claim holds if inequality (3.17) is true for all finite $F \subset N$. So let $F = \{z_1, \ldots, z_n\}$, $x \in \mathsf{X}$ and take $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. Expanding $||x - \sum_{k=1}^n \alpha_k z_k||^2$ in terms of the inner product of X and because of the orthonormality of z_p and z_q, plus the inner product's (anti)linearity, we obtain:

$$\left|\left|x - \sum_{k=1}^n \alpha_k z_k\right|\right|^2 = ||x||^2 + \sum_{k=1}^n |\alpha_k|^2 - \sum_{k=1}^n \alpha_k (x|z_k) - \sum_{k=1}^n \overline{\alpha_k (x|z_k)}\,.$$

The right-hand side is:

$$||x||^2 - \sum_{k=1}^n |(x|z_k)|^2 + \sum_{k=1}^n \left(|(x|z_k)|^2 - \alpha_k(x|z_k) - \overline{\alpha_k(x|z_k)} + |\alpha_k|^2\right)\,.$$

So,

$$\left|\left|x - \sum_{k=1}^n \alpha_k z_k\right|\right|^2 = ||x||^2 - \sum_{k=1}^n |(x|z_k)|^2 + \sum_{k=1}^n |\,(z_k|x) - \alpha_k\,|^2\,.$$

On the right there is only one absolute minimum point $\alpha_k = (z_k|x)$, $k = 1, \ldots, n$. Therefore

$$0 \leq \left|\left|x - \sum_{k=1}^n (z_k|x) z_k\right|\right|^2 = ||x||^2 - \sum_{k=1}^n |(x|z_k)|^2\,,$$

and finally:

$$\sum_{k=1}^n |(x|z_k)|^2 \leq ||x||^2.$$

□

Lemma 3.25. *Let $\{x_n\}_{n\in\mathbb{N}}$ be a countable orthogonal system indexed by $\mathbb{N}$ in the Hilbert space $(\mathsf{H}, (\ |\ ))$, and let $||\ ||$ be the norm induced by $(\ |\ )$. If*

$$\sum_{n=0}^{+\infty} ||x_n||^2 < +\infty\,, \tag{3.18}$$

then:

(a) *there exists, unique, $x \in \mathsf{H}$ such that*

$$\sum_{n=0}^{+\infty} x_n = x\,, \tag{3.19}$$

where convergence is understood as convergence of partial sums in the topology induced by $|| \, ||$;
(b) *the series (3.19) can be re-ordered, i.e.*

$$\sum_{n=0}^{+\infty} x_{f(n)} = x \tag{3.20}$$

for any bijection $f : \mathbb{N} \to \mathbb{N}$.

Proof. (a) Take $A_n := \sum_{k=0}^{n} x_k$; by the orthonormality of the x_k and the definition of norm via the scalar product we have, for $n > m$:

$$||A_n - A_m||^2 = \sum_{k=m+1}^{n} ||x_k||^2 .$$

Since the series converges,

$$||A_n - A_m||^2 = \sum_{k=m+1}^{n} ||x_k||^2 \le \sum_{k=m+1}^{+\infty} ||x_k||^2 \to 0 \quad \text{as } m \to +\infty,$$

which in turn implies that partial sums $\{A_n\}$ are a Cauchy sequence. Since H is complete, there is a limit point $x \in \mathsf{H}$ of the sequence, so of the series. But H is normed, and the Hausdorff property tells that limits, x included, are unique.
(b) Fix a bijective $f : \mathbb{N} \to \mathbb{N}$. Set, as above, $A_n := \sum_{k=0}^{n} x_k$ and $\sigma_n := \sum_{k=0}^{n} x_{f(k)}$. The positive-term series $\sum_{k=0}^{+\infty} ||x_{f(k)}||^2$ converges because its partial sums are smaller than the converging series $\sum_{k=0}^{+\infty} ||x_k||^2$. From part (a) the limit in H of σ_n exists, and the re-ordered series will converge too in H. We claim this limit is exactly x.
Define $r_n := \max\{f(0), f(1), \ldots, f(n)\}$, so

$$||A_{r_n} - \sigma_n||^2 \le \sum_{k \in J_n} ||x_k||^2$$

where J_n arises from

$$\{0, 1, 2, \ldots, \max\{f(0), f(1), \ldots, f(n)\}\}$$

by erasing $f(0), f(1), \ldots, f(n)$. By bijectivity the remaining elements correspond to certain points of the infinite set

$$\{f(n+1), f(n+2), \ldots\} .$$

Therefore

$$||A_{r_n} - \sigma_n||^2 \le \sum_{k \in J_n} ||x_k||^2 \le \sum_{k=n+1}^{+\infty} ||x_{f(k)}||^2 . \tag{3.21}$$

As $\sum_{k=0}^{+\infty} ||x_{f(k)}||^2 < +\infty$, relation (3.21) implies:

$$\lim_{n \to +\infty} (A_{r_n} - \sigma_n) = 0 .$$

On the other hand $r_n \geq n$ (f is injective, and if we had

$$\max\{f(0), f(1), \ldots, f(n)\} < n\,,$$

the various $f(n)$ should be $n+1$ non-negative integers smaller than n, a contradiction). Then $\lim_{n\to+\infty} r_n = +\infty$, so:

$$x = \lim_{n\to+\infty} A_n = \lim_{n\to+\infty} A_{r_n} = \lim_{n\to+\infty} \sigma_n\,. \qquad \square$$

We can now state, and prove, the fundamental theorem about Hilbert bases, according to which Hilbert bases generalise orthonormal bases in inner product spaces *of finite dimension*. The novelty is that, at present, also *infinite linear combinations* are allowed, using the topology of H: any element of a Hilbert space can be written in a unique fashion as infinite linear combination of a basis.

Irrespective of the existence of bases, by Zorn's lemma (or equivalently, the axiom of choice) there are also 'algebraic' bases that require no topology. The difference between a Hilbert basis and an algebraic one lies in that the latter concerns *finite combinations* only: despite the basis has infinite cardinality, any vector in the (Hilbert) space can be decomposed, uniquely, as a finite linear combination of the basis' elements.

Theorem 3.26. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* $N \subset \mathsf{H}$ *an orthonormal system. The following facts are equivalent:*

(a) *N is a basis (N is an orthonormal system and $N^\perp = \{0\}$).*
(b) *Given $x \in \mathsf{H}$, a countable (at most) number of $(z|x)$ is non-zero for all $z \in N$, and:*

$$x = \sum_{z\in N} (z|x)z\,, \tag{3.22}$$

where the series converges in the sense that partial sums converge in the inner product topology.
(c) *Given $x, y \in \mathsf{H}$, a countable (at most) number of $(z|x)$, $(y|z)$ is non-zero for all $z \in N$, and*

$$(x|y) = \sum_{z\in N} (x|z)(z|y)\,. \tag{3.23}$$

(d) *If $x \in \mathsf{H}$:*

$$||x||^2 = \sum_{z\in N} |(z|x)|^2\,. \tag{3.24}$$

in the sense of Definition 3.19.
(e) $\overline{<N>} = \mathsf{H}$, *i.e. the span of N is dense in* H.
Under any of the above properties, in (3.22) and (3.23) the indexing order of non-null coefficients of $(x|z)$, $(z|x) = \overline{(x|z)}$ and $(z|y)$ is irrelevant.

Proof. (a) $\Rightarrow$ (b). By Theorem 3.24 only countably many coefficients $(z|x)$ are non-null, at most. Indicate by $(z_n|x)$, $n \in \mathbb{N}$, these numbers and fix $S_N := \sum_{n=0}^{N} (z_n|x)z_n$. The system $\{(z_n|x)z_n\}_{n\in\mathbb{N}}$ is by construction orthogonal, and because $||z_n|| = 1$ Bessel's inequality implies $\sum_{n=0}^{+\infty}||(z_n|x)z_n||^2 < +\infty$. By Lemma 3.25(a) the series (3.22) converges to a unique $x' \in \mathsf{H}$, $x' = \sum_{n=0}^{+\infty}(z_n|x)z_n$. Moreover, the series can be rearranged, with the same limit x' by Lemma 3.25(b). We claim $x' = x$. The linearity and continuity of the inner product force, for $z' \in N$:

$$(x - x'|z') = (x|z') - \sum_{z\in N}(x|z)(z|z') = (x|z') - (x|z') = 0$$

where we have used the fact that the set of coefficients z is an orthonormal system. Since $z' \in N$ is arbitrary, $x - x' \in N^\perp$ and so $x - x' = 0$, as $N^\perp = \{0\}$ by assumption. This proves that (3.22) holds independently from the thay we index the coefficients $(z|x) \neq 0$.
(b) $\Rightarrow$ (c). If (b) holds, (c) is an obvious consequence, due to continuity and (anti)linearity of the inner product, plus the fact N is orthonormal.
(c) $\Rightarrow$ (d). (d) is a special case of (c) when $y = x$.
(d) $\Rightarrow$ (a). If (d) is true and $x \in \mathsf{H}$ is such that $(x|z) = 0$ for any $z \in N$, then $||x|| = 0$, i.e. $x = 0$. In other words $N^\perp = \{0\}$, that is to say (a).

So, we proved (a), (b), (c) and (d) are equivalent. To finish notice (b) implies immediately (e), while (e) implies (a): if $x \in N^\perp$, the inner product's linearity gives $x \in <N>^\perp \subset \overline{<N>}^\perp$. But Theorem 3.13(b) says $\overline{<N>^\perp} = \overline{<N>}^\perp$. Since $\overline{<N>} = \mathsf{H}$ by hypothesis, $x \in \mathsf{H}^\perp = \{0\}$. Put otherwise we have (a), as $N^\perp = \{0\}$.

The fact that the complex series in (3.23) can be rearranged with the same sum relies on the following argument. Consider the set

$$A := \{z \mid (x|z) \neq 0 \text{ or } (y|z) \neq 0\},$$

which is countable. The Cauchy-Schwarz inequality in $\ell^2(A)$ produces:

$$\sum_{z\in A}|(x|z)||(z|y)| \leq (\sum_{z\in A}|(x|z)|^2)^{1/2}(\sum_{z\in A}|(y|z)|^2)^{1/2} < +\infty$$

by (d). Hence the series $\sum_{z\in N}(x|z)(z|y) = \sum_{z\in A}(x|z)(z|y)$ can be rearranged as one likes because it converges absolutely. □

Zorn's lemma now guarantees each Hilbert space admits a complete orthonormal system.

Theorem 3.27. *Every Hilbert space admits a basis.*

Proof. Let H be a Hilbert space and consider the collection $\mathscr{A}$ of orthonormal systems in H. Define on $\mathscr{A}$ the partial order relation given by the inclusion of sets. By construction any ordered subset $\mathscr{E}$ in $\mathscr{A}$ is upper bounded by the union of all elements of $\mathscr{E}$. Zorn's lemma tells us there is in $\mathscr{A}$ a maximal element N. Therefore there are in H no vectors that are normal to every element in N, non-zero and not belonging to N itself. This means N is a complete orthonormal system. □

Before moving on to separable Hilbert spaces, let us give another important result from the general theory.

Theorem 3.28. *Let* H *be a Hilbert space with basis* N. *Then:*

(a) H *is isomorphic, as Hilbert space, to* $L^2(N,\mu)$, *where* μ *is the positive counting measure of* N *(see Examples 2.26(6, 7) and 3.12(2)); the unitary transformation that identifies the two spaces is*

$$\mathsf{H} \ni x \mapsto \{(z|x)\}_{z\in N} \in L^2(N,\mu)\,; \tag{3.25}$$

(b) *all bases of* H *have the same cardinality (that of* N*), called the* **dimension** *of the Hilbert space.*
(c) *If* H_1 *is a Hilbert space with the same dimension of* H, *the two spaces are isomorphic as Hilbert spaces.*

Proof. (a) The map $U : \mathsf{H} \ni x \mapsto \{(z|x)\}_{z\in N} \in L^2(N,\mu)$ is well defined because if $x \in \mathsf{H}$ and N is a basis, then property (d) of Theorem 3.26 holds, according to which $\{(z|x)\}_{z\in N} \in L^2(N,\mu)$. This function is definitely 1-1: if $x,x' \in \mathsf{H}$ give equal coefficients $(z|x) = (z|x')$ for any $z \in N$, then $x = x'$ by Theorem 3.26(b). The map is onto as well: if $\{\alpha_z\}_{z\in N} \in L^2(N,\mu)$, so $\sum_{z\in N} |\alpha_z|^2 < +\infty$, by Lemma 3.25 there is $x := \sum_{z\in N} \alpha_z z$ and $(z|x) = \alpha_z$ by inner product continuity and orthonormality of N. Now Theorem 3.26(c) implies U is isometric. Therefore $U : \mathsf{H} \to L^2(N,\mu)$ is a unitary operator, making H and $L^2(N,\mu)$ isomorphic Hilbert spaces.
(b) If one Hilbert basis has finite cardinality c, it must be an algebraic basis for H. Elementary geometric techniques allow to prove that if a basis of finite cardinality c exists, then any other set of linearly independent vectors M has cardinality $\leq c$, and the maximum is reached if and only if M spans the whole space. Since a basis, being an orthogonal system, is made of linearly independent vectors, we conclude that any basis of H has cardinality $\leq c$, hence $= c$ because it spans H finitely. This prevents the situation where one basis has finite cardinality and another one infinite. So let N and M be bases di H of infinite cardinality. If $x \in M$, define $N_x := \{z \in N \mid (x|z) \neq 0\}$. As $1 = (z|z) = \sum_{x\in M} |(z|x)|^2$, we must have, for any $z \in N$, an $x \in M$ such that $z \in N_x$. Therefore $N \subset \cup_{x\in M} N_x$ and then the cardinality of N will be less than or equal to that of $\cup_{x\in M} N_x$; but the latter is the cardinality of M because any N_x is countable at most by Theorem 3.26(b). So the cardinality of N does not exceed the cardinality of M. Swapping the roles of N and M we obtain that the cardinality of M is not larger than the one of N and the theorem of Schröder-Bernstein ensures the two cardinalities coincide.
(c) Let N and N_1 be bases of H and H_1 respectively, and suppose they have the same cardinality. Then there is a bijective map taking points in N to points in N_1 that induces a natural isomorphism V of inner product spaces between the L^2 space on N and the L^2 space on N_1 with respect to the counting measure. Therefore V is an isomorphism of Hilbert spaces. If $U_1 : \mathsf{H}_1 \to L^2(N_1,\mu)$ is the isomorphism analogous to the aforementioned $U : \mathsf{H} \to L^2(N,\mu)$, then $UVU_1^{-1} : \mathsf{H}_1 \to \mathsf{H}$ is a unitary transformation, by construction, making H and H_1 isomorphic spaces. □

So-called *separable* Hilbert spaces are particularly interesting in physics.

Definition 3.29. *A Hilbert space is* **separable** *if it admits a countable dense subset.*

There is a well-known characterisation of separability.

Theorem 3.30. *Let* H *be a Hilbert space.*

(a) H *is separable if and only if either it has finite dimension or it has a countable basis.*
(b) *If* H *is separable, every basis is either finite, with cardinality equal to the space's dimension, or countable.*
(c) *If* H *is separable then it is isomorphic either to* $\ell^2(\mathbb{N})$*, or to* $\mathbb{C}^n$ *equipped with standard Hermitian inner product, and the finite number* n *equals the dimension of* H*.*

Proof. (a) If the Hilbert space has a finite or countable basis, Theorem 3.26(b) ensures that a countable dense set exists, because rational numbers are dense in the reals. This set consists clearly of finite linear combinations of basis elements with complex coefficients having rational real and imaginary parts. The easy proof is left to the reader. Conversely, suppose a Hilbert space is separable. By Theorem 3.27, we know bases exist, and we want to show that any basis must be countable at most.

Suppose, by contradiction, that N is an uncountable basis for the separable Hilbert space H. For any chosen $z, z' \in N$, $z \neq z'$, any point $x \in \mathsf{H}$ is such that: $||z-z'|| \leq ||x-z'|| + ||z-x||$ by the triangle inequality induced by the inner product. At the same time $\{z, z'\}$ is an orthonormal system, so $||z-z'||^2 = (z-z'|z-z') = ||z||^2 + ||z'||^2 + 0 = 1+1 = 2$. Hence $||x-z|| + ||x-z'|| \geq \sqrt{2}$. This implies that a pair of open balls of radius $\varepsilon < \sqrt{2}/2$ centred at z and z' are disjoint, irrespective of how we pick $z, z' \in N$ with $z \neq z'$. Call $\{B(z)\}_{z\in N}$ a family of such balls parametrised by their centres $z \in N$. If $D \subset \mathsf{H}$ is a countable dense set (the space is separable), then for any $z \in N$ there exists $x \in D$ with $x \in B(z)$. The balls are pairwise disjoint, so there will be one x for each ball, all different from one another. But the cardinality of $\{B(z)\}_{z\in N}$ is not countable, hence neither D can be countable, a contradiction.

Although (b) and (c) are straighforward consequences of Theorem 3.28, for the sake of the argument let us outline their proof.
(b) From the basic theory, if a (Hilbert or algebraic) basis is finite, the cardinality of any other basis equals the dimension of the space; moreover, any linearly independent set (viewed as basis) cannot contain a number of vectors exceeding the dimension. From this, if a Hilbert space is separable and one of its bases is finite, then all bases are finite with the space's dimension as cardinality. Under the same hypotheses, if a basis is countable then any other is countable by (a).
(c) Fix a basis N. Using Theorem 3.26 one verifies quickly that the map sending $\mathsf{H} \ni x = \sum_{u\in N} \alpha_u u$ to the (infinite or finite, depending on the situation) $\{\alpha_u\}_{u\in N}$ is an isomorphism of inner product spaces from H to either $\mathbb{C}^n$ or $\ell^2(\mathbb{N})$, according to whether the dimension of H is finite or not. □

Here is another useful propostion about separable Hilbert spaces.

Proposition 3.31. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space with* $\mathsf{H} \neq \{0\}$.

(a) *If* $Y := \{y_n\}_{n\in\mathbb{N}}$ *is a set of linearly independent vectors and* $Y^\perp = \{0\}$*, or equivalently* $\overline{<Y>} = \mathsf{H}$*, then* H *is separable and there exists a basis* $X := \{x_n\}_{n\in\mathbb{N}}$ *in* H *such*

that, for any $p \in \mathbb{N}$, the span of $y_0, y_1, \ldots, y_p$ coincides with the span of $x_0, x_1, \ldots, x_p$.
(b) *If* H *is separable and* $\mathsf{S} \subset \mathsf{H}$ *is a (non-closed) dense subspace of* H*, then* S *contains a basis of* H.

Proof. (a) We shall only sketch the proof as the argument essentially duplicates the *Gram-Schmidt orthonormalisation process* known from basic geometry courses [Ser94I]. Since $y_0 \neq 0$, we set $x_0 := y_0/||y_0||$. Consider the non-null (y_0, y_1 are linearly independent) vector $z_1 := y_1 - (x_0|y_1)x_0$. Clearly x_0, z_1 are not null, they are orthogonal (so linearly independent) and span the same subspace as y_0, y_1. Setting $x_1 := z_1/||z_1||$ produces an orthonormal set $\{x_0, x_1\}$ spanning the same space as y_0, y_1. The procedure can be iterated inductively, by defining:

$$z_n := y_n - \sum_{k=0}^{n-1} (x_k|y_n)x_k \,,$$

and considering the set of $x_n := z_n/||z_n||$. By induction it is easy to see $z_0, \ldots, z_k$ are non-null, orthogonal (hence linearly independent) and span the same space generated by the linearly independent $y_0, \ldots, y_k$. If $u \perp y_n$ for any $n \in \mathbb{N}$, then $u \perp x_n$ for any $n \in \mathbb{N}$ (it is enough to express x_n as a linear combination of $y_0, \ldots, y_n$), and conversely (writing y_n as combination of $x_0, \ldots, x_n$). Therefore $X^\perp = Y^\perp = \{0\}$ and X is a basis for H.
(b) S must contain a subset S_0 that is countable and dense in H. In fact, let $\{y_n\}_{n\in\mathbb{N}}$ be countable and dense in H. For any y_n there is a sequence $\{x_{nm}\}_{m\in\mathbb{N}} \subset \mathsf{S}$ such that $x_{nm} \to y_n$ as $m \to +\infty$. It is straighforward that the countable subset $S_0 := \{x_{nm}\}_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ of S is dense in H. Relabelling the elements of S_0 over the naturals so that $x_1 \neq 0$ we have $S_0 = \{x_q\}_{q\in\mathbb{N}}$. Now we can decompose S_0 in two subsets S_1, containing x_1, and S_2 as follows. If x_2 is linearly independent from x_1 we put x_2 in S_1, otherwise in S_2. If x_3 is linearly independent from x_1, x_2 in S_1 we put it in S_1, otherwise in S_2, and we continue like this until we exhaust S_0. Then by construction S_1 contains a set of linearly independent generators of S_0. Thus $\overline{< S_1 >} \supset \overline{S_0} = \mathsf{H}$. Furthermore this process builds a complete orthonormal system by *finite* linear combiantions of $Y := S_1$, as explained in (a), and so it gives a basis made of elements of S since $\mathsf{S} \supset S_1$ is a subspace. □

Examples 3.32. (1) Consider the Hilbert space $L^2([-L/2, L/2], dx)$ (cf. Examples 3.12(2)) where dx is the usual Lebesgue measure on $\mathbb{R}$ and $L > 0$. Take measurable functions (they are continuous)

$$f_n(x) := \frac{e^{i\frac{2\pi n}{L}x}}{\sqrt{L}} \tag{3.26}$$

for $n \in \mathbb{Z}$ and $x \in [-L/2, L/2]$. It is immediate to see the maps f_n belong to the space and form an orthonormal system for the inner product of $L^2([-L/2, L/2], dx)$ (see Examples 3.12(2)):

$$(f|g) := \int_{-L/2}^{L/2} \overline{f(x)} g(x) dx \,. \tag{3.27}$$

Consider the Banach algebra of continuous maps $C([-L/2,L/2])$ (a vector subspace of $L^2([-L/2,L/2],dx)$) with supremum norm (Examples 2.26(4, 5)). The vector subspace S in $C([-L/2,L/2])$ spanned by all f_n, $n \in \mathbb{Z}$, is a subalgebra of $C([-L/2,L/2])$. S contains 1, it is closed under complex conjugation and it is not hard to see that it separates points in $[-L/2,L/2]$ (the set of f_n alone separates points), so the Stone–Weierstrass Theorem 2.27 gurantees S is dense in $C([-L/2,L/2])$. On the other hand it is well known that continous maps on $[-L/2,L/2]$ form a dense space in $L^2([-L/2,L/2],dx)$ in the latter's topology [Rud82, p. 85]; eventually, the topology of $C([-L/2,L/2])$ is finer than that of $L^2([-L/2,L/2],dx)$, because $(f|f) \leq L\sup|f|^2 = L(\sup|f|)^2$ if $f \in C([-L/2,L/2])$. Therefore S is dense in $L^2([-L/2,L/2],dx)$. By Theorem 3.26(e), the vectors f_n form a basis in $L^2([-L/2,L/2],dx)$, making the latter separable.

(2) Consider the Hilbert space $L^2([-1,1],dx)$, dx being the Lebegue measure. As in the previous example the Banach algebra $C([-1,1])$ is dense in $L^2([-1,1],dx)$ in the latter's topology. In contrast to what we had previously, let

$$g_n(x) := x^n\,, \tag{3.28}$$

for $n = 0,1,2,\ldots, x \in [-1,1]$. It can be proved that these vectors are linearly independent, and their span S in $C([-1,1])$ is a subalgebra in $C([-1,1])$. S contains the unit 1, it is closed under complex conjugation and separates points, so the Stone–Weierstrass Theorem 2.27 implies it is dense in $C([-1,1])$; by arguing as in the above example it is also dense in $L^2([-1,1],dx)$. The news is that now the functions g_n do not constitute an orthonormal system. However, using Proposition 3.31 we can immediately build a complete orthonormal system on $L^2([-1,1],dx)$. These basis elements, up to a normalisation, are called **Legendre polynomials**:

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n, \quad n = 0,1,2,\ldots.$$

From this definition follow the orthogonality relations:

$$\int_{[-1,1]} P_n(x)P_m(x)dx = \frac{2\delta_{nm}}{2n+1}\,.$$

(3) In the previous examples we exhibited two L^2 separable spaces. It can be proved that $L^p(\mathsf{X},\mu)$ $(1 \leq p < +\infty)$ is separable if and only if the measure μ is *separable* in the following sense. Take the *metric space* (cf. Definition 2.78) quotient of the subset of the σ-algebra Σ of μ made of all finite-measure sets built by modding out zero-measure sets. Define the distance:

$$d(A,B) := \mu((A\setminus B)\cup(B\setminus A))\,.$$

The measure μ is said to be separable if this metric space admits a dense and countable subset. Concerning separable measures we have the following result [Hal69].

Proposition 3.33 (On separable L^p measures and spaces). *A σ-additive positive measure μ, and hence also $L^p(\mathsf{X},\mu)$, is separable if the following conditions hold: (i) μ is σ-finite (X is the union of at most countably many sets of finite measure) and (ii) the σ-algebra of the measure space of μ is generated by a countable, at most, collection of measurable sets.*

As consequence we have:

Proposition 3.34 (On separable Borel measures and L^p spaces). *Every σ-finite Borel measure on a second-countable topological space produces a separable L^p space.*

Remarks 3.35. This is the case, in particular, of the L^p space relative to the Lebesgue measure on $\mathbb{R}^n$ restricted to Borel sets in $\mathbb{R}^n$. Note, though, that the L^p space obtained is the same we find by using the entire Lebesgue σ-algebra, since the latter is the completion of the Borel σ-algebra for the Lebesgue measure restricted to Borel subsets (see the remark following Proposition 1.65) by Proposition 1.57. ■

Positive and σ-additive Borel measures on locally compact Hausdorff spaces are called **Radon measures** if they are regular and if compact sets have finite measure. A Radon measure is σ-finite if the space on which it is defined is σ-compact, i.e. countable union (at most) of compact sets.

(4) Consider the space $L^2((a,b),dx)$, with $-\infty \le a < b \le +\infty$ and dx being the usual Lebesgue measure on $\mathbb{R}$. From the definitions of the Fourier and Fourier-Plancherel transforms (Proposition 3.90) we have an extremely useful result:

Let $f : (a,b) \to \mathbb{C}$ be measurable and such that: (1) the set $\{x \in (a,b) \mid f(x) = 0\}$ has zero measure, and (2) there exist $C, \delta > 0$ so that $|f(x)| < Ce^{-\delta|x|}$ for any $x \in (a,b)$. Then the finite span of the maps $x \mapsto x^n f(x)$, $n = 0,1,2,\ldots$, is dense in $L^2((a,b),dx)$.

The importance of this fact relies in that it allows to construct with ease bases in $L^2((a,b),dx)$ even when a or b are infinite (a case in which we cannot apply the Stone–Weierstrass Theorem 2.27). In fact, the Gram-Schmidt process applied to $f_n(x) := x^n f(x)$ yields a basis as explained in Proposition 3.31.

For instance, if $f(x) := e^{-x^2/2}$ Gram-Schmidt gives, normalisation apart, the basis of $L^2(\mathbb{R},dx)$ of so-called (normalised) **Hermite functions**:

$$\psi_0(x) = \pi^{-1/4} e^{-x^2/2}$$

and, recursively,

$$\psi_{n+1} = (2(n+1))^{-1/2}\left(x - \frac{d}{dx}\right)\psi_n \quad n = 0,1,2,\ldots$$

A computation, essentially relying on Gram-Schmidt, shows that ψ_n can be obtained alternatively as:

$$\psi_n(x) := (2^n n! \sqrt{\pi})^{-1/2} H_n(x) e^{-x^2/2} \quad n = 0,1,2,\ldots$$

where H_n is a polynomial of degree $n = 0, 1, 2, \ldots$ called nth **Hermite polynomial**:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2} \quad n = 0, 1, 2, \ldots$$

There are orthogonalty relations:

$$\int_{\mathbb{R}} e^{-x^2} H_n(x) H_m(x) dx = \delta_{nm} 2^n n! \sqrt{\pi}\,.$$

In QM this particular basis is important when one studies the physical system known as the *one-dimensional harmonic oscillator*.

With the same procedure on $f(x) := e^{-x/2}$ we find a basis of $L^2((0, +\infty), dx)$ given, up to renormalisation, by **Laguerre's functions** $e^{-x} L_n(x)$, $n = 0, 1, \ldots$. The polynomial L_n has degree n and goes under the name of nth **Laguerre polynomial**. Laguerre polynomials are obtained from the formula:

$$L_n(x) = e^x \frac{d^n}{dx^n} (x^n e^{-x}) \quad n = 0, 1, 2, \ldots$$

Again, we have the normalising relations:

$$\int_{[0,+\infty)} e^{-x} L_n(x) L_m(x) dx = \delta_{nm} (n!)^2\,.$$

In QM the basis of Laguerre functions is important when working with physical systems having a spherical symmerty, like the hydrogen atom, for instance.

(5) Consider the separable Hilbert space $L^2(\mathbb{R}^n, dx)$ (dx being the usual Lebesgue measure on $\mathbb{R}^n$). It is a renowned fact [Vla81] that the spaces of real-valued (or complex-valued) smooth functions on $\mathbb{R}^n$ with compact support (respectively, that decrease at infinity faster than any negative power of $|x|$) are dense subspaces of $L^p(\mathbb{R}^n, dx)$, $(1 \leq p < \infty)$. It falls out of Proposition 3.31(b) that such subspaces contain bases of $L^2(\mathbb{R}^n, dx)$.

(6) We will now construct the so-called *Bargmann-Hilbert space*, also known as *Bargmann-Fock space*. This is a Hilbert space with a host of applications in QM and quantum field theory. Consider the following positive σ-additive measure defined on Borel sets $E \subset \mathbb{C}$, where χ_E is the **characteristic function** of E ($\chi_E(z) = 1$ if $z \in E$, $\chi_E(z) = 0$ if $z \notin E$):

$$\mu(E) := \frac{1}{\pi} \int_{\mathbb{C}} \chi_E(z) e^{-|z|^2} dz d\bar{z}\,.$$

Here, as is customary in this formalism, we denoted by $dzd\bar{z}$ the Lebesgue measure of $\mathbb{R}^2$ identified with $\mathbb{C}$. A function $f : \mathbb{C} \to \mathbb{C}$ is **entire** if it is holomorphic everywhere on $\mathbb{C}$. Let $\mathsf{H}(\mathbb{C})$ be the space of entire functions. Now take the vector subspace of $L^2(\mathbb{C}, \mu)$ given by the intersection $L^2(\mathbb{C}, \mu) \cap \mathsf{H}(\mathbb{C})$ – where the elements of $\mathsf{H}(\mathbb{C})$ represent equivalence classes of maps, as is the case when defining L^p spaces. It is far from obvious that $L^2(\mathbb{C}, \mu) \cap \mathsf{H}(\mathbb{C})$ is a closed subspace of $L^2(\mathbb{C}, \mu)$, because non-evident is that a sequence of entire functions converges *in* $L^2(\mathbb{C}, \mu)$ *sense* to an entire function (i.e. the limit is entire up to zero-measure sets). Bargmann, however, managed to prove [Bar61] that

$$\text{if } f \in \mathsf{H}(\mathbb{C}), \text{ then} \quad \int_{\mathbb{C}} |f(z)|^2 d\mu(z) = \sum_{n=0}^{+\infty} |f_n|^2 \leq +\infty \tag{3.29}$$

where:

$$f_n = \sqrt{n!}a_n \quad \text{with} \quad f(z) = \sum_{n=0}^{+\infty} a_n z^n . \tag{3.30}$$

The power series in (3.30) is just the Taylor expansion of f: it converges absolutely for any $z \in \mathbb{C}$ and uniformly on any compact set in $\mathbb{C}$, and it exists by the mere fact that f is entire. Notice that (3.29) establishes in particular that the positive-term series on the right converges iff the integral of the left-hand-side function converges. Hence $f, g \in L^2(\mathbb{C}, \mu) \cap \mathsf{H}(\mathbb{C})$ if and only if $\{f_n\}_{n\in\mathbb{N}}, \{g_n\}_{n=1,2...} \in \ell^2(\mathbb{N})$ (Example 2.26(7)), and if so the polarisation formula (3.4) and (3.29) give:

$$\int_{\mathbb{C}} \overline{f(z)} g(z) d\mu(z) = \sum_{n=0}^{+\infty} \overline{f_n} g_n . \tag{3.31}$$

In the notation of (3.30), let us consider the map:

$$J : L^2(\mathbb{C}, \mu) \cap \mathsf{H}(\mathbb{C}) \ni f \mapsto \{f_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N}) .$$

This linear isometric (hence injective) transformation is actually onto as well. In fact, since the the series $\sum_{n\in\mathbb{N}} \frac{|z|^{2n}}{(n!)^2}$ converges for any $z \in \mathbb{C}$, the Cauchy-Schwarz inequality implies that the series:

$$\sum_{n\in\mathbb{N}} \frac{c_n}{\sqrt{n!}} \frac{z^n}{\sqrt{n!}} =: f(z)$$

converges absolutely for any $z \in \mathbb{C}$, provided $\{c_n\}_{n\in\mathbb{N}} \in \ell^2(\mathbb{N})$ and if we define an entire map f and $J(f) = \{c_n\}_{n\in\mathbb{N}}$. Since $\ell^2(\mathbb{N})$ is complete we conclude that:

(a) the complex vector space $\mathsf{B}_1 := L^2(\mathbb{C}, \mu) \cap \mathsf{H}(\mathbb{C})$ is a Hilbert space, i.e. a *closed* subspace of $L^2(\mathbb{C}, \mu)$;
(b) this Hilbert space is isomorphic to $\ell^2(\mathbb{N})$ under J (hence in particular separable);
(c) the system of entire maps $\{u_n\}_{n\in\mathbb{N}}$:

$$u_n(z) = \frac{z^n}{\sqrt{n!}} \quad \text{for any } z \in \mathbb{C}, n \in \mathbb{N} \tag{3.32}$$

is a basis for B_1. B_1 is called **Bargmann-Hilbert** or **Bargmann-Fock space**.

To conclude we observe that all constructions have a straightforward generalisation to n copies of $\mathbb{C}$, giving the n-dimensional Bargmann space $\mathsf{B}_n := L^2(\mathbb{C}^n, d\mu_n) \cap \mathsf{H}(\mathbb{C}^n)$ where, for any Borel set $E \in \mathbb{C}^n$:

$$\mu_n(E) := \frac{1}{\pi^n} \int_{\mathbb{C}^n} \chi_E(z) e^{-\sum_{k=1}^n |z_k|^2} dz_1 d\overline{z_1} \otimes \cdots \otimes dz_n d\overline{z_n} ,$$

$\mathsf{H}(\mathbb{C}^n)$ is the space of holomorphic maps in n variables on $\mathbb{C}^n$ and the integral is computed in the product of measures μ on each copy of $\mathbb{C}$. ■

3.3 Hermitian adjoints and applications

We examine here one of the most important notions of the theory of operators on a Hilbert space descending from Riesz's Theorem 3.16: *(Hermitian) adjoint operators*. We have to stress that this notion can be extended to unbounded operators. In this section we consider *only* the bounded case. The general situation will be dealt with extensively in a subsequent chapter. It is also worth recalling that a (related) notion of adjoint operator (or conjugate operator) was given in Definition 2.42, without the need of Hilbert structures. In the sequel we will not use the latter non-Hilbert notion, exception made for a few remarks.

3.3.1 Hermitian conjugation, or adjunction

Let $(\mathsf{H}_1, (\,|\,)_1)$, $(\mathsf{H}_2, (\,|\,)_2)$ be Hilbert spaces and $T \in \mathfrak{B}(\mathsf{H}_1, \mathsf{H}_2)$. For a given $u \in \mathsf{H}_2$, consider:

$$\mathsf{H}_1 \ni v \mapsto (u|Tv)_2 \in \mathbb{C}\,. \tag{3.33}$$

This map is certainly linear and bounded:

$$|(u|Tv)_2| \leq ||u||_2\, ||Tv||_2 \leq ||u||_2\, ||T||\, ||v||_1\,.$$

Hence it belongs to H_1'. By Riesz's Theorem 3.16 there is $w_{T,u} \in \mathsf{H}_1$ such that

$$(u|Tv)_2 = (w_{T,u}|v)_1\,, \quad \text{for any } v \in \mathsf{H}_1\,. \tag{3.34}$$

Moreover, the map $\mathsf{H}_2 \ni u \mapsto w_{T,u} \in \mathsf{H}_1$ is linear. In fact:

$$(w_{T,\alpha u+\beta u'}|v)_1 = (\alpha u + \beta u'|Tv)_2 = \overline{\alpha}(u|Tv)_2 + \overline{\beta}(u'|Tv)_2 = (\alpha w_{T,u} + \beta w_{T,u'}|v)_1\,,$$

so, for any $v \in \mathsf{H}_1$,

$$0 = (w_{T,\alpha u+\beta u'} - \alpha w_{T,u} - \beta w_{T,u'}|v)_1\,.$$

Choosing $v := w_{T,\alpha u+\beta u'} - \alpha w_{T,u} - \beta w_{T,u'}$, we have $w_{T,\alpha u+\beta u'} - \alpha w_{T,u} - \beta w_{T,u'} = 0$ hence

$$w_{T,\alpha u+\beta u'} = \alpha w_{T,u} + \beta w_{T,u'}$$

for any $\alpha, \beta \in \mathbb{C}$, $u, u' \in \mathsf{H}_2$. Therefore there exists a linear operator:

$$T^* : \mathsf{H}_2 \ni u \mapsto w_{T,u} \in \mathsf{H}_1\,.$$

By construction the latter satisfies $(u|Tv)_2 = (T^*u|v)_1$ for any pair $u \in \mathsf{H}_2$, $v \in \mathsf{H}_1$ and T^* is the unique linear operator with such property. If there were another such $B \in \mathfrak{L}(\mathsf{H}_2, \mathsf{H}_1)$, then $(T^*u|v)_1 = (Bu|v)_1$ for any $v \in \mathsf{H}_1$. Consequently $((T^* - B)u|v)_1 = 0$ for any $v \in \mathsf{H}_1$. Choosing $v := (T^* - B)u$ it follows $||(T^* - B)u||_1^2 = 0$, so $T^*u - Bu = 0$. Since $u \in \mathsf{H}_2$ is arbitrary, $T^* = B$. Overall, we proved the following fact.

Proposition 3.36. *Let* $(\mathsf{H}_1, (\,|\,)_1)$, $(\mathsf{H}_2, (\,|\,)_2)$ *be Hilbert spaces, and* $T \in \mathfrak{B}(\mathsf{H}_1, \mathsf{H}_2)$. *There exists a unique linear operator* $T^* : \mathsf{H}_2 \to \mathsf{H}_1$ *such that:*

$$(u|Tv)_2 = (T^*u|v)_1\,, \quad \textit{for any pair } u \in \mathsf{H}_2,\ v \in \mathsf{H}_1. \tag{3.35}$$

We are ready to define *adjoint Hermitian operators*. From now on we will drop the adjective "Hermitian", given that this textbook will never use non-Hermitian adjoint operators as we said at the beginning.

Definition 3.37. *Let* $(\mathsf{H}_1, (\,|\,)_1)$, $(\mathsf{H}_2, (\,|\,)_2)$ *be Hilbert spaces and* $T \in \mathfrak{B}(\mathsf{H}_1, \mathsf{H}_2)$*. The unique linear operator* $T^* \in \mathfrak{L}(\mathsf{H}_2, \mathsf{H}_1)$ *fulfilling (3.35) is called the* **(Hermitian) adjoint***, or* **Hermitian conjugate** *to the operator* T.

Recall that given a linear operator $T : \mathsf{X} \to \mathsf{Y}$ between vector spaces, $Ran(T) := \{T(x) \mid x \in \mathsf{X}\}$ and $Ker(T) := \{x \in \mathsf{X} \mid T(x) = 0\}$ denote the subspaces of Y and X called **range** (or **image**) and **kernel** (or **null space**) of T.

The operation of Hermitian conjugation enjoys the following elementary properties.

Proposition 3.38. *Let* $(\mathsf{H}_1, (\,|\,)_1)$, $(\mathsf{H}_2, (\,|\,)_2)$ *be Hilbert spaces and* $T \in \mathfrak{B}(\mathsf{H}_1, \mathsf{H}_2)$*. Then:*

(a) $T^* \in \mathfrak{B}(\mathsf{H}_2, \mathsf{H}_1)$*, and more precisely:*

$$||T^*|| = ||T||\,, \tag{3.36}$$

$$||T^*T|| = ||T||^2 = ||TT^*||\,. \tag{3.37}$$

(b) *Hermitian conjugation is* **involutive***:*

$$(T^*)^* = T\,.$$

(c) *If* $S \in \mathfrak{B}(\mathsf{H}_1, \mathsf{H}_2)$ *and* $\alpha, \beta \in \mathbb{C}$*:*

$$(\alpha T + \beta S)^* = \overline{\alpha} T^* + \overline{\beta} S^*\,, \tag{3.38}$$

and if $S \in \mathfrak{B}(\mathsf{H}, \mathsf{H}_1)$*, with* H *Hilbert space:*

$$(TS)^* = S^* T^*\,. \tag{3.39}$$

(d) *We have:*

$$Ker(T) = [Ran(T^*)]^{\perp}\,, \quad Ker(T^*) = [Ran(T)]^{\perp}\,. \tag{3.40}$$

(e) T *is bijective if and only if* T^* *is bijective, in which case* $(T^*)^{-1} = (T^{-1})^*$.

Proof. From now on we will write $||\;||$ to denote both $||\;||_1$ and $||\;||_2$, and similarly for inner products. Which norm or inner product will be clear from the context.
(a) For any pair $u \in \mathsf{H}_2$, $x \in \mathsf{H}_1$ we have $|(T^*u|x)| = |(u|Tx)| \leq ||u||\,||T||\,||x||$. By choosing $x := T^*u$ we have in particular $||T^*u||^2 \leq ||T||\,||u||\,||T^*u||$, so $||T^*u|| \leq ||T||\,||u||$. Hence T^* is bounded and $||T^*|| \leq ||T||$. Therefore it makes sense to define $(T^*)^*$, so $||(T^*)^*|| \leq ||T^*||$. This inequality becomes $||T|| \leq ||T^*||$ by (b) (whose proof only uses the boundedness of T^*). As $||T^*|| \leq ||T||$ and $||T|| \leq ||T^*||$, equation (3.36) follows. Let us pass to (3.37). It suffices to prove the first identity, since the second descends from the first one and (3.36), by (b) (which does not depend on (a)).

By Theorem 2.41(b), case (i), whose conclusion holds for $S \in \mathfrak{B}(\mathsf{Y},\mathsf{X})$, $T \in \mathfrak{B}(\mathsf{Z},\mathsf{Y})$ with $\mathsf{X},\mathsf{Y},\mathsf{Z}$ normed, we have $||T^*T|| \leq ||T^*||\,||T|| = ||T||^2$. At the same time:

$$||T||^2 = (\sup_{||x||\leq 1} ||Tx||)^2 = \sup_{||x||\leq 1} ||Tx||^2 = \sup_{||x||\leq 1} (Tx|Tx)\,.$$

By definition of adjoint and by Cauchy-Schwarz (on the last term) we obtain:

$$||T||^2 = \sup_{||x||\leq 1} (Tx|Tx) = \sup_{||x||\leq 1} |(T^*Tx|x)| \leq \sup_{||x||\leq 1} ||T^*Tx|| = ||T^*T||\,.$$

Therefore $||T^*T|| \leq ||T||^2$ and $||T||^2 \leq ||T^*T||$, so $||T^*T|| = ||T||^2$.
(b) This follows immediately from the uniqueness of the adjoint operator. By known properties of the inner product and the definition of adjoint to T, in fact, we have:

$$(v|T^*u) = \overline{(T^*u|v)} = \overline{(u|Tv)} = (Tv|u)\,.$$

(c) If $u \in \mathsf{H}_2$, $v \in \mathsf{H}_1$ then

$$(u|(\alpha T + \beta S)v) = \alpha(u|Tv) + \beta(u|Sv) = \alpha(T^*u|v) + \beta(S^*u|v) = ((\overline{\alpha}T^* + \overline{\beta}S^*)u|v)\,.$$

The adjoint's uniqueness gives (3.38). If $v \in \mathsf{H}$, $u \in \mathsf{H}_2$,

$$(u|(TS)v) = (T^*u|Sv) = ((S^*T^*)u|v)\,.$$

By uniqueness (3.39) holds.
(d) It is enough to prove the first identity, as the second one is a consequence of it and of part (b). Since $(T^*u|v) = (u|Tv)$, if $v \in Ker(T)$ then $(T^*u|v) = 0$ for any $u \in \mathsf{H}_2$, so $v \in [Ran(T^*)]^\perp$. Conversely, still for $(T^*u|v) = (u|Tv)$, if $v \in [Ran(T^*)]^\perp$ then $(u|Tv) = 0$ for any $u \in \mathsf{H}_2$. If we choose $u := Tv$ then $Tv = 0$ and so $v \in Ker(T)$.
(e) If T is bijective T^{-1} is bounded by Banach's inverse operator theorem. Therefore $(T^{-1})^*$ exists. We have: $T^{-1}T = TT^{-1} = I$. Let us compute the adjoint of both sides, using the second property of (c) and remembering $I^* = I$: $T^*(T^{-1})^* = (T^{-1})^*T^* = I$. These are equivalent to saying T^* is bijective and $(T^*)^{-1} = (T^{-1})^*$. Eventually, if T^* is bijective, then also $(T^*)^* = T$ is bijective, for what we have just proved, by (b). □

Remarks 3.39. The relationship between Hermitian adjoints and conjugate operators seen in Definition 2.42 arises as follows. Start with $T \in \mathfrak{B}(\mathsf{H}_1,\mathsf{H}_2)$ and compute the conjugate $T' \in \mathfrak{B}(\mathsf{H}_2',\mathsf{H}_1')$ and the adjoint $T^* \in \mathfrak{B}(\mathsf{H}_2,\mathsf{H}_1)$. Then:

$$(T^*y_f|x)_1 = (y_f|Tx)_2 = (T'f)(x) \quad \text{for any } f \in \mathsf{H}_2',\, x \in \mathsf{H}_1,$$

where $f \in \mathsf{H}_2'$, and $y_f \in \mathsf{H}_2$ is the element in H_2 representing f under Riesz's Theorem 3.16. Because $x \in \mathsf{H}_1$ is arbitrary, we may write:

$$T'f = (T^*y_f|\ \ )_1 \quad \text{for any } f \in \mathsf{H}_2'. \tag{3.41}$$

Given that the Riesz map $\mathsf{H}_2' \ni f \mapsto y_f \in \mathsf{H}_2$ is bijective, the above equation determines T' completely whenever T^* is given, and conversely. ■

3.3.2 *-algebras and C^*-algebras

The operation of Hermitian conjugation is an excuse for introducing one of the most useful mathematical concepts in advanced formulations of QM: we are talking about *C^*-algebras* (also known as *B^*-algebras*). We shall return to this notion in Chapter 8 when discussing the spectral decomposition theorem.

Definition 3.40. *Let $\mathfrak{A}$ be a (commutative, Banach) algebra (with unit, normed by $|| \; ||$) over the field $\mathbb{C}$. A map $^* : \mathfrak{A} \to \mathfrak{A}$ such that:*

I1. *(antilinearity)* $(\alpha x + \beta y)^* = \overline{\alpha} x^* + \overline{\beta} y^*$ *for any* $x, y \in \mathfrak{A}$, $\alpha, \beta \in \mathbb{C}$;
I2. *(involutivity)* $(x^*)^* = x$ *for any* $x \in \mathfrak{A}$;
I3. $(xy)^* = y^* x^*$ *for any* $x, y \in \mathfrak{A}$,

is called **involution** *and the structure* $(\mathfrak{A}, ^*)$ *is a* **-algebra (respectively commutative, Banach, with unit, normed). A Banach* **-algebra (with unit) is a* C^***-algebra** *(with unit) if, further:*

$$||x^* x|| = ||x||^2 \, . \tag{3.42}$$

A homomorphism of **-algebras* $f : \mathfrak{A}_1 \to \mathfrak{A}_2$ *is a* ***-homomorphism** *if it preserves the involution:* $f(x^{*_1}) = f(x)^{*_2}$ *for any* $x \in \mathfrak{A}_1$ *(*^{*_1} *is the involution of* $\mathfrak{A}_1$ *and* ^{*_2} *the involution in* $\mathfrak{A}_2$*), and a* **-homomorphism is a* ***-isomorphism** *if it is additionally bijective. An element* x *in a* **-algebra* $\mathfrak{A}$ *(with unit* $\mathbb{I}$ *in cases (iii), (iv) below) is said:*

(i) **normal** *if* $xx^* = x^* x$;
(ii) **Hermitian** *or* **self-adjoint** *if* $x^* = x$;
(iii) **isometric** *if* $x^* x = \mathbb{I}$;
(iv) **unitary** *if* $x^* x = xx^* = \mathbb{I}$.

A ***-subalgebra** (C^***-subalgebra**) of a *-algebra (C^*-algebra) $\mathfrak{A}$ is the natural object: a subalgebra that is a *-algebra (C^*-algebra) for the restricted involution (and for the restricted Banach structure in case of a C^*-subalgebra). If the *-algebra (C^*-algebra) has a unit, any *-subalgebra (C^*-subalgebra) with unit is required to have the same unit of the *-algebra.

Remark 3.41. (1) If $\mathfrak{A}$ is a *-algebra (with unit), and $\{\mathfrak{A}_i\}_{i \in I}$ is a collection of *-subalgebras (with unit), it is easy to see $\bigcap_{i \in \mathrm{i}} \mathfrak{A}_i$ is a *-subalgebra (with unit) of $\mathfrak{A}$. If we add the topological structure and $\{\mathfrak{A}_i\}_{i \in I}$ are C^*-subalgebras (with unit) of the C^*-algebra (with unit) $\mathfrak{A}$, then $\bigcap_{i \in \mathrm{i}} \mathfrak{A}_i$ is a C^*-subalgebra (with unit) of $\mathfrak{A}$. The only not completely obvious fact is the completeness of $\bigcap_{i \in \mathrm{i}} \mathfrak{A}_i$, but this follows directly from the fact it is closed, hence complete, being intersection of closed (complete) sets $\mathfrak{A}_i$.
(2) If $S \subset \mathfrak{A}$ is a subset in $\mathfrak{A}$ *-algebra (with unit), the *-algebra (with unit) **generated** by S is the intersection of all *-subalgebras (with unit) in $\mathfrak{A}$ that contain S. The same holds for C^*-algebras (with unit), *mutatis mutandis*. ■

Before we return to $\mathfrak{B}(\mathsf{H})$ (and show it is a C^*-algebra), let us see a few general features of *-algebras that descend from the definition.

Proposition 3.42. *Let $\mathfrak{A}$ be a *-algebra with involution *.*

(a) *If $(\mathfrak{A}, ||\ ||)$ is a C^*-algebra and $x \in \mathfrak{A}$ is normal, then for any $m = 1, 2, \ldots$:*

$$||x^m|| = ||x||^m .$$

(b) *If $(\mathfrak{A}, ||\ ||)$ is a C^*-algebra and $x \in \mathfrak{A}$,*

$$||x^*|| = ||x|| .$$

(c) *If $\mathfrak{A}$ has unit $\mathbb{I}$, then $\mathbb{I}^* = \mathbb{I}$. Moreover, $x \in \mathfrak{A}$ has an inverse if and only if x^* has an inverse, in which case $(x^{-1})^* = (x^*)^{-1}$.*

Proof. (a) If $||x|| = 0$ the claim is trivial, so assume $x \neq 0$. A repeated use of (3.42), **I2**, **I3** and the fact that $xx^* = x^*x$ gives:

$$||x^2||^2 = ||(x^2)^*x^2|| = ||(x^*)^2x^2|| = ||(x^*x)^*(x^*x)|| = ||x^*x||^2 = (||x||^2)^2$$

whence $||x^2|| = ||x||^2$ by norm positivity. Iterating we obtain $||x^{2^k}|| = ||x||^{2^k}$ for any natural number k. If $m = 3, 4, \ldots$ there exist two natural numbers n, k with $m + n = 2^k$, thus:

$$||x||^m||x||^n = ||x||^{n+m} = ||x^{n+m}|| \leq ||x^m||\,||x^n|| \leq ||x^m||\,||x||^n \leq ||x||^m||x||^n .$$

But then all inequalities are equalities, so in particular:

$$||x^m||\,||x||^n = ||x||^m||x||^n,$$

dividing which by $||x||^n$ (non-zero since $x \neq 0$ and $||\cdot||$ is a norm) proves the claim.
(b) Equation (3.42) implies $||x||^2 = ||xx^*|| \leq ||x||\,||x^*||$ so $||x|| \leq ||x^*||$. Similarly $||x^*|| \leq ||(x^*)^*||$, and then $(x^*)^* = x$ finishes this part.
(c) $\mathbb{I}\mathbb{I}^* = \mathbb{I}^*$ by definition of unit; on the other hand $\mathbb{I}\mathbb{I}^* = (\mathbb{I}^*)^*\mathbb{I}^* = (\mathbb{I}^*\mathbb{I})^*$. From these two descends $\mathbb{I}^* = (\mathbb{I}^*\mathbb{I})^* = (\mathbb{I}^*)^* = \mathbb{I}$. The other statement follows from this, **I2** and uniqueness of the inverse. □

Remarks 3.43. The structure of a C^*-algebra is remarkable in that its topological and algebraic properties are deeply intertwined. We will prove later (Corollary 8.18) that if $\phi : \mathfrak{A} \to \mathfrak{B}$ is a *-homomorphism between C^*-algebras with unit, then it is continuous because $||\phi(a)||_{\mathfrak{B}} \leq ||a||_{\mathfrak{A}}$ for any $a \in \mathfrak{A}$. Moreover, ϕ is isometric, i.e. $||\phi(a)||_{\mathfrak{B}} = ||a||_{\mathfrak{A}}$ for any $a \in \mathfrak{A}$, if and only if it is injective (Theorem 8.22). ■

Examples 3.44. (1) The Banach algebras of complex-valued functions seen in Examples 2.26(2), (3), (4), (8) and (9) are all instances of commutative C^*-algebras whose involution is the complex conjugation of functions.

(2) By virtue of (a), (b), (c) in Proposition 3.38 we have this result.

Theorem 3.45. *If H is Hilbert space, $\mathfrak{B}(\mathsf{H})$ is a C^*-algebra with unit if the involution is defined as the Hermitian conjugation.*

(3) An example of C^*-algebra, absolutely fundamental for the applications in quantum field theory (but not only) is the *von Neumann algebra*. Before we introduce it, let us define first the *commutant* of an operator algebra and state an important preliminary theorem. If $\mathfrak{M} \subset \mathfrak{B}(\mathsf{H})$ is a subset in the algebra of bounded operators on the complex Hilbert space $\mathfrak{B}(H)$, the **commutant** of $\mathfrak{M}$ is:

$$\mathfrak{M}' := \{T \in \mathfrak{B}(\mathsf{H}) \mid TA - AT = 0 \quad \text{for any } A \in \mathfrak{M}\} .$$

If $\mathfrak{M}$ is closed under Hermitian conjugation (i.e. $A^* \in \mathfrak{M}$ if $A \in \mathfrak{M}$) the commutant $\mathfrak{M}'$ is surely a *-algebra with unit. In general: $\mathfrak{M}_1' \subset \mathfrak{M}_2'$ if $\mathfrak{M}_2 \subset \mathfrak{M}_1$ and $\mathfrak{M} \subset (\mathfrak{M}')'$, which imply $\mathfrak{M}' = ((\mathfrak{M}')')'$. Hence we cannot reach beyond the second commutant by iteration.

The continuity of the product of operators says that the commutant $\mathfrak{M}'$ is closed in the uniform topology, so if $\mathfrak{M}$ is closed under Hermitian conjugation, its commutant $\mathfrak{M}'$ is a C^*-algebra (C^*-subalgebra) in $\mathfrak{B}(\mathsf{H})$.

$\mathfrak{M}'$ has other pivotal topological properties in this general setup. It is easy to prove $\mathfrak{M}'$ is both strongly and weakly closed. This holds, despite the product of operators is not continuous, because separate continuity in each variable is sufficient.

In the sequel we shall adopt the standard convention used for von Neumann algebras and write $\mathfrak{M}''$ in the stead of $(\mathfrak{M}')'$ *et c*. The next crucial result is due to von Neumann [BrRo02]. In the following we will make use of some notions presented in the Section 3.4.

Theorem 3.46 (Double commutant theorem). *If* H *is a complex Hilbert space and* $\mathfrak{A}$ *a* **-subalgebra in* $\mathfrak{B}(\mathsf{H})$ *(in particular* $\mathfrak{A}$ *contains the identity operator), the following statements are equivalent.*

(a) $\mathfrak{A} = \mathfrak{A}''$.
(b) $\mathfrak{A}$ *is weakly closed and* $\mathbb{I} \in \mathfrak{A}$.
(c) $\mathfrak{A}$ *is strongly closed and* $\mathbb{I} \in \mathfrak{A}$.

Proof. (a) implies (b) because the commutant S' of a set $S \subset \mathfrak{B}(\mathsf{H})$ is closed in the weak topology, as we can see directly: if $A_n \in S'$ for any $n \in \mathbb{N}$ and $A_n \to A \in \mathfrak{B}(\mathsf{H})$ weakly, then $A_n^* \to A^*$ weakly (this is immediate), so for $B \in S$:

$$0 = (\psi|(A_nB - BA_n)\phi) = \overline{(B\phi|A_n^*\psi)} - (B^*\psi|A_n\phi) \to (A\psi|B\phi) - (B^*\psi|A\phi)$$

$$= (\psi|(AB - BA)\phi) ,$$

hence $(\psi|(AB-BA)\phi) = 0$. Being ψ, ϕ arbitrary in H, A commutes with B, and $A \in S'$ because $B \in S$ is generic.
(b) implies (c) because strong convergence implies weak convergence, and therefore a limit point in the strong topology remains so in the weak one.
To finish we have to show (c) implies (a). Since $S \subset S''$ for any set and the commutant is strongly closed, we only need to prove that (c) forces $\mathfrak{A}''$ to be in the strong closure of $\mathfrak{A}$. This proof relies on facts that will be introduced in Section 3.4 which are independent of the present theorem. Now we prove that, under (c), $X \in \mathfrak{A}''$ implies $X \in \mathfrak{A}$.

For this, for $\psi \in \mathsf{H}$ define the closed subspace: $\mathsf{H}_\psi := \overline{\{A\psi \,|\, A \in \mathfrak{A}\}}$, and by P denote the orthogonal projector onto H_ψ. By construction, if $\phi \in \mathsf{H}_\psi$ then $PB\phi = B\phi$ for any $B \in \mathfrak{A}$. That is to say $PBP = BP$ for $B \in \mathfrak{A}$. Taking the adjoint gives $PB^* = B^*P$, since $P = P^*$ by definition of orthogonal projector. Since $B^* = A$ for some $A \in \mathfrak{A}$, and since B varies in the whole $\mathfrak{A}$ as $A \in \mathfrak{A}$, we conclude $P \in \mathfrak{A}'$. Therefore for any $X \in \mathfrak{A}''$ we have $PX = XP$. But $I \in \mathfrak{A}$, so $\psi \in \mathsf{H}_\psi$ and $X\psi \in \mathsf{H}_\psi$ (since $PX\psi = XP\psi = X\psi$). By definition of H_ψ, $X\psi \in \mathsf{H}_\psi$ implies that for any $\varepsilon > 0$ there exists $A \in \mathfrak{A}$ with $||A\psi - X\psi|| < \varepsilon$. Consider then a finite collection of vectors $\psi_1, \psi_2, \ldots, \psi_n$ and the n-fold orthogonal sum $\mathsf{H}_n := \mathsf{H} \oplus \cdots \oplus \mathsf{H}$. On this Hilbert space take the algebra $\mathfrak{A}_n \subset \mathfrak{B}(\mathsf{H}_n)$ of operators of the form $A \oplus \cdots \oplus A : (v_1, \ldots, v_n) \mapsto (Av_1, \ldots, Av_n)$, with $v_k \in \mathsf{H}$ for $k = 1, \ldots, n$, $A \in \mathfrak{A}$. It is immediate that $\mathfrak{A}_n$ is a *-subalgebra (with unit) in $\mathfrak{B}(\mathsf{H}_n)$, and strongly closed ($\mathfrak{A}$ is, by assumption). If $X \in \mathfrak{A}''$, then $X \oplus \cdots \oplus X \in \mathfrak{A}''_n$. With the same argument as before, for any $\varepsilon > 0$ there exists $A \in \mathfrak{A}$, $||A\psi_k - X\psi_k|| < \varepsilon$, for $k = 1, \ldots, n$. By definition of strong topology, this implies that if $X \in \mathfrak{A}''$ then X belongs to the strong closure of $\mathfrak{A}$. By (c), the same implies $X \in \mathfrak{A}$. □

At this juncture we are ready to define von Neumann algebras.

Definition 3.47. *A* **von Neumann algebra** *in* $\mathfrak{B}(\mathsf{H})$ *is a* **-subalgebra of* $\mathfrak{B}(\mathsf{H})$*, with unit, that satisfies either equivalent property appearing in von Neumann's Theorem 3.46.*

In particular $\mathfrak{M}'$ is a von Neumann algebra provided $\mathfrak{M}$ is a *-closed subset of $\mathfrak{B}(\mathsf{H})$, because $(\mathfrak{M}')'' = \mathfrak{M}'$ as we saw above. Note how, by construction, a von Neumann algebra in $\mathfrak{B}(\mathsf{H})$ is a C^*-algebra with unit, or better, a C^*-subalgebra with unit of $\mathfrak{B}(\mathsf{H})$.

It is not hard to see that the intersection of two von Neumann algebras is a von Neumann algebra. If $\mathfrak{M} \subset \mathfrak{B}(\mathsf{H})$ is closed under Hermitian conjugation, $\mathfrak{M}''$ turns out to be the smallest (set-theoretically) von Neumann algebra containing $\mathfrak{M}$ as a subset [BrRo02]. Thus $\mathfrak{M}''$ is called the **von Neumann algebra generated** by $\mathfrak{M}$.

A von Neumann algebra $\mathfrak{R}$ is a **factor** when its **centre**, the subset $\mathfrak{R} \cap \mathfrak{R}'$ of elements commuting with the whole algebra, is trivial: $\mathfrak{R} \cap \mathfrak{R}' = \{cI\}_{c \in \mathbb{C}}$. The classification of factors, started by von Neumann and Murray, is one of the key chapters in the theory of operator algebras, and has enourmous consequences in the algebraic theory of quantum fields.

(4) The algebra $\mathbb{H}$ of **quaternions** is a 4-dimensional real vector space with a privileged basis $\{1, i, j, k\}$. $\mathbb{H}$ is equipped with a product that makes it an $\mathbb{R}$-algebra, and has a unit, given by the basis element 1. The product is determined, keeping in mind Definition 2.23, by the relations $ii = jj = kk = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$.

As the field $\mathbb{R}$ identifies naturally with the Abelian subalgebra of $\mathbb{H}$ of quaternions of the form $a1$, $a \in \mathbb{R}$, $\mathbb{H}$ can be viewed as a real normed algebra with unit: it is enough to think real numbers as quaternions, and define the product of a real scalar by a quaternion using the product in $\mathbb{H}$. The norm is the usual Euclidean norm for the standard basis of $\mathbb{H}$: $||a1 + bi + cj + dk|| := \sqrt{a^2 + b^2 + c^2 + d^2}$. It is easy to check that $\mathbb{H}$ becomes thus a real Banach algebra with unit. Although the field is $\mathbb{R}$, it is possible to define an involution on $\mathbb{H}$ via quaternionic conjugation: $(a1+bi+cj+dk)^* = a1-bi-cj-dk$, $a,b,c,d \in \mathbb{R}$. Then the usual properties of involutions hold (the field is real, so the involution is $\mathbb{R}$-linear) also concerning the norm, and also the property typical of C^*-algebras: $||a^*a|| = ||a||^2$. The relationship between product and norm is precisely $||ab|| = ||a||\,||b||$, $a,b \in \mathbb{H}$, reminiscent of $\mathbb{C}$ with modulus and $\mathbb{R}$ with absolute value. A further property, valid for $\mathbb{R}$ and $\mathbb{C}$ as well, is that the quaternion algebra is a **division algebra**, i.e. an algebra with unit where any non-zero element is invertible.

A concrete representation of $\mathbb{H}$ is given by the real subalgebra of $M(2,\mathbb{C})$ (2×2 complex matrices) spanned over $\mathbb{R}$ by the identity I and the three **Pauli matrices** $-i\sigma_1, -i\sigma_2, -i\sigma_3$: these correspond to the quaternionic units 1 and i,j,k (see Remark 7.26(3)). ■

To conclude the section let us see a very relevant definition in advanced formalisations of, for example, quantum field theories.

Definition 3.48. *Given a* **-algebra* $\mathfrak{A}$ *(not necessarily with unit nor* C^**) and a Hilbert space* H*, a* **-homomorphism* $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ *(preserving the unit if present) is called a* **representation of** $\mathfrak{A}$ **on** H.

(a) π *is* **faithful** *if one-to-one.*
(b) *A subspace* $\mathsf{M} \subset \mathsf{H}$ *is* **invariant** *under* π *if* $\pi(a)(\mathsf{M}) \subset \mathsf{M}$ *for any* $a \in \mathfrak{A}$.
(c) π *is* **irreducible** *if there are no* π*-invariant closed subspaces other than* $\{0\}$ *and* H *itself.*
(d) *If* $\pi' : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}')$ *is another representation of* $\mathfrak{A}$ *on* H'*,* π *and* π' *are said* **unitarily equivalent***:*

$$\pi \simeq \pi'$$

if there exists a surjective isometry $U : \mathsf{H} \to \mathsf{H}'$ *such that:*

$$U\pi(a)U^{-1} = \pi'(a) \text{ for any } a \in \mathfrak{A}.$$

(e) *A vector* $\psi \in \mathsf{H}$ *is* **cyclic** *for* π *if* $\overline{\{\pi(a)\psi \,|\, a \in \mathfrak{A}\}} = \mathsf{H}$.

Remark 3.49. (1) One can consider representations of *-algebras in terms of unbounded operators and operators defined on a common invariant domain of the Hilbert space.
(2) In case $\mathfrak{A}$ is a C^*-algebra with unit, every representation is automatically continuous with respect to the norm of $\mathfrak{A}$ on the domain and the operator norm on the codomain, as $||\pi(a)|| \leq ||a||$ for any $a \in \mathfrak{A}$. Then π is faithful iff isometric: $||\pi(a)|| = ||a||$ for any $a \in \mathfrak{A}$. All this will be proved in Theorem 8.22. ■

Here is an elementary yet important fact.

Proposition 3.50. *If* $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ *is an irreducible representation of the* **-algebra* $\mathfrak{A}$ *on* H*, then every non-zero vector in* H *is cyclic for* π*, or* π *is the zero representation (that maps all of* $\mathfrak{A}$ *to the null operator) on* $\mathsf{H} = \mathbb{C}$.

Proof. Suppose π is not the trivial representation. Since π is irreducible, if $\psi \in \mathsf{H} \setminus \{0\}$ the subspace $\pi(\mathfrak{A})\psi$ must coincide with $\{0\}$; but then the closed subspace generated by ψ would be invariant, contradicting the assumption; thus $\pi(\mathfrak{A})\psi$ is dense in H, i.e. ψ is cyclic. □

3.3.3 Normal, self-adjoint, isometric, unitary and positive operators

Returning to the C^*-algebra $\mathfrak{B}(\mathsf{H})$ (or more generally to $\mathfrak{B}(\mathsf{H},\mathsf{H}_1)$), we can recall the most important types of operators we will encounter in subsequent chapters.

Definition 3.51. *Let* $(\mathsf{H},(\,|\,))$, $(\mathsf{H}_1,(\,|\,)_1)$ *be Hilbert spaces and* I_{H}, I_{H_1} *their respective identity operators.*

(a) $T \in \mathfrak{B}(\mathsf{H})$ *is said* **normal** *if* $TT^* = T^*T$.
(b) $T \in \mathfrak{B}(\mathsf{H})$ *is* **self-adjoint** *if* $T = T^*$.
(c) $T \in \mathfrak{L}(\mathsf{H},\mathsf{H}_1)$ *is* **isometric** *if bounded and* $T^*T = I_{\mathsf{H}}$*; equivalently,* $T \in \mathfrak{L}(\mathsf{H},\mathsf{H}_1)$ *is isometric if* $(Tx|Ty)_1 = (x|y)$ *for any pair* $x,y \in \mathsf{H}$.
(d) $T \in \mathfrak{L}(\mathsf{H},\mathsf{H}_1)$ *is* **unitary** *if bounded and* $T^*T = I_{\mathsf{H}}$, $TT^* = I_{\mathsf{H}_1}$*; equivalently,* $T \in \mathfrak{L}(\mathsf{H},\mathsf{H}_1)$ *is unitary if isometric and onto, i.e. an isomorphism of Hilbert spaces.*
(e) $T \in \mathfrak{L}(\mathsf{H})$ *is* **positive***, written* $T \geq 0$*, if* $(u|Tu) \geq 0$ *for any* $u \in \mathsf{H}$.
(f) *If* $U \in \mathfrak{L}(\mathsf{H})$*, we write* $T \geq U$ *in case* $T - U \geq 0$.

Remark 3.52. (1) Let us comment on the equivalence in (c): if $T \in \mathfrak{B}(\mathsf{H},\mathsf{H}_1)$ and $T^*T = I_{\mathsf{H}}$, then $(Tx|Ty)_1 = (x|y)$ for any pair $x,y \in \mathsf{H}$, since $(x|y) = (x|T^*Ty) = (Tx|Ty)_1$. On the other hand, if $T \in \mathfrak{L}(\mathsf{H},\mathsf{H}_1)$ and $(Tx|Ty)_1 = (x|y)$ for any $x,y \in \mathsf{H}$, then T is bounded (set $y = x$), so T^* is well defined; eventually, $T^*T = I_{\mathsf{H}}$ because $(x|T^*Ty) = (Tx|Ty)_1 = (x|y)$ for any pair $x,y \in \mathsf{H}$, so in particular $(x|(T^*T - I)y) = 0$ with $x = (T^*T - I)y$.

As for the equivalence in (d), notice that any isometric operator T is obviously injective, for $Tu = 0$ implies $||u|| = 0$ and hence $u = 0$. Thus surjectivity is equivalent to the existence of a right inverse that coincides with the left inverse (the latter exists by injectivity, and equals T^*). From this it follows immediately that $T^*T = I_{\mathsf{H}}$ and $TT^* = I_{\mathsf{H}_1}$ are together equivalent to saying $T \in \mathfrak{L}(\mathsf{H},\mathsf{H}_1)$ is isometric (hence bounded) and surjective. Our definition of a unitary operator agrees with Definition 3.10.
(2) There are isometric operators in $\mathfrak{B}(\mathsf{H})$ that are not unitary (this cannot happen if H has finite dimension). For instance, the operator on $\ell^2(\mathbb{N})$:

$$A : (z_0, z_1, z_2, \ldots) \mapsto (0, z_0, z_1, , \ldots),$$

for any $(z_0, z_1, z_2, \ldots) \in \ell^2(\mathbb{N})$.
(3) Unitary and self-adjoint operators in $\mathfrak{B}(\mathsf{H})$ are normal, but not conversely in general. ■

To close the section we consider one Hilbert space and discuss properties of normal, self-adjoint, unitary and positive operators. First, though, a definition that should be known from elementary courses.

Definition 3.53. *Let* X *be a vector space over* $\mathbb{K} = \mathbb{C}$, *or* $\mathbb{R}$, *and take* $T \in \mathfrak{L}(\mathsf{X})$; $\lambda \in \mathbb{K}$ *is an* **eigenvalue** *of* T *if:*

$$Tu = \lambda u \tag{3.43}$$

for some $u \in \mathsf{X} \setminus \{0\}$ *called an* **eigenvector** *of* T *relative (or associated) to* λ. *The subspace of* X *made of the null vector and all eigenvectors relative to a given eigenvalue* λ *is called the* **eigenspace** *of* T *with eigenvalue* λ *(of, associated to, relative to* λ*).*

Now here is the proposition summarising the aforementioned properties.

Proposition 3.54. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space.*
(a) *If* $T \in \mathfrak{B}(\mathsf{H})$ *is self-adjoint:*

$$||T|| = \sup\{|(x|Tx)| \mid x \in \mathsf{H}\,,\ ||x|| = 1\}\ . \tag{3.44}$$

More generally, if $T \in \mathfrak{L}(\mathsf{H})$ *satifies* $(x|Tx) = (Tx|x)$ *for any* $x \in \mathsf{H}$ *and the right-hand side of (3.44) is finite, then* T *is bounded.*
(b) *If* $T \in \mathfrak{B}(\mathsf{H})$ *is normal (in particular self-adjoint or unitary):*

- (i) $\lambda \in \mathbb{C}$ *is an eigenvalue of* T *with eigenvector* u *iff* $\overline{\lambda}$ *is an eigenvalue for* T^* *with the same eigenvector* u;
- (ii) *eigenspaces of* T *relative to distinct eigenvalues are orthogonal;*
- (iii) *the identity:*

$$||Tx|| = ||T^*x|| \quad \textit{for any } x \in \mathsf{H} \tag{3.45}$$

holds, so $Ker(T) = Ker(T^*)$ *and* $\overline{Ran(T)} = \overline{Ran(T^*)}$.

(c) *Let* $T \in \mathfrak{L}(\mathsf{H})$:

- (i) *if* T *is positive, its possible eigenvalues are real and non-negative;*
- (ii) *if* T *is bounded and self-adjoint, its possible eigenvalues are real;*
- (iii) *if* T *is isometric (in particular unitary), its possible eigenvalues are complex numbers of norm one.*

(d) *If* $T \in \mathfrak{L}(\mathsf{H})$ *satisfies* $(y|Tx) = (Ty|x)$ *for any pair* $x, y \in \mathsf{H}$, T *is bounded and self-adjoint.*
(e) *If* $T \in \mathfrak{B}(\mathsf{H})$ *satisfies* $(x|Tx) = (Tx|x)$ *for any* $x \in \mathsf{H}$, T *is self-adjoint.*
(f) *If* $T \in \mathfrak{B}(\mathsf{H})$ *is positive, it is self-adjoint.*
(g) *The relation* $\geq$ *is a partial order on* $\mathfrak{L}(\mathsf{H})$ *(hence on* $\mathfrak{B}(\mathsf{H})$*).*

Proof. (a) Set $Q := \sup\{|(x|Tx)| \mid x \in \mathsf{H}\,,\ ||x|| = 1\}$. Since we take $||x|| = 1$

$$|(x|Tx)| \leq ||Tx||||x|| \leq ||Tx|| \leq ||T||\,,$$

hence $Q \leq ||T||$. To conclude it suffices to show $||T|| \leq Q$. The immediate identity

$$4(x|Ty) = (x+y|T(x+y)) - (x-y|T(x-y)) - i(x+iy|T(x+iy)) \\ + i(x-iy|T(x-iy))\,,$$

together with the fact that $\overline{(z|Tz)} = (Tz|z) = (z|Tz)$, allow to rephrase $4Re(x|Ty) = 2(x|Ty) + 2\overline{(x|Ty)}$ as:

$$4Re(x|Ty) = (x+y|T(x+y)) - (x-y|T(x-y)) \leq Q||x+y||^2 + Q||x-y||^2$$
$$= 2Q||x||^2 + 2Q||y||^2 .$$

We proved:

$$4Re(x|Ty) \leq 2Q||x||^2 + 2Q||y||^2 .$$

Let $y \in \mathsf{H}$, $||y|| = 1$. If $Ty = 0$, it is clear that $||Ty|| \leq Q$; otherwise, define $x := Ty/||Ty||$ and we obtain the above inequality:

$$4||Ty|| = 4Re(x|Ty) \leq 2Q(||x||^2 + ||y||^2) = 2Q(1+1) = 4Q,$$

from which $||Ty|| \leq Q$ once again. Overall, $||Ty|| \leq Q$ if $||y|| = 1$, so

$$||T|| = \sup\{||Ty|| \mid y \in \mathsf{H},\ ||y|| = 1\} \leq Q .$$

The more general statement follows from the second part of the above proof ($||T|| \leq Q$).
(b) (iii) The claim follows from the observation that $TT^* = T^*T$ implies $||Tx||^2 = (Tx|Tx) = (x|T^*Tx) = (x|TT^*x) = ||T^*x||^2$. The remaining identities are now obvious, in the light of Proposition 3.38(d). Let us prove (i). $T - \lambda I$ is normal with adjoint $T^* - \overline{\lambda} I$, so (iii) gives

$$||Tu - \lambda u|| = ||T^*u - \overline{\lambda} u||$$

and the claim is proved. (ii) Let u be an eigenvector of T with eigenvalue λ, v an eigenvector of T with eigenvalue μ. By (i), $\lambda(v|u) = (v|Tu) = (T^*v|u) = (\overline{\mu}v|u) = \mu(v|u)$, so $(\lambda - \mu)(v|u) = 0$. But $\lambda \neq \mu$, so $(v|u) = 0$.
(c) If $T \geq 0$ and $Tu = \lambda u$ with $u \neq 0$, then $0 \leq (u|Tu) = \lambda(u|u)$; since $(u|u) > 0$, $\lambda \geq 0$. Let now $T = T^*$ and $Tu = \lambda u$ with $u \neq 0$: then $\lambda(u|u) = (u|Tu) = (Tu|u) = \overline{\lambda}(u|u)$. From $(u|u) \neq 0$ we have $\lambda = \overline{\lambda}$, i.e. $\lambda \in \mathbb{R}$. If, instead, T is isometric: $(u|u) = (Tu|Tu) = |\lambda|^2(u|u)$, so $|\lambda| = 1$ as $u \neq 0$.
(d) It is enough to prove T is bounded. The adjoint's uniqueness implies that $T = T^*$ because $(y|Tx) = (Ty|x)$ for any pair $x, y \in \mathsf{H}$. By the closed graph Theorem 2.95, to prove T bounded we can just show it is closed. Let then $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{H}$ be a sequence converging to x and suppose the vectors Tx_n form a converging sequence: the claim is $Tx_n \to Tx$. Given $y \in \mathsf{H}$, with our assumptions we have

$$(y|Tx_n) = (Ty|x_n) \to (Ty|x) = (y|Tx) .$$

The inner product is continuous, and by hypothesis $\lim_{n\to+\infty} Tx_n$ exists, so

$$\left(y \,\middle|\, \lim_{n\to+\infty}(Tx - Tx_n)\right) = 0 .$$

Given that y is arbitrary, and choosing precisely $y := \lim_{n\to+\infty}(Tx - Tx_n)$ we obtain $\lim_{n\to+\infty}(Tx - Tx_n) = 0$.

(e)–(f) We have $((T^* - T)x|x) = 0$ for any $x \in \mathsf{H}$. By Exercise 3.18 $T^* - T = 0$ i.e. $T = T^*$. If $T \in \mathfrak{B}(\mathsf{H})$ is positive, then $(x|Tx)$ is real and coincides with its complex conjugate $(Tx|x)$ (by the properties of the inner product), so we fall back into the previous case.
(g) We have to prove three things. (i) $T \geq T$: this is obvious because it means $((T - T)x|x) \geq 0$ for any $x \in \mathsf{H}$. (ii) if $T \geq U$ and $U \geq S$ then $T \geq S$: this is immediate by noting $T - S = (T - U) + (U - S)$, so $((T - S)x|x) = ((T - U)x|x) + ((U - S)x|x) \geq 0$ for any $x \in \mathsf{H}$, since $T \geq U$ and $U \geq S$. (iii) if $T \geq U$ and $U \geq T$ then $T = U$. For this last one notice $(x|(T - U)x) = 0$ for any $x \in \mathsf{H}$. Exercise 3.18 forces $T - U = 0$, so $T = U$. □

Remarks 3.55. Observe that on real Hilbert spaces the relation $\geq$ is not a partial order, because $A \geq 0$ and $0 \geq A$ do not imply $A = 0$. For example consider a skew-symmetric matrix A acting on $\mathbb{R}^n$ (seen as real vector space with ordinary inner product). Then $A \geq 0$ and also $0 \geq A$, since $(x|Ax) = 0$ for any $x \in \mathbb{R}^n$, but A can be very different from the null matrix. ■

3.4 Orthogonal projectors and partial isometries

The last elementary notion we would like to introduce are *orthogonal projectors* and, related to this, *partial isometries*. Th former will play a role to construct the formalism of QM.

Referring to Definition 2.96 and the subsequent Propositions 4.5, 5.17, we provide this definition.

Definition 3.56. *If* $(\mathsf{H}, (\,|\,))$ *is a Hilbert space, a projector* $P \in \mathfrak{B}(\mathsf{H})$ *is called* **orthogonal projector** *if* $P^* = P$.

Remarks 3.57. With that in place, orthogonal projectors are precisely those bounded operators from H to H that are defined by $P = PP$ (P is idempotent) and $P = P^*$ (P is self-adjoint). An immediate consequence is the *positivity* of orthogonal projectors: for any $x \in \mathsf{H}$

$$(u|Pu) = (u|PPu) = (P^*u|Pu) = (Pu|Pu) = ||Pu||^2 \geq 0\,.$$ ■

The next couple of propositions characterise orthogonal projectors.

Proposition 3.58. *If* $P \in \mathfrak{B}(\mathsf{H})$ *is an orthogonal projector* (H *Hilbert) onto the space* M *the following hold.*

(a) $Q := I - P$ *is an orthogonal projector.*
(b) $Q(\mathsf{H}) = M^\perp$*, so the direct sum decomposition associated to* P *and* Q *as of Proposition 2.97(b) is given by* M *and its orthogonal* $M^\perp$*:*

$$\mathsf{H} = M \oplus M^\perp\,.$$

(c) *For any* $x \in \mathsf{H}$, $||x - P(x)|| = \min\{||x - y|| \mid y \in M\}$.
(d) *If N is a basis on M, then:*

$$P = s\text{-}\sum_{u \in N} u\,(u|\ )\,,$$

where the "s-" denotes a series computed in the strong topology in case the sum is infinite.
(e) $I \geq P$*; moreover,* $||P|| = 1$ *if P is not the null operator (the projector onto* $\{0\}$*).*

Proof. (a) We know already that $Q := I - P$ is a projector (Proposition 2.4). By part (c) in Proposition 3.38, since $I^* = I$, we have $Q^* = Q$, so Q is an orthogonal projector.
(b) By Proposition 2.97(b) it is enough to show $Q(\mathsf{H}) = M^\perp$. For that notice that if $x \in Q(\mathsf{H})$ and $y \in M$, then $(x|y) = (Qx|y) = (x|Qy) = (x|y - Py) = (x|y - y) = 0$, so $Q(\mathsf{H}) \subset M^\perp$. We claim $M^\perp \subset Q(\mathsf{H})$, hence $M^\perp = Q(\mathsf{H})$. By Proposition 2.97, we have a direct sum decomposition:

$$\mathsf{H} = M \oplus Q(\mathsf{H})\,.$$

At the same time Theorem 3.13(d) gives the (orthogonal) decomposition:

$$\mathsf{H} = M \oplus M^\perp\,.$$

If $y \in M^\perp$, the first decomposition induces $y = y_M + z$ with $y_M \in M$ and $z \in Q(\mathsf{H})$. As we saw, $Q(\mathsf{H}) \subset M^\perp$, so the uniqueness of the above splitting implies $y = y_M + z$ must also be the decomposition of y induced by $\mathsf{H} = M \oplus M^\perp$. Thus $y_M \in M$ and $z \in M^\perp$. Then by assumption $y_M = 0$, and $y = z \in Q(\mathsf{H})$. Since $y \in M^\perp$ is arbitrary, we have proved $M^\perp \subset Q(\mathsf{H})$.
(c) The statement is a straighforward consequence of Theorem 3.13(d) when $K := M$, because the decomposition is unique.
(d) We may extend N to a basis of H by adding a basis N' of $M^\perp$ (in fact $N \cup N'$ is an orthonormal system by construction; moreover, part (b) gives $\mathsf{H} = M \oplus M^\perp$, so any $x \in \mathsf{H}$ orthogonal to both N and N' must be null. Then, by definition, $N \cup N'$ is basis for H.) We can immediately verfy that:

$$R = \sum_{u \in N} u\,(u|\ )$$

and

$$R' = \sum_{u \in N'} u\,(u|\ )$$

(if the sums are infinite, we use the strong topology) are bounded operators, and they satisfy $RR = R$, $R(\mathsf{H}) = M$, $R'R' = R'$, $R'(\mathsf{H}) = M^\perp$ and also $R'R = RR' = 0$ and $R + R' = I$. By Proposition 2.97 R and R' are projectors associated to the decomposition $M \oplus M^\perp$. By uniqueness of the decomposition of any vector we must have $R = P$ (and $R' = Q$).
(e) $Q = I - P$ is an orthogonal projector such that:

$$0 \leq (Qx|Qx) = (x|QQx) = (x|Qx) = (x|Ix) - (x|Px)\,,$$

for any $x \in \mathsf{H}$. This means $I \geq P$. What we have just seen also implies

$$||Px||^2 = (Px|Px) = (x|PPx) = (x|Px) \leq (x|x) = ||x||^2 .$$

Therefore taking the supremum over all x with $||x|| = 1$ yields $||P|| \leq 1$. If $P \neq 0$, there will be a unit $x \in \mathsf{H}$ so that $Px = x$, hence $||Px|| = 1$. If so, $||P|| = 1$. □

Proposition 3.59. *Let* H *be a Hilbert space and* $\mathsf{M} \subset \mathsf{H}$ *a closed subspace. The projectors* P *and* Q *that yield the orthogonal decomposition* $\mathsf{H} = \mathsf{M} \oplus \mathsf{M}^\perp$*, as in Proposition 2.98 (with* $\mathsf{N} := \mathsf{M}^\perp$*), and project onto* M *and* $\mathsf{M}^\perp$ *respectively, are orthogonal.*

Proof. It is enough to prove $P = P^*$. That $Q = Q^*$ follows from $Q = I - P$.
If $x \in \mathsf{H}$ we have a unique decomposition $x = y + z$, $y = P(x) \in \mathsf{M}$ and $z = Q(x) \in \mathsf{M}^\perp$. Let $x' = y' + z'$ be the analogous splitting of $x' \in \mathsf{H}$. Then $(x'|Px) = (y' + z'|y) = (y'|y)$. We also have $(Px'|x) = (y'|y+z) = (y'|y)$, hence $(x'|Px) = (Px'|x)$ i.e. $((P^* - P)x'|x) = 0$ for any $x, x' \in \mathsf{H}$. By choosing $x = (P^* - P)x'$ we obtain $Px' = P^*x'$ for any x', so $P = P^*$. □

Our last result characterises commuting orthogonal projetions.

Proposition 3.60. *Two orthogonal projectors* P *and* P' *on a Hilbert space* H ***commute****,*

$$PP' = P'P,$$

if and only if there is a basis N *of* H *such that:*

$$P = s\text{-}\sum_{u \in N_P} u\,(u|\;) \quad \textit{and, simultaneously,} \quad P' = s\text{-}\sum_{u \in N_{P'}} u\,(u|\;),$$

for some pair of subsets $N_P, N_{P'} \subset N$.

Proof. If $P = s\text{-}\sum_{u \in N_P} u\,(u|\;)$ and $P' = s\text{-}\sum_{u \in N_{P'}} u\,(u|\;)$ for subsets $N_P, N_{P'} \subset N$, where N is a basis in H, then trivially $PP' = P'P$, as a direct computation, involving the orthogonality relations of $u \in N$, shows. Conversely, assume $PP' = P'P$; if $M := P(\mathsf{H})$, it is not hard to see $P'(M) \subset M$ and $P'(M^\perp) \subset M^\perp$. In addition $P'\restriction_M$ and $P'\restriction_{M^\perp}$ are orthogonal projectors in Hilbert spaces M and $M^\perp$ respectively, plus $P' = P'\restriction_M \oplus P'\restriction_{M^\perp}$ corresponding to the orthogonal splitting $\mathsf{H} = M \oplus M^\perp$. By Proposition 3.58(d) (as any orthonormal system can be completed to a basis) we can fix a basis N_M of M and a basis $N_{M^\perp}$ of $M^\perp$ such that, for suitable subsets $N'_M \subset N_M$, $N'_{M^\perp} \subset N_{M^\perp}$:

$$P'\restriction_M = s\text{-}\sum_{u \in N'_M} u(u|\;), \quad P'\restriction_{M^\perp} = s\text{-}\sum_{v \in N'_{M^\perp}} v(v|\;).$$

Therefore

$$P' = P'\restriction_M \oplus P'\restriction_{M^\perp} = s\text{-}\sum_{w \in N'_M \cup N'_{M^\perp}} w(w|\;).$$

By construction, from the orthogonal decomposition $\mathsf{H} = M \oplus M^\perp$ we have that $N_M \cup N_{M^\perp}$ is a basis of H, and

$$P = s\text{-}\sum_{w \in N_M} w(w|\ ),$$

again from Proposition 3.58(d). The basis $N := N_M \cup N_{M^\perp}$ of H satisfies the requirements once we set $N_P := N_M$ and $N_{P'} := N'_M \cup N'_{M^\perp}$. □

We can pass to the useful notion of a *partial isometry*, a weaker version of isometry seen earlier.

Definition 3.61. *A bounded operator $U : \mathsf{H} \to \mathsf{H}$, with H a Hilbert space, is a* **partial isometry** *when:*

$$||Ux|| = ||x||, \quad \textit{for } x \in [Ker(U)]^\perp.$$

Then $[Ker(U)]^\perp$ is called the **initial space** *of U and $Ran(U)$ the* **final space**.

Any unitary operator $U : \mathsf{H} \to \mathsf{H}$ is a special partial isometry whose initial and final spaces coincide with the entire Hilbert space H. Observe also that if $U : \mathsf{H} \to \mathsf{H}$ is a partially isometric operator then H decomposes orthogonally into $Ker(U) \oplus [Ker(U)]^\perp$, and U restricts to an honest isometry on the second summand (with values in $Ran(U)$), while it is null on the first summand. This self-evident fact can be made stronger by proving that $Ran(U)$ is closed, hence showing $U\restriction_{[Ker(U)]^\perp} : [Ker(U)]^\perp \to Ran(U)$ is indeed a unitary operator between Hilbert spaces (closed subspaces in H). The second statement in the ensuing proposition shows U^* is a partial isometry if U is, and its initial and final spaces are those of U, but swapped.

Proposition 3.62. *Let $U : \mathsf{H} \to \mathsf{H}$ be a partial isometry on the Hilbert space H. Then:*

(a) *$Ran(U)$ is closed.*
(b) *$U^* : \mathsf{H} \to \mathsf{H}$ is a partial isometry with initial space $Ran(U)$ and final space $[Ker(U)]^\perp$.*

Proof. (a) Let $y \in \overline{Ran(U)} \setminus \{0\}$. There is a sequence of vectors $x_n \in [Ker(U)]^\perp$ such that $Ux_n \to y$ as $n \to +\infty$. Since $||U(x_n - x_m)|| = ||x_n - x_m||$ by definition of partial isometry, the sequence x_n is Cauchy and converges to some $x \in \mathsf{H}$. By continuity $y = Ux$, so $y \in Ran(U)$. But $y = 0$ clearly belongs to $Ran(U)$, so we have proved $Ran(U)$ contains all its limits points, and as such it is closed.
(b) Begin by observing $Ker(U^*) = \overline{Ran(U)}$ and $\overline{Ran(U^*)} = [Ker(U)]^\perp$ by Proposition 3.38, so if we use part (a) there remains only to prove U^* is an isometry when restricted to $\overline{Ran(U)}$. Notice preliminarly that if $z, z' \in [Ker(U)]^\perp$, U being a partial isometry implies:

$$(Uz|Uz') = \frac{1}{4}\left[||U(z+z')||^2 - ||U(z-z')||^2 - i||U(z+iz')||^2 + i||U(z-iz')||^2)\right]$$

$$= \frac{1}{4}\left[||z+z'||^2 - ||z-z'||^2 - i||z+iz'||^2 + i||z-iz'||^2)\right] = (z|z')\,.$$

From what we have seen, suppose $y = Ux$ with $x \in [Ker(U)]^\perp$. Then

$$||U^* y||^2 = (U^*Ux|U^*Ux) = (Ux|U(U^*Ux)) = (x|U^*Ux) = (Ux|Ux) = ||y||^2 .$$

In other terms U^* is isometric on $Ran(U)$, and it is so also on the closure, by continuity. This proves (b). □

At last, we present a relationship between partial isometries and orthogonal projectors.

Proposition 3.63. *Let $U : \mathsf{H} \to \mathsf{H}$ be a bounded linear operator on the Hilbert space H.*

(a) *U is a partial isometry if and only if U^*U is an orthogonal projector. In such a case UU^* is an orthogonal projector as well.*
(b) *If U is a partial isometry, U^*U projects onto the initial space of U, and UU^* projects onto the final space of U.*

Proof. Suppose U is partially isometric, and let us show U^*U is an orthogonal projector. Since the latter is patently self-adjoint, it suffices to show it is idempotent. Decompose $\mathsf{H} \ni x = x_1 + x_2$ by $x_1 \in Ker(U)$ and $x_2 \in [Ker(U)]^\perp$. Then

$$(x|(U^*U)^2 x) = (Ux|U^*UUx) = (Ux_2|U(U^*Ux_2)) = (x_2|U^*Ux_2) = (Ux_2|Ux_2)$$

$$= (Ux|Ux) .$$

I.e., $(x|((U^*U)^2 - U^*U)x) = 0$ whichever $x \in \mathsf{H}$. Choose $x = y \pm iz$ and $x = x \pm y$, and then $(y|((U^*U)^2 - U^*U)z) = 0$ for any $y, z \in \mathsf{H}$; therefore U^*U is idempotent, so an orthogonal projector. Conversely if U^*U is an orthogonal projector, let N be the closed subspace onto which it projects. If $U^*Ux = 0$ then $Ux \in Ker(U^*) = [Ran(U)]^\perp$. But $Ux \in Ran(U)$, hence $Ux = 0$. Therefore $U^*Ux = 0$ if and only if $x \in Ker(U)$, so $\mathsf{N}^\perp = Ker(U^*U) = Ker(U)$ and $\mathsf{N} = [Ker(U)]^\perp$. If additionally $x \in [Ker(U)]^\perp = \mathsf{N}$, then $U^*Ux = x$ and $||Ux||^2 = (U^*Ux|x) = ||x||^2$, proving U is a partial isometry. Throughout we also proved U^*U projects on the initial space $\mathsf{N} = [Ker(U)]^\perp$. The remaining part follows easily from Proposition 3.62(b): if U is a partial isometry, U^* is partially isometric and so $UU^* = (U^*)^*U^*$ is an orthogonal projector; from the previous part it projects onto the closed subspace $[Ker(U^*)]^\perp = \overline{Ran(U)}$. □

3.5 Square roots of positive operators and polar decomposition of bounded operators

This section is rather technical and contains notions useful in the theory of bounded operators on a Hilbert space. The central result is the so-called *polar decomposition theorem for bounded operators* that generalises the polar form of a complex number, whereby $z = |z|e^{i \arg z}$ splits as product of the modulus times an exponential with imaginary logarithm. In the analogy z corresponds to a bounded operator, $|z|$ plays the role of a certain positive operator called the *modulus* of the operator, and $e^{i\arg z}$ is represented by a unitary operator when restricted to a subspace. The *modulus of an operator* is useful to introduce a generalisation of "absolute convergence" for numerical series, built using operators and bases. We shall use these series to define *Hilbert–Schmidt operators* and *operators of trace class*, some of which represent states in QM. Part of the ensuing proofs are taken from [Mar82] and [KaAk80].

Definition 3.64. *Given $A \in \mathfrak{B}(\mathsf{H})$ with H a Hilbert space, $B \in \mathfrak{B}(\mathsf{H})$ is a* **square root** *of A if $B^2 = A$. If additionally $B \geq 0$, we call B a* **positive square root**.

We will show in a moment that any bounded positive operator has one, and one only, positive square root. For this we need the preliminary result below, about sequences and series of orthogonal projectors in the strong topology, which is on its own a useful fact in spectral theory.

Proposition 3.65. *Let H be a Hilbert space and $\{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{B}(\mathsf{H})$ a non-decreasing (resp. non-increasing) sequence of self-adjoint operators. If $\{A_n\}$ is bounded from above (below) by $K \in \mathfrak{B}(\mathsf{H})$, there exists $A \in \mathfrak{B}(\mathsf{H})$ self-adjoint with $A \leq K$ ($A \geq K$) and such that:*

$$A = s\text{-}\lim_{n\to+\infty} A_n \,. \tag{3.46}$$

Proof. We prove it in the non-decreasing case, for the other case falls back to this situation if one considers the sequence $K - A_n$. Set $B_n := A_n + ||A_0||I$. Then we have these facts.

(i) The B_n form a *non-decreasing sequence of positive operators*. If $||x|| = 1$ in fact, $(x|A_n x) + ||A_0|| \geq (x|A_0 x) + ||A_0||$, but $-||A_0|| \leq (x|A_0 x) \leq ||A_0||$ by Proposition 3.54(a). Therefore $(x|A_n x) + ||A_0|| \geq 0$ for any unit x. That is to say $(y|A_n y) + ||A_0||(y|y) \geq 0$ for any $y \in \mathsf{H}$, i.e. $B_n = A_n + ||A_0||I \geq 0$.
(ii) $B_n \leq K + ||A_0||I =: K_1$, and K_1 is *positive* (K cannot be).
(iii) $(x|K_1 x) \geq (x|B_n x) - (x|B_m x) \geq 0$ for any $x \in \mathsf{H}$ if $n \geq m$. In fact, $(x|K_1 x) \geq (x|B_n x)$, and also $-(x|B_m x) \leq 0$ and $(x|B_n x) - (x|B_m x) \geq 0$.
Since any positive operator T defines a semi-inner product that satisfies Schwarz's inequality:

$$|(x|Ty)|^2 \leq (x|Tx)(y|Ty)\,, \tag{3.47}$$

we have, if $n \geq m$:

$$|(x|(B_n - B_m)y)|^2 \leq (x|(B_n - B_m)x)(y|(B_n - B_m)y) \leq (x|K_1 x)(y|K_1 y) \leq ||K_1||^2||x||^2||y||^2 \,.$$

Hence

$$|(x|(B_n - B_m)y)|^2 \leq ||K_1||^2||x||^2||y||^2 \,.$$

If we set $x = (B_n - B_m)y$ and take the supremum over unit $y \in \mathsf{H}$, we find:

$$||B_n - B_m|| \leq ||K_1|| \,. \tag{3.48}$$

From (3.47), putting $y = (B_n - B_m)x$ and $T = B_n - B_m$, we obtain

$$||(B_n - B_m)x||^4 = ((B_n - B_m)x|(B_n - B_m)x)^2 \\ \leq (x|(B_n - B_m)x)((B_n - B_m)x|(B_n - B_m)^2 x)$$

for $x \in \mathsf{H}$, $n \geq m$. By (3.48), the last term is bounded by

$$(x|(B_n - B_m)x)||B_n - B_m||^3||x||^2 \leq ||K_1||^3||x||^2[(x|B_n x) - (x|B_m x)],$$

and so

$$||(B_n - B_m)x||^4 \leq ||K_1||^3||x||^2[(x|B_n x) - (x|B_m x)].$$

The non-decreasing, bounded sequence of positive numbers $(x|B_k x)$ has to converge, so it is Cauchy. Therefore also $B_k x$ must be a Cauchy sequence, and the limit exists as $k \to +\infty$. Define

$$B : \mathsf{H} \ni x \mapsto \lim_{n \to +\infty} B_n x.$$

By construction B is linear, and it satisfies

$$0 \leq (Bx|x) = (x|Bx) \leq (x|K_1 x)$$

since $0 \leq (B_k x|x) = (x|B_k x) \leq (x|K_1 x)$ for any $k \in \mathbb{N}$.

Now, K_1 is bounded and self-adjoint (it is positive), so Proposition 3.54(a) forces B to be bounded, since:

$$\sup\{|(x|Bx)| \mid x \in \mathsf{H}, ||x|| = 1\} \leq \sup\{|(x|K_1 x)| \mid x \in \mathsf{H}, ||x|| = 1\} = ||K_1||.$$

B is also self-adjoint because of Proposition 3.54(e). Therefore $A := B - ||A_0||I$ is a bounded, self-adjoint operator and

$$Ax = \lim_{n \to +\infty} (B_n - ||A_0||I)x = \lim_{n \to +\infty} A_n x.$$

Eventually, $A \leq K$ because for any $x \in \mathsf{H}$ we have $(x|A_n x) \leq (x|Kx)$ by assumption, and this is still true when taking the limit as $n \to +\infty$. □

The above result allows us to prove that bounded positive operators admit square roots.

Theorem 3.66. *Let* H *be a Hilbert space and* $A \in \mathfrak{B}(\mathsf{H})$ *a positive operator. Then there exists a unique positive square root, indicated by* $\sqrt{A}$*. Furthermore:*

(a) $\sqrt{A}$ *commutes with any* $B \in \mathfrak{B}(\mathsf{H})$ *that commutes with* A*:*

$$\text{if } AB = BA \text{ with } B \in \mathfrak{B}(\mathsf{H}), \text{ then } \sqrt{A}B = B\sqrt{A}.$$

(b) *if* A *is bijective,* $\sqrt{A}$ *is bijective.*

Proof. We may as well suppose $||A|| \leq 1$ without any loss of generality, so let us set $A_0 := I - A$. First of all let us show $A_0 \geq 0$ and $||A_0|| \leq 1$. $A_0 \geq 0$ because $(x|A_0x) = (x|x) - (x|Ax) \geq ||x||^2 - ||A||||x||^2$, where we have used $A = A^*$, so (Proposition 3.54(a)) $||A|| = \sup\{|(z|Az)| \mid ||z|| = 1\}$ and recalling $|(z|Az)| = (z|Az)$ by positivity. Since $(x,y) \mapsto (x|A_0y)$ is a semi-inner product, as $A_0 \geq 0$, the Cauchy-Schwarz inequality:

$$|(x|A_0y)|^2 \leq (x|A_0x)(y|A_0y) \leq ||x||^2||y||^2$$

holds, having used the positivity of $A = I - A_0$ and A_0 in the final step. Since $A = A^*$, using $y = A_0x$ in the inequality gives

$$|(A_0x|A_0x)|^2 \leq ||x||^2||A_0x||^2\,,$$

hence $||A_0x|| \leq ||x||$, and so:

$$||A_0|| \leq 1\,. \tag{3.49}$$

Let us now define a sequence of bounded operators $B_n : \mathsf{H} \to \mathsf{H}$, $n = 1, 2, \ldots$:

$$B_1 := 0\,, \quad B_{n+1} := \frac{1}{2}(A_0 + B_n^2)\,. \tag{3.50}$$

From (3.49), using the norm's properties,

$$||B_n|| \leq 1 \quad \text{for any } n \in \mathbb{N}\,. \tag{3.51}$$

By induction the operators B_n are polynomials in A_0 with non-negative coefficients. *Recall, here and in the sequel, that all operators B_k commute with one another and with A_0 by construction.* Equation (3.50) implies

$$B_{n+1} - B_n = \frac{1}{2}(A_0 + B_n^2) - \frac{1}{2}(A_0 + B_{n-1}^2) = \frac{1}{2}(B_n^2 - B_{n-1}^2)$$

i.e.

$$B_{n+1} - B_n = \frac{1}{2}(B_n + B_{n-1})(B_n - B_{n-1})\,.$$

This identity implies, by induction, that $B_{n+1} - B_n$ are polynomials in A_0 with non-negative coefficients: every $B_n + B_{n-1}$ is the sum of polynomials with non-negative coefficients, and is itself a polynomial with non-negative coefficients; moreover, the product of two such is of the same kind.

Since $A_0 \geq 0$, any polynomial in A_0 with non-negative coefficients is a positive operator: the polynomial is the sum of terms $a_{2n}A_0^{2n}$ (all positive, as $a_{2n} \geq 0$ and $A_0^{2n} = A_0^nA_0^n$ with A_0^n self-adjoint, so $a_{2n}(x|A_0^{2n}x) = a_{2n}(A_0^nx|A_0^nx) \geq 0$), and of positive terms $a_{2n+1}A_0^{2n+1}$ ($a_{2n+1} \geq 0$ and $(x|A_0^{2n+1}x) = (x|A_0^nAA_0^nx) = (A_0^nx|A\,A_0^nx) \geq 0$).

We conclude the bounded operators B_n and $B_{n+1} - B_n$ are positive. So the sequence of positive, bounded (and self-adjoint) operators B_n is non-decreasing. The sequence is also bounded from above by I. In fact, $B_n^* = B_n \geq 0$ implies, due to Proposition 3.54(a), that $(x|B_nx) = |(x|B_nx)| \leq ||B_n||||x||^2$. From (3.51) follows $B_n \leq I$.

So, we may apply Proposition 3.65 to detect a positive bounded operator $B_0 \leq I$ such that

$$B_0 = s - \lim_{n\to+\infty} B_n\,.$$

By definition of strong topology, and because the continuous B_k commute:

$$B_0B_mx = (\lim_{n\to+\infty} B_n)B_mx = \lim_{n\to+\infty} B_nB_mx = \lim_{n\to+\infty} B_mB_nx = B_m \lim_{n\to+\infty} B_nx = B_mB_0x\,.$$

Thus B_0 commutes with every B_m,

$$B_0^2 - B_n^2 = (B_0 + B_n)(B_0 - B_n)$$

and so, as $n \to +\infty$:

$$||B_0^2x - B_n^2x|| \leq ||B_0 + B_n||||B_0x - B_nx|| \leq (||B_0|| + ||B_n||)||B_0x - B_nx|| \\ \leq 2||B_0x - B_nx|| \to 0\,.$$

Rephrasing,

$$B_0^2x = \lim_{n\to+\infty} B_n^2x\,.$$

Taking the limit in

$$B_{n+1}x = \frac{1}{2}(A_0x + B_n^2x)\,,$$

obtained from (3.50), we find

$$B_0x = \frac{1}{2}(A_0x + B_0^2x)\,,$$

for any $x \in \mathsf{H}$, i.e.

$$2B_0 = A_0 + B_0^2\,.$$

To conclude, let us write the above identity in terms of $B := I - B_0$:

$$B^2 = I - A_0\,,$$

i.e.

$$B^2 = A\,.$$

Thus B is a square root of A. Note that $B \geq 0$ because $B_0 \leq I$ and $B = I - B_0$, so B is a positive root di A. Moreover, if C is bounded and commutes with A, it commutes with A_0 and hence with any B_n. Therefore C commutes also with B_0 and $B = I - B_0$. Let us see to uniqueness of a positive square root V of A. The above positive root B commutes with all operators that commute with A. Since

$$AV = V^3 = VA\,,$$

V and A commute, forcing B to commute with V. Fix an arbitrary $x \in \mathsf{H}$ and set $y := Bx - Vx$. Then:

$$||Bx - Vx||^2 = ([B-V]x|[B-V]x) = ([B-V]x|y) = (x|[B^* - V^*]y) = (x|[B-V]y). \tag{3.52}$$

We will show that $By = 0$ and $Vy = 0$ independently. This will end the proof, because then $||Bx - Vx|| = 0$ will imply $B = V$.

Now,

$$(y|By) + (y|Vy) = (y|[B+V][B-V]x) = (y|[B^2 - V^2]x) = (y|[A-A]x) = 0\,.$$

Since $(y|Vy) \geq 0$ and $(y|By) \geq 0$,

$$(y|Vy) = (y|By) = 0\,.$$

This means $Vy = By = 0$, for if W is a positive root of V, from

$$||Wy||^2 = (Wy|Wy) = (y|W^2y) = (y|Vy) = 0$$

it follows $Wy = 0$ and *a fortiori* $Vy = W(Wy) = 0$. In the same way $By = 0$.

There remains to prove $\sqrt{A}$ is bijective if A is. If A is bijective, it commutes with A^{-1}, so $\sqrt{A}$ as well commutes with A^{-1}. Then, immediately, $A^{-1}\sqrt{A} = \sqrt{A}A^{-1}$ is the left and right inverse of $\sqrt{A}$, which becomes bijective. □

Corollary 3.67. *Let* H *be a Hilbert space. If* $A, B \in \mathfrak{B}(\mathsf{H})$ *are positive and commute, their product is a positive bounded operator.*

Proof. $\sqrt{B}$ commutes with A, hence

$$(x|ABx) = (x|A\sqrt{B}^2x) = (x|\sqrt{B}A\sqrt{B}x) = (\sqrt{B}x|A\sqrt{B}x) \geq 0\,.$$ □

Remarks 3.68. That the square root of $0 \leq A \in \mathfrak{B}(\mathsf{H})$ commutes with every operator of $\mathfrak{B}(\mathsf{H})$ that commutes with A can be expressed, equivalently, by saying $\sqrt{A}$ belongs to the *von Neumann algebra* spanned by I and A in $\mathfrak{B}(\mathsf{H})$. ■

To conclude the section we will show that any bounded operator A in a Hilbert space admits a decomposition $A = UP$ as product of a uniquely-determined positive operator P with an isometric operator U, defined and unique on the image of P. The splitting is called *polar decomposition* and has a host of applications in mathematical physics. A preparatory definition is needed first.

Definition 3.69. *Let* H *be a Hilbert space and* $A \in \mathfrak{B}(\mathsf{H})$. *The bounded, positive and hence self-adjoint operator*

$$|A| := \sqrt{A^*A} \tag{3.53}$$

is called **modulus** *of* A.

Remarks 3.70. For any $x \in \mathsf{H}$: $||\,|A|x||^2 = (x|\,|A|^2x) = (x|A^*Ax) = ||Ax||^2$, so:

$$||\,|A|x|| = ||Ax||\,, \tag{3.54}$$

whence:

$$Ker(|A|) = Ker(A) \tag{3.55}$$

and so $|A|$ is injective if and only if A is. Another useful property is

$$\overline{Ran(|A|)} = (Ker(A))^\perp\,, \tag{3.56}$$

consequence of $\overline{Ran(|A|)} = ((Ran(|A|))^\perp)^\perp = (Ker(|A|^*))^\perp = (Ker(|A|))^\perp = (Ker(A))^\perp$. ■

Now to the polar decomposition theorem. We present the version for bounded operators. A more general statement will be discussed in Theorem 10.38 concerning a special class of unbounded operators.

Theorem 3.71 (Polar decomposition of bounded operators). *Let* H *be a Hilbert space and* $A \in \mathfrak{B}(\mathsf{H})$.

(a) *There exists a unique pair of operators* $P, U \in \mathfrak{B}(\mathsf{H})$ *such that:*

(i) *the decomposition*

$$A = UP \tag{3.57}$$

holds;

(ii) *P is positive;*

(iii) *U is isometric on $Ran(P)$;*

(iv) *U is null on $Ker(P)$.*

(b) $P = |A|$, *so* $Ker(U) = Ker(A) = Ker(P) = [Ran(P)]^{\perp}$.

(c) *If A is bijective, U coincides with the unitary operator* $A|A|^{-1}$.

Proof. (a) Start with uniqueness. If we have (3.57), $A = UP$ with $P \geq 0$ (beside bounded) and U bounded, then $A^* = PU^*$, since P is self-adjoint as positive (Theorem 3.66(c)); hence

$$A^*A = PU^*UP\,. \tag{3.58}$$

That U is isometric on $Ran(P)$ is written $(UPx|UPy) = (Px|Py)$ for any $x, y \in \mathsf{H}$, or $(x|[PU^*UP - P^2]y) = 0$ for any $x, y \in \mathsf{H}$. Therefore $PU^*UP = P^2$. Substituting in (3.58) we have $P^2 = A^*A$; as P is positive and extracting the only positive root (Theorem 3.66) on both sides we get $P = |A|$. So if a decomposition as claimed exists, necessarily $P = |A|$. Let us prove U is unique, as well. From $\mathsf{H} = Ker(P) \oplus (Ker(P))^{\perp}$, Proposition 3.38(d) and Theorem 3.13(e) imply $(Ker(P))^{\perp} = \overline{Ran(P^*)} = \overline{Ran(P)}$ because P is self-adjoint. Hence $\mathsf{H} = Ker(P) \oplus \overline{Ran(P)}$. To define an operator on H it suffices to have it on both summands above: $U = 0$ on $Ker(P)$, while $UPx = Ax$ for any $x \in \mathsf{H}$ fixes U on $Ran(P)$ uniquely. By assumption, on the other hand, U is bounded, and it remains bounded if restricted to $Ran(P)$. A bounded operator over a dense domain can be extended to a unique bounded operator on the closed domain (cf. Proposition 2.44). Therefore U is completely determined on all $\overline{Ran(P)}$, hence all H. This concludes the proof of uniqueness, so let us deal with existence.

We have seen it is necessary to show that $UP = A$, $P = |A|$, or better $U : |A|x \mapsto Ax$ for any $x \in \mathsf{H}$, actually defines an operator, say U_0, on $Ran(|A|)$. To make it well defined, it is necessary and sufficient that $|A|x = |A|y \Rightarrow Ax = Ay$, or else it would not be a function. By (3.54), if $|A|x = |A|y$, then $Ax = Ay$, so $U_0 : Ran(|A|) \ni |A|x \mapsto Ax$ is well defined (not multi-valued). That U_0 is linear is obvious by construction, and the same that it is an isometry, for U_0 preserves norms by (3.54) (cf. Exercise 3.5). Being an isometry on $Ran(|A|)$ implies, by continuity, that we can extend it uniquely to an isometry on the closure of $Ran(|A|)$, still denoted U_0. Now define $U : \mathsf{H} \to \mathsf{H}$ by setting $U\restriction_{Ker(|A|)} := 0$ and $U\restriction_{\overline{Ran(|A|)}} := U_0$, with respect to the splitting $\mathsf{H} = Ker(|A|) \oplus \overline{Ran(|A|)}$. It is immediate to see $U \in \mathfrak{B}(\mathsf{H})$ and that U satisfies (3.57). Furthermore, by construction $Ker(U) \supset Ker(|A|)$. We claim there latter

two are equal. Any u with $Uu = 0$ splits into $u_0 + x$, with $u_0 \in Ker(|A|)$ annihilated by U, and $x \in \overline{Ran(|A|)}$ for which $U_0 x = 0$. Since on $\overline{Ran(|A|)}$ U_0 is isometric, then $x = 0$ and so $u = u_0 \in Ker(|A|)$. Therefore $Ker(U) \subset Ker(|A|)$, so overall $Ker(U) = Ker(|A|) = Ker(A)$ by (3.55).
(b) is already proved within (a).
(c) If A is injective, using (b) we see $Ker(A) = Ker(U)$ is trivial and so U is injective. Directly from $A = UP$, though, we have $Ran(U) \supset Ran(A)$, so if A is onto also U is. Hence, if A is bijective U must be as well. If so, U is an onto isometry on $\overline{Ran(P)} = (Ker(P))^\perp = \{0\}^\perp = \mathsf{H}$ by (b), hence unitary. At last, from $A = U|A|$ follows that $|A|$ is bijective because A and U are, whence we can write $U = A|A|^{-1}$. □

Remarks 3.72. (1) Observe that the operator U showing up in (3.58) is a *partial isometry* (Definition 3.61) with *initial space*

$$[Ker(U)]^\perp = \overline{Ran(|A|)} = [Ker(A)]^\perp = \overline{Ran(A^*)}\,.$$

Bearing in mind Proposition 3.62(a) we see easily that the *final space* of U is

$$Ran(U) = \overline{Ran(A)}\,.$$

(2) With Theorem 10.38 we will prove the polar decomposition theorem under much weaker assumptions on A. We will also prove that the partial isometry U has the same initial and final spaces above, and is unitary precisely when A is injective and, simultaneously, $Ran(A)$ is dense in H. ■

Definition 3.73. *Let* H *be a Hilbert space and* $A \in \mathfrak{B}(\mathsf{H})$. *The splitting*

$$A = UP\,,$$

with $P \in \mathfrak{B}(\mathsf{H})$ *positive,* $U \in \mathfrak{B}(\mathsf{H})$ *isometric on* $Ran(P)$ *and null on* $Ker(P)$*, is called* **polar decomposition** *of the operator* A.

A corollary of the polar decomposition, useful in several applications, is this.

Corollary 3.74 (to Theorem 3.71). *Under the assumptions of Theorem 3.71, if* $U|A| = A$ *is the polar decomposition of* A*, then:*

$$|A^*| = U|A|U^*\,. \tag{3.59}$$

Proof. From $A = U|A|$ follows $A^* = |A|U^* = U^*U|A|U^*$, where we used $U^*U|A| = |A|$, since U is isometric on $Ran(|A|)$. Thus the self-adjoint operator AA^* satisfies

$$AA^* = U|A|U^*U|A|U^*\,.$$

As $U|A|U^*$ is clearly positive, by uniqueness of the root we have

$$|A^*| = \sqrt{(A^*)^*A^*} = \sqrt{AA^*} = U|A|U^*\,,$$

proving the claim. □

We cite, in the form of the next theorem, yet another consequence of the polar decomposition theorem valid when $A \in \mathfrak{B}(\mathsf{H})$ is normal, i.e. commuting with A^*.

Theorem 3.75 (Polar decomposition of normal operators). *Let $A \in \mathfrak{B}(\mathsf{H})$ be a normal operator on the Hilbert space H, and $W_0 : Ker(A) \to Ker(A)$ a given unitary operator. There exists a unique pair $W, P \in \mathfrak{B}(\mathsf{H})$ such that $P \geq 0$, W is unitary and:*

$$A = WP \quad with \quad W\restriction_{Ker(A)} = W_0 \,.$$

Moreover, $P = |A|$, $W\restriction_{Ker(A)^\perp}$ does not depend on W_0, and W commutes with A, A^* and $|A|$.

Proof. Under the assumptions made, $A = WP$ implies $A^*A = PW^*WP = P^2$, so $P = |P|$. Then consider the polar decomposition $A = U|A|$. As we know (Remark 3.72(1)) U is partially isometric, with initial space $Ker(A)^\perp$ and final space $\overline{Ran(A)}$. We have $\mathsf{H} = Ker(A) \oplus Ker(A)^\perp = Ker(A) \oplus \overline{Ran(A^*)}$, and since A is normal, by (iii) in Proposition 3.54(b) we can write $\mathsf{H} = Ker(A) \oplus \overline{Ran(A)}$. So U is unitary from $\overline{Ran(A)}$ to $\overline{Ran(A)}$, and is null from $Ker(A)$ to itself. Notice $Ker(|A|) = Ker(A)$, as seen earlier, so $\overline{Ran(|A|)} = \overline{Ran(A)}$. Now if there exists W unitary with $A = W|A|$, it must be isometric on $Ker(A)^\perp = \overline{Ran(A)}$, so it must coincide with U there by the polar decomposition theorem; thus the restriction of W to $\overline{Ran(A)}$ gives a unitary operator from $\overline{Ran(A)}$ to $\overline{Ran(A)}$. The condition $W\restriction_{Ker(A)} = W_0$ fixes W completely on the whole Hilbert space as unitary operator, ending the proof of uniqueness. As far as existence is concerned, it is enough to verify that $W := W_0 \oplus U$, defined according to $\mathsf{H} = Ker(A) \oplus \overline{Ran(A)}$, is an operator that commutes with A, A^*, and that $|A|$ fulfills $A = W|A|$. The latter request is true by the polar decomposition theorem, for $Ker(|A|) = Ker(A)$. Since $A^*(Ker(A^*)) \subset Ker(A^*)$, we have $A(Ker(A^*)^\perp) \subset Ker(A^*)^\perp$, i.e. $A(\overline{Ran(A)}) \subset \overline{Ran(A)}$. With respect to the usual orthogonal splitting of the Hilbert space, $A = 0 \oplus A\restriction_{\overline{Ran(A)}}$. Since we have $W = W_0 \oplus U\restriction_{\overline{Ran(U)}}$ and $A = 0 \oplus A\restriction_{\overline{Ran(A)}}$, the condition $AW = WA$ holds if $AU = UA$. So let us prove the latter. By the polar decomposition theorem $A = U|A|$. Normality ($AA^* = A^*A$) can be rephrased as $U|A|^2U^* = |A|U^*U|A| = |A|^2$, since U is isometric on $Ran(|A|)$. Therefore $U^*U|A|^2U^* = U^*|A|^2$ i.e. $|A|^2U^* = U^*|A|^2$. Taking adjoints gives $U|A|^2 = |A|^2U$. Theorem 3.66(a) says the root of an operator commutes with everything that commutes with the given operator, so $U|A| = |A|U$. Still by polar decomposition we infer $UA = AU$, so $WA = AW$, as required. Taking adjoints: $A^*W^* = W^*A^*$, and using W on both sides produces $WA^* = A^*W$. Consequently W commutes with $A^*A = |A|^2$, and hence also with its square root $|A|$. □

3.6 The Fourier-Plancherel transform

In the last section of the chapter we introduce, rather concisely, the basics in the theory of the Fourier and Fourier-Plancherel transforms, without any mention to Schwartz distributions [Rud91, ReSi80, Vla81].

Notation 3.76. From now on we will use the notations of Example 2.87 introduced for differential operators: in particular $x_k \in \mathbb{R}$ will denote the kth component of $x \in \mathbb{R}^n$, dx the ordinary Lebesgue measure on $\mathbb{R}^n$, and

$$M^\alpha(x) := x_1^{\alpha_1} \cdots x_n^{\alpha_n} \quad \text{for any multi-index } \alpha = (\alpha_1, \ldots, \alpha_n)\,.$$

By $\mathscr{D}(\mathbb{R}^n)$ we shall denote the space of smooth complex-valued functions with compact support (in the literature this is also called $C_c^\infty(\mathbb{R}^n)$ or $C_0^\infty(\mathbb{R}^n)$). $\mathscr{S}(\mathbb{R}^n)$ will indicate the **Schwartz space** on $\mathbb{R}^n$ (cf. Example 2.87). In these notations $\mathscr{S}(\mathbb{R}^n)$ is the $\mathbb{C}$-vector space of complex maps $C^\infty(\mathbb{R}^n)$ with this property: for any $f \in \mathscr{S}(\mathbb{R}^n)$ and for any pair of multi-indices α, β, there exists $K < +\infty$ (depending on f, α, β) such that

$$|M^\alpha(x)\partial_x^\beta f(x)| \leq K, \quad \text{for any } x \in \mathbb{R}^n\,. \tag{3.60}$$

The norms $|| \; ||_1$, $|| \; ||_2$ and $|| \; ||_\infty$ will denote, throughout the section, the norms of $L^1(\mathbb{R}^n, dx)$, $L^2(\mathbb{R}^n, dx)$ and $L^\infty(\mathbb{R}^n, dx)$ and the corresponding seminorms of $\mathscr{L}^1(\mathbb{R}^n, dx)$, $\mathscr{L}^2(\mathbb{R}^n, dx)$ and $\mathscr{L}^\infty(\mathbb{R}^n, dx)$ (see Examples 2.26(6) and (8)). ■

Below we recall a number of known properties.

(1) The spaces $\mathscr{D}(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$ are invariant under $M^\alpha(x)$ (seen as multiplicative operator) and ∂_x^α. Put otherwise, functions stay in their respective spaces when acted upon by $M^\alpha(x)$ and ∂_x^α.
(2) Clearly $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{L}^p(\mathbb{R}, dx)$ as subspace, for any $1 \leq p \leq \infty$, since compact sets in $\mathbb{R}^n$ have finite Lebesgue measure and any $f \in \mathscr{D}(\mathbb{R}^n)$ is continuous, hence bounded on compact sets.
(3) For any $1 \leq p \leq \infty$ we have $\mathscr{S}(\mathbb{R}^n) \subset \mathscr{L}^p(\mathbb{R}, dx)$ as subspace. In fact, if $C \subset \mathbb{R}^n$ is a compact set containing the origin, $f \in \mathscr{S}(\mathbb{R}^n)$ is bounded on C because continuous, while outside C we have $|f(x)| \leq C_m |x|^{-m}$ for any $n = 0, 1, 2, 3, \ldots$ as long as we choose $C_n \geq 0$ big enough. In summary, $|f|$ is bounded on $\mathbb{R}^n$, so it belongs to $\mathscr{L}^\infty$. But it is also bounded by a map in $\mathscr{L}^p$, for any $p \in [1, +\infty)$: this bounding function is constant on C, and equals $C_m/|x|^m$, $m > n/p$, outside C.
(4) Beside the obvious $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$, recall a notorious fact (independent of this section) that we will use shortly [KiGv82]:

Proposition 3.77. *The spaces $\mathscr{D}(\mathbb{R}^n)$, $\mathscr{S}(\mathbb{R}^n)$ are dense subspaces in $\mathscr{L}^p(\mathbb{R}, dx)$ for any $1 \leq p < \infty$.*

(5) The next important lemma, whose proof can be found in [Bre10, Corollary IV.24], is independent of the section's results.

Lemma 3.78. *Suppose $f \in \mathscr{L}^1(\mathbb{R}^n, dx)$ satisfies*

$$\int_{\mathbb{R}^n} f(x) g(x)\, dx = 0 \quad \textit{for any } g \in \mathscr{D}(\mathbb{R}^n)\,.$$

Then $f(x) = 0$ almost everywhere for the Lebesgue measure dx on $\mathbb{R}^n$.

Let us introduce the first elementary definitions concerning the Fourier transform.

Definition 3.79. *The linear maps*

$$(\mathscr{F}f)(k) := \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x)\,dx\,, \quad f \in \mathscr{L}^1(\mathbb{R}^n, dx),\ k \in \mathbb{R}^n\,, \tag{3.61}$$

$$(\mathscr{F}_- g)(x) := \int_{\mathbb{R}^n} \frac{e^{ik\cdot x}}{(2\pi)^{n/2}} g(k)\,dk\,, \quad g \in \mathscr{L}^1(\mathbb{R}^n, dk),\ x \in \mathbb{R}^n \tag{3.62}$$

from $\mathscr{L}^1(\mathbb{R}^n, dx)$ *to* $\mathscr{L}^\infty(\mathbb{R}^n, dx)$ *are respectively called* **Fourier transform** *and* **inverse Fourier transform**.

Remark 3.80. (1) Above dk always denotes the Lebesgue measure on $\mathbb{R}^n$. We have used a different name for the variable in $\mathbb{R}^n$ (k not x), in the inverse Fourier formula, only to respect the traditional notation and to simplify subsequent calculations.
(2) By the integral's properties it is obvious that

$$|(\mathscr{F}f)(k)| \leq \left| \int_{\mathbb{R}^n} e^{-ik\cdot x} f(x) \frac{dx}{(2\pi)^{n/2}} \right| \leq \int_{\mathbb{R}^n} |e^{-ik\cdot x}|\,|f(x)| \frac{dx^n}{(2\pi)^{n/2}}$$
$$= \int_{\mathbb{R}^n} |f(x)| \frac{dx}{(2\pi)^{n/2}} = \frac{||f||_1}{(2\pi)^{n/2}},$$

and similarly $|(\mathscr{F}_- g)(x)| \leq ||g||_1/(2\pi^{n/2})$ for any $x, k \in \mathbb{R}^n$; thus it makes sense to define the Fourier transform and inverse Fourier transform as operators with values in $\mathscr{L}^\infty(\mathbb{R}^n, dx)$. ■

In the sequel we will discuss straighforward features of the Fourier transform that are most related to the Fourier-Plancherel transform. We will overlook many results, like the continuity in the seminorm topology in the Schwartz space, for which we refer to any text on functional analysis or distributions [Rud91, ReSi80, Vla81] (see also Chapter 2.3.4).

Proposition 3.81. *The Fourier and inverse Fourier transforms enjoy the following properties.*

(a) *They are continuous in the natural norms of domain and codomain:*

$$||\mathscr{F}f||_\infty \leq \frac{||f||_1}{(2\pi)^{n/2}} \quad \text{and} \quad ||\mathscr{F}_- g||_\infty \leq \frac{||g||_1}{(2\pi)^{n/2}}.$$

(b) *The Schwartz space is invariant under* $\mathscr{F}$ *and* $\mathscr{F}_-$: $\mathscr{F}(\mathscr{S}(\mathbb{R}^n)) \subset \mathscr{S}(\mathbb{R}^n)$ *and* $\mathscr{F}_-(\mathscr{S}(\mathbb{R}^n)) \subset \mathscr{S}(\mathbb{R}^n)$.
(c) *When restricted to the invariant space* $\mathscr{S}(\mathbb{R}^n)$ *they are one the inverse of the other: if* $f \in \mathscr{S}(\mathbb{R}^n)$, *then*

$$g(k) = \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x)\,dx$$

if and only if

$$f(x) = \int_{\mathbb{R}^n} \frac{e^{ik\cdot x}}{(2\pi)^{n/2}} g(k)\, dk.$$

(d) *When restricted to the invariant space $\mathscr{S}(\mathbb{R}^n)$ they are isometric for the semi-inner product of $\mathscr{L}^2(\mathbb{R}^n, dx^n)$: if $f_1, f_2, g_1, g_2 \in \mathscr{S}(\mathbb{R}^n)$, then*

$$\int_{\mathbb{R}^n} \overline{(\mathscr{F} f_1)(k)} (\mathscr{F} f_2)(k) dk = \int_{\mathbb{R}^n} \overline{f_1(x)} f_2(x) dx$$

and

$$\int_{\mathbb{R}^n} \overline{(\mathscr{F}_- g_1)(x)} (\mathscr{F}_- g_2)(x) dx = \int_{\mathbb{R}^n} \overline{g_1(k)} g_2(k) dk\,.$$

(e) *They determine bounded maps from $L^1(\mathbb{R}^n, dx)$ to $C_0(\mathbb{R}^n)$ (continuous maps that vanish at infinity, cf. Example 2.26(4)), and so the* **Riemann-Lebesgue lemma** *holds: for any $f \in L^1(\mathbb{R}^n, dx)$*

$$(\mathscr{F} f)(k) \to 0 \quad as\ |k| \to +\infty$$

and analogously for $\mathscr{F}_-$.
(f) *They are injective if defined on $L^1(\mathbb{R}^n, dx)$.*

Remarks 3.82. Concering statement (f), more can be proved [Rud91], namely: if $f \in L^1(\mathbb{R}^n, dx)$ is such that $\mathscr{F} f \in L^1(\mathbb{R}^n, dk)$, then $\mathscr{F}_-(\mathscr{F} f) = f$. The same holds if we swap $\mathscr{F}$ and $\mathscr{F}_-$. ■

Proof of Proposition 3.81. (a) was proved in Remark 3.80(2). (b) Let us prove the claim about $\mathscr{F}$, the one about $\mathscr{F}_-$ being similar. Set

$$g(k) := \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x)\, dx.$$

The right-hand side can be differentiated in k by passing the operator ∂_k^α inside the integral. In fact,

$$\left|\partial_k^\alpha e^{-ik\cdot x} f(x)\right| = \left|i^{|\alpha|} M^\alpha(x) e^{-ik\cdot x} f(x)\right| \leq |M^\alpha(x) f(x)|\,.$$

The function $x \mapsto |M^\alpha(x) f(x)|$ is in $\mathscr{L}^1$ as $f \in \mathscr{S}(\mathbb{R}^n)$. Since the absolute value of the derivative of the integrand is uniformly bounded by an integrable positive map, known theorems on exchanging derivatives and integrals allow to say:

$$\partial_k^\alpha g(k) = (-i)^{|\alpha|} \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} M^\alpha(x) f(x)\, dx\,. \tag{3.63}$$

Since f vanishes faster than any inverse power of $|x|$, as $|x| \to +\infty$, we have:

$$M^\beta(k) g(k) = \int_{\mathbb{R}^n} i^{|\beta|} \partial_x^\beta \left(\frac{e^{-ik\cdot x}}{(2\pi)^{n/2}}\right) f(x)\, dx$$

and, integrating by parts,

$$M^{\beta}(k)g(k) = (-i)^{|\beta|}\int_{\mathbb{R}^n}\frac{e^{-ik\cdot x}}{(2\pi)^{n/2}}\partial_x^{\beta} f(x)\,dx\,. \tag{3.64}$$

Writing $\partial_k^{\alpha} g$ instead of g in (3.64), and by (a), we have:

$$|M^{\beta}(k)\partial_k^{\alpha} g(k)| \leq \left|\left|\partial^{\beta}\left(M^{\alpha} f\right)\right|\right|_1 ,$$

for any $k \in \mathbb{R}^n$. The right-hand term is finite, since $f \in \mathscr{S}(\mathbb{R}^n)$; and because α and β are arbitrary, we conclude $g \in \mathscr{S}(\mathbb{R}^n)$.
(c) Identities (3.63) and (3.64) read:

$$\partial^{\alpha}\,\mathscr{F} = (-i)^{|\alpha|}\,\mathscr{F} M^{\alpha}\,, \tag{3.65}$$

$$M^{\beta}\,\mathscr{F} = (-i)^{|\beta|}\,\mathscr{F}\partial^{\beta}\,, \tag{3.66}$$

where $\mathscr{F}$ is the restriction of the Fourier transform to $\mathscr{S}(\mathbb{R}^n)$. Observing that

$$\overline{\mathscr{F}h} = \mathscr{F}_-\overline{h}$$

for any $h \in \mathscr{S}(\mathbb{R}^n)$, it is easy to obtain

$$\partial^{\alpha}\,\mathscr{F}_- = i^{|\alpha|}\,\mathscr{F}_- M^{\alpha}\,, \tag{3.67}$$

$$M^{\beta}\,\mathscr{F}_- = i^{|\beta|}\,\mathscr{F}_-\partial^{\beta}\,. \tag{3.68}$$

Then (3.65), (3.66), (3.67) and (3.68) imply in particular that:

$$\mathscr{F}\mathscr{F}_- M^{\alpha} = M^{\alpha}\mathscr{F}\mathscr{F}_-\,, \tag{3.69}$$

$$\mathscr{F}_-\mathscr{F} M^{\alpha} = M^{\alpha}\mathscr{F}_-\mathscr{F}\,, \tag{3.70}$$

where M^{α} is thought of as *multiplicative operator* $(M^{\alpha} f)(x) := M^{\alpha}(x)f(x)$, and

$$\mathscr{F}\mathscr{F}_-\partial^{\alpha} = \partial^{\alpha}\mathscr{F}\mathscr{F}_-\,, \tag{3.71}$$

$$\mathscr{F}_-\mathscr{F}\partial^{\alpha} = \partial^{\alpha}\mathscr{F}_-\mathscr{F}\,. \tag{3.72}$$

By virtue of those commuting relations, we claim $J := \mathscr{F}\mathscr{F}_-$ and $J_- := \mathscr{F}_-\mathscr{F}$ are the identity of $\mathscr{S}(\mathbb{R}^n)$. To begin with, we show, given $x_0 \in \mathbb{R}^n$ and $f \in \mathscr{S}(\mathbb{R}^n)$, that the value $(Jf)(x_0)$ depends only on $f(x_0)$. If $f \in \mathscr{S}(\mathbb{R}^n)$ we can write:

$$f(x) = f(x_0) + \int_0^1 \frac{df(x_0+t(x-x_0))}{dt}dt = f(x_0) + \sum_{i=1}^{n}(x_i - x_{0i})g_i(x)\,,$$

where the g_i (in $C^{\infty}(\mathbb{R}^n)$, as is easy to see) are:

$$g_i(x) := \frac{\partial}{\partial x_i}\int_0^1 f(x_0+t(x-x_0))dt.$$

Hence if $f_1, f_2 \in \mathscr{S}(\mathbb{R}^n)$ and $f_1(x_0) = f_2(x_0)$:

$$f_1(x) = f_2(x) + \sum_{i=1}^{n} (x_i - x_{0i}) h_i(x) , \tag{3.73}$$

where, subtracting, the map $x \mapsto \sum_{i=1}^{n} (x_i - x_{0i}) h_i(x)$ and also the h_i belong to $\mathscr{S}(\mathbb{R}^n)$. Using J on both sides of (3.73) and recalling J commutes with polynomials in x by (3.69), we have:

$$(Jf_1)(x) = (Jf_2)(x) + \sum_{i=1}^{n} (x_i - x_{0i})(Jh_i)(x) .$$

Taking $x = x_0$ shows $(Jf_1)(x_0) = (Jf_2)(x_0)$ under the initial assumption $f_1(x_0) = f_2(x_0)$. Hence, as claimed, $(Jf)(x_0)$ is a map of $f(x_0)$ only. This map must be linear, as J is linear by construction. Consequently $(Jf)(x_0) = j(x_0) f(x_0)$, where j is a function on $\mathbb{R}^n$ with values in $\mathbb{C}$. Given that x_0 was arbitrary, J acts as multiplication by a function j. The latter must be C^∞. To justify this, choose $f \in \mathscr{S}(\mathbb{R}^n)$ equal to 1 on a neighbourhood $I(x_0)$ of x_0. If $x \in I(x_0)$, then $(Jf)(x) = j(x)$. The left-hand side is $C^\infty(I(x_0))$, so also the right term. That being valid around any point in $\mathbb{R}^n$, we have $j \in C^\infty(\mathbb{R}^n)$. Equation (3.71) implies

$$j(x) \frac{\partial}{\partial x^i} f(x) = \frac{\partial}{\partial x^i} j(x) f(x)$$

for any $f \in \mathscr{S}(\mathbb{R}^n)$, $x \in \mathbb{R}^n$. Choose as before f equal 1 on an open set, so the above identity forces all derivatives of j to vanish there. This holds around any point, and $\mathbb{R}^n$ is connected, so the continuous map j is constant on $\mathbb{R}^n$. The constant value clearly does not depend on the argument of J, and may be computed by evaluating J on an arbitrary function $\mathscr{S}(\mathbb{R}^n)$. Computing J on $x \mapsto e^{-x^2}$ is a useful exercise, and reveals the constant value is exactly 1. The argument for J_- is similar.
(d) Using (c) the claim is immediate. Let us carry out the proof for $\mathscr{F}$; the one for $\mathscr{F}_-$ is the same, essentially. Let $f_1, f_2 \in \mathscr{S}(\mathbb{R}^n)$ and set, $i = 1, 2$:

$$g_i(k) := \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f_i(x)\, dx .$$

With the assumptions made the theorem of Fubini–Tonelli gives:

$$\int_{\mathbb{R}^n} \overline{g_1(k)} g_2(k) dk = \int_{\mathbb{R}^n} \overline{g_1(k)} \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f_2(x) dx dk$$

$$= \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} \overline{g_1(k)} f_2(x) dx \otimes dk .$$

Now we rephrase the last integral and apply Fubini–Tonelli again:

$$\int_{\mathbb{R}^n} \overline{g_1(k)} g_2(k) dk$$

$$= \int_{\mathbb{R}^n\times\mathbb{R}^n} \overline{\frac{e^{ik\cdot x}}{(2\pi)^{n/2}} g_1(k)} f_2(x) dx \otimes dk = \int_{\mathbb{R}^n} f_2(x) \int_{\mathbb{R}^n} \overline{\frac{e^{ik\cdot x}}{(2\pi)^{n/2}} g_1(k)} dk dx$$

$$= \int_{\mathbb{R}^n} \overline{f_1(x)} f_2(x) dx ,$$

where part (c) was used. This was what we wanted.
(e) We prove for $\mathscr{F}$, and leave the similar assertion on $\mathscr{F}_-$ for the reader. Notice that both transformations are well defined on $L^1(\mathbb{R}^n, dx)$ since the integral is invariant by altering the maps by sets of zero Lebesgue measure. The estimate $||\mathscr{F}f||_\infty \leq \frac{||f||_1}{(2\pi)^{n/2}}$ guarantees the linear map $\mathscr{F} : \mathscr{S}(\mathbb{R}^n) \to \mathscr{S}(\mathbb{R}^n) \subset C_0(\mathbb{R}^n)$ is continuous when the domain has the L^1 norm and the codomain $||\cdot||_\infty$. $\mathscr{S}(\mathbb{R}^n)$ is dense in L^1 in the given norm, and the codomain is complete in the second norm; therefore the Fourier transform, initially defined on $\mathscr{S}(\mathbb{R}^n)$, can be extended by continuity and in a unique way to a bounded linear map from $L^1(\mathbb{R}^n, dx)$ to $C_0(\mathbb{R}^n)$ that preserves the same norm by Proposition 2.44 (and coincides with the aforementioned linear transformation on $L^1(\mathbb{R}^n, dx)$). If $f \in L^1$, $\mathscr{F}f \in C_0(\mathbb{R}^n)$, then for any $\varepsilon > 0$ there is a compact set $K_\varepsilon \subset \mathbb{R}^n$ such that $|(\mathscr{F}f)(k)| < \varepsilon$ if $k \notin K_\varepsilon$. Choose, for any $\varepsilon > 0$, a ball of radius r_ε at the origin big enough to contain K_ε: then there exists, for any $\varepsilon > 0$, a real number $r_\varepsilon > 0$ such that $|(\mathscr{F}f)(k)| < \varepsilon$ if $|k| > r_\varepsilon$.
(f) We prove the claim for $\mathscr{F}$, as the one for $\mathscr{F}_-$ is analogous. Since $\mathscr{F}$ is a linear operator, it suffices to show that if $\mathscr{F}f$ is the zero map, f is null almost everywhere. Thus assume:

$$\int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x)\, dx = 0 , \quad \text{for any } k \in \mathbb{R}^n.$$

If $g \in \mathscr{S}(\mathbb{R}^n)$, Fubini–Tonelli gives

$$0 = \int_{\mathbb{R}^n} g(k) \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x)\, dx\, dk = \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} g(k)\, dk \right) f(x)\, dx .$$

Since $\mathscr{F}$ is bijective on $\mathscr{S}(\mathbb{R}^n)$, what we have proved is equivalent (note $\psi f \in \mathscr{L}^1(\mathbb{R}^n, dx)$ for any $\psi \in \mathscr{S}(\mathbb{R}^n)$, as ψ is bounded) to:

$$\int_{\mathbb{R}^n} \psi(x) f(x)\, dx = 0 \quad \text{for any } \psi \in \mathscr{S}(\mathbb{R}^n).$$

As $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$, Lemma 3.78 forces f to vanish almost everywhere. □

We move on to the *Fourier-Plancherel transform*. As $\mathscr{S}(\mathbb{R}^n)$ is dense in $\mathscr{L}^2(\mathbb{R}^n)$, considering equivalence classes we can say $\mathscr{S}(\mathbb{R}^n)$ determines a dense subset, still indicated by $\mathscr{S}(\mathbb{R}^n)$, in the Hilbert space $L^2(\mathbb{R}^n)$. The operators $\mathscr{F}$ and $\mathscr{F}_-$ can be seen as defined on that dense subspace of $L^2(\mathbb{R}, dx)$. Proposition 3.81(d) says in particular that these operators are bounded with norm 1, since they are isometric. As Proposition 2.44 tells us, $\mathscr{F}$ and $\mathscr{F}_-$ determine unique bounded linear operators on $L^2(\mathbb{R}^n, dx)$. For instance, the operator extending $\mathscr{F}$ to $L^2(\mathbb{R}^n, dx)$ is defined as

$$\widehat{\mathscr{F}} f := \lim_{n\to+\infty} \mathscr{F} f_n ,$$

for $f \in L^2(\mathbb{R}^n, dx)$. Above, $\{f_n\}_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^n)$ is an arbitrary sequence converging to f in the topology of $L^2(\mathbb{R}^n, dx)$. By inner product continuity, the extended operator $\widehat{\mathscr{F}}$ will preserve the inner product of $L^2(\mathbb{R}^2, dx)$, and as such $\widehat{\mathscr{F}}$ will be 1-1 on $L^2(\mathbb{R}^n, dx)$. The following elementary argument explains why $\widehat{\mathscr{F}}$ is surjective too. Beside $\widehat{\mathscr{F}}$, we can construct the operator that extends to $L^2(\mathbb{R}^n, dx)$ the inverse Fourier transform $\widehat{\mathscr{F}_-}$. On $\mathscr{S}(\mathbb{R}^n, dx)$

$$\mathscr{F}\mathscr{F}_- = I_{\mathscr{S}(\mathbb{R}^n)}.$$

Now pass to the L^2 extensions, by linearity and continuity, and recall that the unique linear extension of the identity from $\mathscr{S}(\mathbb{R}^n, dx)$ to $L^2(\mathbb{R}^n, dx)$ is the latter's identity operator I (constructed in the general way explained above). Then

$$\widehat{\mathscr{F}}\widehat{\mathscr{F}_-} = I\,.$$

The equation implies $\widehat{\mathscr{F}}$ is onto.

Definition 3.83. *The unique operator $\widehat{\mathscr{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dx)$ that extends linearly and continuously the Fourier transform restricted to $\mathscr{S}(\mathbb{R}^n)$ is called* **Fourier-Plancherel transform**.

Theorem 3.84 (Plancherel). *The Fourier-Plancherel transform:*

$$\widehat{\mathscr{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dx)$$

is a bijective and isometric linear operator.

Proof. The proof was given immediately before Definition 3.83. □

There is still one issue we have to deal with. If $f \in L^1(\mathbb{R}^n, dx) \cap L^2(\mathbb{R}^n, dx)$ (but $f \notin \mathscr{S}(\mathbb{R}^n)$), *a priori* $\mathscr{F} f$ and $\widehat{\mathscr{F}} f$ are different, because to define $\widehat{\mathscr{F}}$ we did not extend $\mathscr{F}$ from $L^1(\mathbb{R}^n, dx)$, but rather from the subspace $\mathscr{S}(\mathbb{R}^n)$. This was the only possible choice because $L^1(\mathbb{R}^n, dx) \not\subset L^2(\mathbb{R}^n, dx)$.

The next proposition sheds light on the matter, and provides a practical method to compute the Fourier-Plancherel transform by means of limits of Fourier transforms.

Remarks 3.85. Recall that if $K \subset \mathbb{R}^n$ is a finite-measure set, in particular compact, (compact sets have finite Lebesgue measure):

(1) $L^2(K, dx) \subset L^1(K, dx)$.
(2) If $\{f_n\}_{n\in\mathbb{N}} \subset L^2(K, dx)$ converges in norm $|| \, ||_2$ to $f \in L^2(K, dx)$, it converges in norm $|| \, ||_1$ to f.
(3) $L^\infty(K, dx) \subset L^p(K, dx)$, $1 \le p < \infty$.
(4) If $\{f_n\}_{n\in\mathbb{N}} \subset L^\infty(K, dx)$ converges in norm $|| \, ||_\infty$ to $f \in L^\infty(K, dx)$, it converges to f in norm $|| \, ||_p$ as well.

These four statements are proved as follows: concerning the first two, recall the constant map 1 on a compact (of finite measure) set is integrable. Since

$$2|f(x)| \le |f(x)|^2 + 1\,,$$

the integral of the left is bounded by the integral of the right, so we have statement one. As for the second claim, the Cauchy-Schwarz inequality

$$\left(\int_K |g(x)|\, 1\, dx\right)^2 \leq \left(\int_K |g(x)|^2 dx\right)\left(\int_K 1 dx\right)$$

with $f(x) - f_n(x)$ replacing $g(x)$ proves it. To get the last two, note that by definition of Lebesgue integral:

$$\int_K |g|^p dx \leq \operatorname{ess\,sup}_K |g|^p \int_K dx = (||g||_\infty)^p \int_K dx$$

for any measurable map g on K. ■

Proposition 3.86. *The Fourier-Plancherel and Fourier transforms satisfy the following properties.*

(a) *If $f \in L^2(\mathbb{R}^n, dx) \cap L^1(\mathbb{R}^n, dx)$, the Fourier-Plancherel transform reduces to the Fourier transform $\mathscr{F} f$ computed by the integral formula (3.61).*
(b) *If $f \in L^2(\mathbb{R}^n, dx)$, the Fourier-Plancherel transform can be computed as the limit (understood in $L^2(\mathbb{R}^n, dk)$)*

$$\hat{\mathscr{F}} f = \lim_{n\to+\infty} g_n$$

of

$$g_n(k) := \int_{K_n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x) dx\,, \tag{3.74}$$

where $K_n \subset \mathbb{R}^n$ are compact, $K_{m+1} \supset K_m$, $m = 1, 2, \ldots$, and $\cup_{m=1}^\infty K_m = \mathbb{R}^n$.

Proof. (a) Begin with proving the claim for $f \in L^2(\mathbb{R}^n, dx)$ different from 0 on a zero-measure set outside a compact K_0. Such f belongs to $L^1(\mathbb{R}^n, dx)$. Let then $\{s_n\}_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^n)$ be a sequence converging to f in $L^2(\mathbb{R}^n, dx)$. If B, B' are open balls of finite radius with $B \supset \overline{B'} \supset B' \supset K_0$, we can construct a function $h \in \mathscr{D}(\mathbb{R}^n)$ equal to 1 on B' and null outside B. Obviously, if $f_n := h \cdot s_n$, the sequence $\{f_n\}$ is in $\mathscr{D}(\mathbb{R}^n)$, and hence $\mathscr{S}(\mathbb{R}^n)$, with support inside the compact set $K := \overline{B}$. Therefore every f_n belongs to $L^1(\mathbb{R}^n, dx)$ and the sequence $\{f_n\}$ tends to f in $L^2(\mathbb{R}^n, dx)$ and $L^1(\mathbb{R}^n, dx)$.
By definition, as $f_n \to f$ in norm $||\,||_2$,

$$||\mathscr{F} f_n - \hat{\mathscr{F}} f||_2 \to 0 \tag{3.75}$$

as $n \to +\infty$. At the same time, since $f_n \to f$ in norm $||\,||_1$, by Proposition 3.81(a) we have $||\mathscr{F} f_n - \mathscr{F} f||_\infty \to 0$ as $n \to +\infty$. But on finite-measure sets convergence for $||\,||_\infty$ implies convergence for $||\,||_2$, so

$$||\mathscr{F} f_n - \mathscr{F} f||_2 \to 0 \tag{3.76}$$

and hence $\widehat{\mathscr{F}} f = \mathscr{F} f$ by (3.75) and by uniqueness of the limit.

Suppose now $f \in L^2(\mathbb{R}^n, dx) \cap L^1(\mathbb{R}^n, dx)$, and nothing more. Consider a collection of compact sets $\{K_n\}$ as in part (b). Define maps $f_n := \chi_{K_n} \cdot f$, where χ_E is the characteristic function of E ($\chi_E(x) = 0$ if $x \notin E$ and $\chi_E(x) = 1$ if $x \in E$). It is clear that pointwise $f_n \to f$ as $n \to +\infty$. Moreover $|f(x) - f_n(x)|^p \leq |f(x)|^p$, $p = 1, 2, \ldots$. By Lebesgue's dominated convergence theorem, $f_n \to f$, as $n \to +\infty$, for $|| \; ||_1$ and $|| \; ||_2$. On the other hand what we have proved just above tells:

$$\mathscr{F} f_n = \widehat{\mathscr{F}} f_n \,.$$

Proposition 3.81(a) gives $||\mathscr{F} f - \mathscr{F} f_n||_\infty \to 0$ and at the same time $||\widehat{\mathscr{F}} f - \mathscr{F} f_n||_2 \to 0$. These facts hold also when restricting $\widehat{\mathscr{F}} f$, $\mathscr{F} f$, $\mathscr{F} f_n$ to any compact set K. For maps that are zero outside a compact set uniform convergence implies L^2 convergence, so if x belongs to a compact set, $(\mathscr{F} f)(x) = (\widehat{\mathscr{F}} f)(x)$ almost everywhere. But every point $x \in \mathbb{R}^n$ belongs to some compact set, so $\mathscr{F} f = \widehat{\mathscr{F}} f$ as elements in $L^2(\mathbb{R}^n, dx)$.
(b) This was proved in the final part of (a). □

Examples 3.87. (1) There is an important property distinguishing $\mathscr{D}(\mathbb{R}^n)$ from $\mathscr{S}(\mathbb{R}^n)$: only the former is *not* invariant under Fourier transform (and inverse Fourier transform). Since $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$, in fact, it is clear that $\mathscr{F}(\mathscr{D}(\mathbb{R}^n)) \subset \mathscr{F}(\mathscr{S}(\mathbb{R}^n)) \subset \mathscr{S}(\mathbb{R}^n)$. This cannot be sharpened, by this reason:

Proposition 3.88. *Let $f \in \mathscr{D}(\mathbb{R}^n)$. If $\mathscr{F} f \in \mathscr{D}(\mathbb{R}^n)$ then $f = 0$. The same holds for the inverse Fourier transform.*

Proof. The proof is easy, and we show it only for $\mathscr{F}$, because the case $\mathscr{F}_-$ is similar. If

$$g(k) = \int_{\mathbb{R}^n} \frac{e^{-ik\cdot x}}{(2\pi)^{n/2}} f(x)\, dx \,,$$

where f has compact support, the integral converges also for $k \in \mathbb{C}^n$. Using Lebesgue's dominated convergence, moreover, we can differentiate the components k_i of k inside the integral, and their real and imaginary parts. Since $k \mapsto e^{-ik\cdot x}$ is analytic (in each variable k_i separately), it solves the Cauchy-Riemann equations in each k_i. Consequently also g will solve those equations in each k_i, becoming analytic on $\mathbb{C}^n$. The restriction of g to $\mathbb{R}^n$ defines, via its real and imaginary parts, real analytic maps on $\mathbb{R}^n$. If g has compact support, there will be an open, non-empty set in $\mathbb{R}^n$ where $Re\, g$ and $Im\, g$ vanish. A known property of real-analytic maps (of one real variable) on open connected sets (here $\mathbb{R}^n$) is that they vanish everywhere if they vanish on an open non-empty set of the domain. Therefore if g has compact support it is the zero map. Then also f is zero, since $\mathscr{F}$ is invertibile on $\mathscr{S}(\mathbb{R}^n)$. □

(2) Related to (1) is the known *Paley–Wiener theorem* (see for instance [KiGv82]):

Theorem 3.89 (Paley–Wiener). *Take $a > 0$ and consider $L^2([-a,a], dx)$ as subspace of $L^2(\mathbb{R}, dx)$. The space $\hat{\mathscr{F}}(L^2([-a,a], dx))$ consists of maps $g = g(k)$ that can be extended uniquely to an analytic map on the complex plane ($k \in \mathbb{C}$) such that*

$$|g(k)| \leq C e^{2\pi a |Im k|}\,, \quad k \in \mathbb{C}$$

for some constant $C \geq 0$ *depending on* g.

Since $\hat{\mathscr{F}}(L^2([-a,a],dx)) \subset L^2(\mathbb{R},dk)$ by Plancherel's theorem, the result of Paley–Wiener implies that analytic maps g bounded as above determine elements of $L^2(\mathbb{R},dk)$ when k restricts to $\mathbb{R}$. ■

To conclude, consider the space $L^2((a,b),dx)$, where $-\infty \leq a < b \leq +\infty$ and dx is the usual Lebesgue measure on $\mathbb{R}$. The following extremely practical fact, used in Example 3.32(4) to build bases, descends from the Fourier-Plancherel theory.

Proposition 3.90. *Let* $f : (a,b) \to \mathbb{C}$ *be a measurable map such that:*

(i) *the set* $\{x \in (a,b) \mid f(x) = 0\}$ *has zero measure;*
(ii) *there exist* $C, \delta > 0$ *for which* $|f(x)| < Ce^{-\delta|x|}$ *for any* $x \in (a,b)$.

Then the space finitely generated by $x \mapsto x^n f(x) =: f_n(x)$, $n = 0,1,2,\ldots$, *is dense in* $L^2((a,b),dx)$.

Proof. Let $S := \{f_n\}_{n\in\mathbb{N}}$. It is enough to prove $S^\perp = \{0\}$, because $S^\perp \oplus \overline{<S>} = L^2((a,b),dx)$ by Theorem 3.13. So take $h \in L^2((a,b),dx)$ such that

$$\int_a^b x^n f(x)\overline{h(x)}dx = 0$$

for any $n = 0,1,2,\ldots$. Extend h to the whole real line by setting it to zero outside (a,b). The above condition reads:

$$\int_{\mathbb{R}} x^n f(x)\overline{h(x)}dx = 0\,, \tag{3.77}$$

for any $n = 0,1,2,\ldots$. Moreover, the following three facts hold.

(i) $f \cdot \overline{h} \in L^1(\mathbb{R},dx)$: in fact both maps are in $L^2(\mathbb{R},dx)$, so their product is in $L^1(\mathbb{R},dx)$;
(ii) $f \cdot \overline{h} \in L^2(\mathbb{R},dx)$, because $|f(x)|^2 < C^2e^{-2\delta|x|} < C^2 < +\infty$ and $|h|^2$ is integrable by assumption;
(iii) The map sending $x \in \mathbb{R}$ to $e^{\delta'|x|}f(x)\overline{h}(x)$ is in $L^1(\mathbb{R},dx)$ for any $\delta' < \delta$. In fact, since $x \mapsto |e^{\delta'|x|}f(x)| \leq Ce^{-(\delta-\delta')|x|}$, the function $x \mapsto e^{\delta'|x|}f(x)$ is in $L^2(\mathbb{R},dx)$, and $h \in L^2(\mathbb{R},dx)$ by hypothesis, so the product belongs in $L^1(\mathbb{R},dx)$.
Using (i) we compute the Fourier transform:

$$g(k) = \int_{\mathbb{R}} \frac{e^{-ik\cdot x}}{\sqrt{2\pi}} f(x)\overline{h}(x)\,dx\,.$$

This concides with the Fourier-Plancherel transform of $f \cdot \overline{h}$ by (i), (ii) and Proposition 3.86(a). Using (iii), if k is complex and $|Imk| < \delta$, then $g = g(k)$ is well defined and analytic on the open strip $B \subset \mathbb{C}$ given by $Rek \in \mathbb{R}$, $|Imk| < \delta$; this is proved similarly to what we did in example (1). Lebesgue's dominated convergence and exchanging derivatives and integrals allow to see that

$$\frac{d^n g}{dk^n}|_{k=0} = \frac{(-i)^n}{\sqrt{2\pi}} \int_{\mathbb{R}} x^n f(x)\overline{h(x)}dx$$

for any $n = 0, 1, \ldots$. All derivatives vanish by (3.77), and so the Taylor expansion of g at the origin is zero. This annihilates g on an open disc contained in B, so analyticity guarantees g is zero on the open connected set B, and in particular on the real axis k. Therefore the Fourier-Plancherel transform of $f \cdot \overline{h}$ is the null vector of $L^2(\mathbb{R}, dk)$. Since the transform is unitary we conclude $f \cdot \overline{h} = 0$ almost everywhere on $\mathbb{R}$: in particular on (a, b), where by assumption $f \neq 0$ almost everywhere. But then $h = 0$ almost everywhere on (a, b), which is to say each $h \in S^\perp$ coincides with the null element in $L^2((a, b), dx)$, ending the proof. □

Exercises

3.1. The definition of a (semi-)inner product makes sense on real vector spaces as well, simply by replacing **H2** in Definition 3.1 with $S(u, v) = S(v, u)$, and using real linear combinations in **H1**.

Show that with this definition Proposition 3.3 still holds, provided the polarisation formula is written as in (3.7).

3.2. (*Hard*) Consider a real vector space and prove that if a (semi)norm p satisfies the parallelogram identity (3.3):

$$p(x+y)^2 + p(x-y)^2 = 2(p(x)^2 + p(y)^2) ,$$

then there exists a unique (semi-)inner product S, defined in 3.1, inducing p via (3.2).

Solution. If S is a (semi-)inner product on the real vector space X, we have the polarisation identity (3.7):

$$S(x, y) = \frac{1}{4} \left(S(x+y, x+y) - S(x-y, x-y) \right) .$$

This implies uniqueness of S, as $S(z, z) = p(z)^2$. For the existence from a given norm p set:

$$S(x, y) := \frac{1}{4} \left(p(x+y)^2 - p(x-y)^2 \right) .$$

We prove S is a semi-inner product or an inner product according to whether p is a norm or a seminorm. If this is true and if p is a norm, substituting S to p on the right above gives $S(x, x) = 0$ and $x = 0$, making S an inner product.

To finish we need to prove, for any $x, y, z \in \mathsf{X}$:
(a) $S(\alpha x, y) = \alpha S(x, y)$ if $\alpha \in \mathbb{R}$;
(b) $S(x+y, z) = S(x, z) + S(y, z)$;
(c) $S(x, y) = S(y, x)$;
(d) $S(x, x) = p(x)^2$.

Properties (c) and (d) are straightforward from the definition of S. Let us prove (a) and (b). By (3.3) and the definition of S:

$$S(x,z)+S(y,z) = 4^{-1}\left(p(x+z)^2 - p(x-z)^2 + p(y+z)^2 - p(y-z)^2\right)$$
$$= 2^{-1}\left(p\left(\frac{x+y}{2}+z\right)^2 - p\left(\frac{x+y}{2}-z\right)^2\right) = 2S\left(\frac{x+y}{2},z\right).$$

Hence

$$S(x,z)+S(y,z) = 2S\left(\frac{x+y}{2},z\right). \tag{3.78}$$

Then (a) clearly implies (b), and we have to prove (a) only. Setting $y=0$ in (3.78) and recalling $S(0,z)=0$ by definition of S,

$$S(x,z) = 2S(x/2,z).$$

Iterating the formula gives (a) for $\alpha = m/2^n$, $m,n = 0,1,2,\dots$. These numbers are dense in $[0,+\infty)$. At the same time $\mathbb{R} \ni \alpha \mapsto p(\alpha x+z)$ and $\mathbb{R} \ni \alpha \mapsto p(\alpha x - z)$ are both continuous (in the topology induced by p), so $S(x,y) := \frac{1}{4}\left(p(x+y,x+y) - p(x-y,x-y)\right)$ allows to conclude $\mathbb{R} \ni \alpha \mapsto S(\alpha x,y)$ is continuous in α. That is to say, (a) holds for any $\alpha \in [0,+\infty)$. Again by definition of S we have $S(-x,y) = -S(x,y)$, so the previous result is valid for any $\alpha \in \mathbb{R}$ and the proof is over.

3.3. (*Hard*) Suppose a (semi)norm p satisfies the parallelogram rule (3.3):

$$p(x+y)^2 + p(x-y)^2 = 2(p(x)^2 + p(y)^2),$$

on a $\mathbb{C}$-vector space. Show that there is a unique (semi-)inner product S inducing p via (3.2).

Solution. If S is a (semi-)inner product on the complex vector space X we have the polarisation formula (3.4):

$$4S(x,y) = S(x+y,x+y) - S(x-y,x-y) - iS(x+iy,x+iy) + iS(x-iy,x-iy).$$

Since $S(z,z) = p(z)^2$, as in the real case, that implies uniqueness of S for a given norm p on X. Existence: define, for given (semi)norm p and $x,y \in \mathsf{X}$:

$$S_1(x,y) := 4^{-1}(p(x+y)^2 - p(x-y)^2), \quad S(x,y) := S_1(x,y) - iS_1(x,iy).$$

Notice $S(x,x) = p(x)^2$, and if p is a norm, by construction $S(x,x)=0$ implies $x=0$. There remains to show S as above is a Hermitian (semi-)inner product. By Definition 3.1 we have to show:

(a) $S(\alpha x,y) = \alpha S(x,y)$ if $\alpha \in \mathbb{C}$;
(b) $S(x+y,z) = S(x,z) + S(y,z)$;
(c) $S(x,y) = \overline{S(y,x)}$;
(d) $S(x,x) = p(x)^2$.

The last one is true by construction. Proceeding as in the previous exercise, using S_1 instead of S, we can prove (b) for S_1, (a) for S_1 *with* $\alpha \in \mathbb{R}$, and also $S_1(x,y) = S_1(y,x)$. These, using the definition of S in terms of S_1, imply (a), (b) and (c).

3.4. Prove the claim in Remark 3.4(1) on a (semi-)inner product space (X, S): the (semi-)inner product $S : \mathsf{X} \times \mathsf{X} \to \mathbb{C}$ is continuous in the product topology of $\mathsf{X} \times \mathsf{X}$, having on X the topology induced by the (semi-)inner product itself. Consequently it is continuous in either argument separately.

Hint. Suppose $\mathsf{X} \times \mathsf{X} \ni (x_n, y_n) \to (x, y) \in \mathsf{X} \times \mathsf{X}$ as $n \to +\infty$. Use the Cauchy-Schwarz inequality to show that if S is the (semi-)inner product associated to p, then:

$$|S(x,y) - S(x_n, y_n)| \leq p(x_n)p(y_n - y) + p(x_n - x)p(y)\,.$$

Recall $p(x_n) \to p(x)$ and canonical projections are continuous in the product topology.

3.5. Prove Proposition 3.8: a linear operator $L : \mathsf{X} \to \mathsf{Y}$ between inner product spaces is an isometry, in the sense of Definition 3.6, if and only if

$$||Lx||_{\mathsf{Y}} = ||x||_{\mathsf{X}} \quad \text{for any } x \in \mathsf{X},$$

where norms are associated to the respective inner product spaces.

Hint. Polarise.

3.6. Consider the Banach space $\ell^p(\mathbb{N})$, for $p \geq 1$. Show that for $p \neq 2$ one cannot define any Hermitian inner product so to induce the usual norm $||\ ||_p$. Conclude that $\ell^p(\mathbb{N})$ cannot be rendered a Hilbert space for $p \neq 2$.

Hint. Show there are pairs of vectors f, g violating the parallelogram rule. E.g. $f = (1,1,0,0,\ldots)$ and $g = (1,-1,0,0,\ldots)$.

3.7. Prove that the Banach space $(C([0,\pi/2]), ||\ ||_\infty)$ does not admit a Hermitian inner product inducing $||\ ||_\infty$, i.e.: $(C([0,\pi/2]), ||\ ||_\infty)$ cannot be made into a Hilbert space.

Hint. Show there are pairs of vectors f, g violating the parallelogram rule. Consider for example $f(x) = \cos x$ and $g(x) = \sin x$.

3.8. In the Hilbert space H consider a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{H}$ weakly converging to $x \in \mathsf{H}$. I.e., $f(x_n) \to f(x)$, $n \to +\infty$, for any $f \in \mathsf{H}'$. Show that, in general, $x_n \not\to x$ in the topology of H. However, if we assume additionally $||x_n|| \to ||x||$, $n \to +\infty$, then $x_n \to x$, $n \to +\infty$, in the topology of H.

Hint. Riesz's theorem implies that $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{H}$ weakly converges to $x \in \mathsf{H}$ iff $(z|x_n) \to (z|x)$, $n \to +\infty$, for any $z \in \mathsf{H}$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a basis of H, thought of as separable. Then $x_n \to 0$ weakly but not in the topology of H. For the second claim, note $||x - x_n||^2 = ||x||^2 + ||x_n||^2 - 2Re(x|x_n)$.

3.9. Consider the basis of $L^2([-L/2, L/2], dx)$ formed by the functions (up to zero-measure sets):

$$e_n(x) := \frac{e^{i2\pi nx/L}}{\sqrt{L}} \quad n \in \mathbb{Z}\,.$$

Suppose, for $f \in L^2([-L/2,L/2],dx)$, that the series

$$\sum_{n\in\mathbb{Z}} (e_n|f)e_n(x)$$

converges to some g in norm $||\ ||_\infty$. Prove $f(x) = g(x)$ a.e.

Hint. Compute the components $(e_n|g)$ using the fact that the integral of an absolutely convergent series on $[a,b]$ is the series of the integrated summands. Check that $(e_n|g) = (e_n|f)$ for any $n \in \mathbb{Z}$.

3.10. Consider the basis of $L^2([-L/2,L/2],dx)$ made by the functions e_n of Exercise 3.9. Suppose $f : [-L/2,L/2] \to \mathbb{C}$ is continuous, $f(-L/2) = f(L/2)$, and f is piecewise C^1 on $[-L/2,L/2]$ (i.e. $[-L/2,L/2] = [a_1,a_2] \cup [a_2,a_3] \cup \cdots \cup [a_{n-2},a_{n-1}] \cup [a_{n-1},a_n]$ and $f\restriction_{[a_i,a_{i+1}]} \in C^1([a_i,a_{i+1}])$ for any i, understanding boundary derivatives as left and right derivatives). Show:

$$f(x) = \sum_{n\in\mathbb{Z}} (e_n|f)e_n(x) \quad \text{for any } x \in [-L/2,L/2]$$

where

$$e_n(x) := \frac{e^{i2\pi nx/L}}{\sqrt{L}} \quad n \in \mathbb{Z}.$$

Prove the series converges uniformly.

Hint. Compute the components $(e_n|df/dx)$ by integration by parts: $|n(e_n|f)| = 2c|(e_n|df/dx)|$, where $c = L/(4\pi)$. Then

$$|(e_n|f)| = c2|(e_n|df/dx)||1/n| \leq c(|(e_n|df/dx)|^2 + 1/n^2), \quad n \neq 0.$$

Now, df/dx gives an L^2 map, the series of generic term $1/n^2$ converges, and $|e_n(x)| = 1$ for any x. Therefore the series

$$\sum_{n\in\mathbb{Z}} (e_n|f)e_n(x)$$

converges uniformly, i.e. in norm $||\ ||_\infty$. Apply Exercise 3.9.

3.11. Rephrase and prove Exercise 3.10, replacing the requirement that f be continuous and piecewise C^1 with the demand that f be *absolutely continuous* on $[-L/2,L/2]$ either with *essentially bounded* derivative, or with derivative in $\mathscr{L}^2([-L/2,L/2],dx)$.

Hint. Remember Theorem 1.75(a).

3.12. Consider the basis $\{e_n\}$ of $L^2([-L/2,L/2],dx)$ of Exercise 3.9. Let $f : [-L/2,L/2] \to \mathbb{C}$ be of class C^N, suppose $d^kf/dx^k|_{-L/2} = d^kf/dx^k|_{L/2}$, $k = 0,1,\ldots,N$ and that f is piecewise C^{N+1} on $[-L/2,L/2]$. Prove

$$\frac{d^kf(x)}{dx^k} = \sum_{n\in\mathbb{Z}} (e_n|f)\frac{d^k}{dx^k}e_n(x) \quad \text{for any } x \in [-L/2,L/2]$$

where

$$e_n(x) := \frac{e^{i2\pi nx/L}}{\sqrt{L}} \quad n \in \mathbb{Z},$$

and the series' convergence is uniform, $k = 0, 1, 2, \ldots, N$.

Hint. Iterate the procedure of Exercise 3.10, bearing in mind that we can swap derivatives and sum in a convergent series of C^1 maps whose derivative series converges uniformly.

3.13. Prove that the functions $[0,L] \ni x \mapsto s_n(x) := \sqrt{\frac{2}{L}} \sin\left(\frac{\pi nx}{L}\right)$, $n = 1, 2, 3, \ldots$, form an orthonormal system in $L^2([0,L], dx)$.

Sketch. A direct computation tells $||s_n|| = 1$. Then observe that if $\Delta := -\frac{d^2}{dx^2}$ then $\Delta s_n = \left(\frac{\pi nx}{L}\right)^2 s_n$. Thus if $(\,|\,)$ is the inner product in $L^2([0,L], dx)$:

$$(s_n|s_m) = \frac{1}{n}(\Delta s_n|s_m) = \frac{1}{n}(s_n|\Delta s_m) = \frac{m}{n}(s_n|s_m)$$

where, in the middle, we integrated twice by parts to shift Δ from the left to the right, and we used $s_k(0) = s_k(L) = 0$ to annihilate the boundary terms.
Therefore

$$\left(1 - \frac{m}{n}\right)(s_n|s_m) = 0,$$

implying $(s_n|s_m) = 0$ if $n \neq m$.

3.14. Prove that the maps $[0,L] \ni x \mapsto c_n(x) := \sqrt{\frac{2}{L}} \cos\left(\frac{\pi nx}{L}\right)$, $n = 0, 1, 2, \ldots$, form an orthonormal system in $L^2([0,L], dx)$.

Hint. Proceed exactly as in Exercise 3.13. If $\Delta := -\frac{d^2}{dx^2}$ we still have $\Delta c_n = \left(\frac{\pi nx}{L}\right)^2 c_n$, but the difference is that now it is the derivatives of c_n that vanish on the boundary of $[0,L]$.

3.15. Recall that the space $\mathscr{D}((0,L))$ of smooth maps with compact support in $(0,L)$ is dense in $L^2([0,L], dx)$ in the latter's topology. Using Exercise 3.10 prove that the functions $[0,L] \ni x \mapsto s_n(x) := \sqrt{\frac{2}{L}} \sin\left(\frac{\pi nx}{L}\right)$, $n = 1, 2, 3, \ldots$, are a basis of $L^2([0,L], dx)$.

Outline. It suffices to prove that the space $< s_n >_{n=1,2,\ldots}$ of finite linear combinations of the s_n is dense in $\mathscr{D}((0,L))$ in norm $||\ ||_\infty$, because this would imply, by elementary integral properties, that they are dense in the topology of $L^2([0,L], dx)$. Since $\mathscr{D}((0,L))$ is dense in $L^2([0,L], dx)$, we would have $< s_n >_{n=1,2,\ldots}$ dense in $L^2([0,L], dx)$. Because $\{s_n\}_{n=1,2,\ldots}$ is an orthonormal system (Exercise 3.13), this would in turn imply the claim, by Theorem 3.26. To show $< s_n >_{n=1,2,\ldots}$ is dense in $\mathscr{D}((0,L))$ with respect to $||\ ||_\infty$, fix $f \in \mathscr{D}((0,L))$ and extend it to F on $[-L,L]$ by imposing F be an odd map. By construction F is in $\mathscr{D}((-L,L))$ and satisfies

$F(-L) = F(L)$, because it and its derivatives vanish around $x = 0$ and $x = \pm L$. *A fortiori* F is continuous and piecewise C^1 on $[-L, L]$. Applying Exercise 3.10 we conclude:

$$F(x) = \sum_{n \in \mathbb{N}} (F|e_n) \frac{e^{i\pi nx/L}}{\sqrt{2L}}$$

where now

$$e_n(x) := \frac{e^{i\pi nx/L}}{\sqrt{2L}} ,$$

and the series' convergence is in norm $|| \; ||_\infty$. Since F is odd:

$$F(x) = -F(-x) = -\sum_{n \in \mathbb{N}} (F|e_n) \frac{e^{-i\pi nx/L}}{\sqrt{2L}}$$

adding which to the previous expression of $F(x)$ gives:

$$2F(x) = \sum_{n \in \mathbb{N}} \frac{2i(F|e_n)}{\sqrt{2L}} \sin\left(\frac{\pi nx}{L}\right) .$$

Restricting to $x \in [0, L]$:

$$f(x) = \sum_{n \in \mathbb{N}} \frac{i(F|e_n)}{\sqrt{2L}} \sin\left(\frac{\pi nx}{L}\right) .$$

Since the convergence is in norm $|| \; ||_\infty$, we have the claim.

3.16. Recall that the space $\mathscr{D}((0, L))$ of smooth maps with compact support in $(0, L)$ is dense in $L^2([0, L], dx)$ in the latter's topology. Using Exercise 3.10 prove that the functions $[0, L] \ni x \mapsto c_n(x) := \sqrt{\frac{2}{L}} \cos\left(\frac{\pi nx}{L}\right)$ are a basis of $L^2([0, L], dx)$.

Hint. Proceed as in Exercise 3.15, extending f to an even function on $[-L, L]$.

3.17. Let $C \subset \mathsf{H}$ be a closed subspace in the Hilbert space H. Prove C is weakly closed. Put otherwise, show that if $\{x_n\}_{n \in \mathbb{N}} \subset C$ converges weakly (cf. Exercise 3.8) to $x \in \mathsf{H}$, then $x \in C$.

Hint. If $P_C : \mathsf{H} \to C$ is the orthogonal projector onto C, show $P_C x_n \to P_C x$ weakly.

3.18. Let H be a Hilbert space and $T : D(T) \to \mathsf{H}$ a linear operator, where $D(T) \subset \mathsf{H}$ is a dense subspace in H ($D(T) = \mathsf{H}$, possibly). Prove that if $(u|Tu) = 0$ for any $u \in D(T)$ then $T = 0$, i.e. T is the null operator (sending everything to 0).

Solution. We have

$$0 = (u + v|T(u + v)) = (u|Tu) + (v|Tv) + (u|Tv) + (v|Tu) = (u|Tv) + (v|Tu)$$

and similarly

$$0 = i(u + iv|T(u + iv)) = i(u|Tu) + i(v|Tv) - (u|Tv) + (v|Tu) = -(u|Tv) + (v|Tu) .$$

Adding these two gives $(v|Tu) = 0$ for any $u, v \in D(T)$. Choose $\{v_n\}_{n\in\mathbb{N}} \subset D(T)$ such that $v_n \to Tu$, $n \to +\infty$; then $||Tu||^2 = (Tu|Tu) = \lim_{n\to+\infty}(v_n|Tu) = 0$ for any $u \in D(T)$, i.e. $Tu = 0$ for any $u \in D(T)$, hence $T = 0$.

3.19. Consider $L^2([0,1],m)$ where m is the Lebesgue measure, and take $f \in \mathscr{L}^2([0,1],m)$. Let $T_f : L^2([0,1],m) \ni g \mapsto f \cdot g$, with $\cdot$ being the standard pointwise product of functions. Prove T_f is well defined, bounded with norm $||T_f|| \leq ||f||$ and normal. Moreover, show T_f is self-adjoint iff f is real-valued up to a zero-measure set in $[0,1]$.

3.20. Let $T \in \mathfrak{B}(\mathsf{H})$ be *self-adjoint*. For $\lambda \in \mathbb{R}$ consider the series of operators

$$U(\lambda) := \sum_{n=0}^{\infty} (i\lambda)^n \frac{T^n}{n!},$$

where $T^0 := I$, $T^1 := T$, $T^2 := TT$ and so on, and the convergence is uniform. Prove the series converges to a unitary operator.

Hint. Proceed as when proving the properties of the exponential map from its definition as series.

3.21. Referring to the previous exercise, show that $\lambda, \mu \in \mathbb{R}$ imply $U(\lambda)U(\mu) = U(\lambda + \mu)$.

3.22. Show that the series of Exercise 3.20 converges for any $\lambda \in \mathbb{C}$ to a bounded operator, and that $U(\lambda)$ is always normal.

3.23. Show that the operator $U(\lambda)$ in Exercise 3.20 is positive if $\lambda \in i\mathbb{R}$. Are there values $\lambda \in \mathbb{C}$ for which $U(\lambda)$ is a projector (not necessarily orthogonal)?

3.24. Compute explicitly $U(\lambda)$ in Exercise 3.20 if T is defined by T_f of Exercise 3.19 with $f = \overline{f}$.

3.25. In $\ell^2(\mathbb{N})$ consider the operator $T : \{x_n\} \mapsto \{x_{n+1}/n\}$. Prove T is bounded and compute T^*.

3.26. Consider the Volterra operator $T : L^2([0,1],dx) \to L^2([0,1],dx)$:

$$(Tf)(x) = \int_0^x f(t)dt\,.$$

Prove it is well defined, bounded and its adjoint satisfies:

$$(T^*f)(x) = \int_x^1 f(t)dt \quad \text{for any } f \in L^2([0,1],dx).$$

Hint. Since $[0,1]$ has finite Lebesgue measure, $L^2([0,1],dx) \subset L^1([0,1],dx)$. Then use Theorem 1.75.

3.27. Let $\mathfrak{A}$ be a C^*-algebra without unit. Consider the direct sum $\mathfrak{A} \oplus \mathbb{C}$ and define the product:

$$(x,c) \cdot (y,c') := (x \circ y + cy + c'x, cc'), \quad (x',c'),(x,c) \in \mathfrak{A} \oplus \mathbb{C}$$

on it, where $\circ$ is the product on $\mathfrak{A}$. Define the norm:

$$||(x,c)|| := \sup\{||cy + xy|| \,|\, y \in \mathfrak{A}, ||y|| = 1\}$$

and the involution: $(x,c)^* = (x^*, \overline{c})$, where $\overline{c}$ is the complex conjugate of c and the involution on the right is the one of $\mathfrak{A}$. Prove that the vector space $\mathfrak{A} \oplus \mathbb{C}$ with the above extra structure is a C^*-algebra with unit $(0,1)$.

Hint. The triangle inequality is easy. The proof that $||(x,c)|| = 0$ implies $c = 0$ and $x = 0$ goes as follows. If $c = 0$, $||(x,0)|| = 0$ means $||x|| = 0$, so $x = 0$. If $c \neq 0$, we can simply look at $c = 1$. In that case $||y + xy|| \leq ||y||\, ||(x,1)||$, so $||(x,1)|| = 0$ implies $y = xy$ for any $y \in \mathfrak{A}$. Using the involution we have $y = yx^*$ for any $y \in \mathfrak{B}$. In particular $x^* = xx^* = x$, and then $y = xy = yx$ for any $y \in \mathfrak{A}$. Therefore x is the identity of $\mathfrak{A}$, which has no unit. The contradiction says $c = 0$ is the only possibility, and we fall back to the previous case. Let us see to the C^* properties of the norm. By definition of norm: $||(c,x)||^2 = \sup\{||cy + xy||^2 \,|\, y \in \mathfrak{A}, ||y|| = 1\} = \sup\{||y^*(\overline{c}cy + \overline{c}xy + cx^*y + x^*xy)|| \,|\, y \in \mathfrak{A}, ||y|| = 1\}$. Hence $||(c,x)||^2 \leq ||(c,x)^*(c,x)|| \leq ||(c,x)^*||\,||(c,x)||$. In particular $||(c,x)|| \leq ||(c,x)^*||$, and replacing (c,x) with $(c,x)^*$ gives $||(c,x)^*|| = ||(c,x)||$. The inequality $||(c,x)||^2 \leq ||(c,x)^*(c,x)|| \leq ||(c,x)^*||\, ||(c,x)||$ implies $||(c,x)||^2 \leq ||(c,x)^*(c,x)|| \leq ||(c,x)||^2$, and so $||(c,x)||^2 = ||(c,x)^*(c,x)||$.

4

Families of compact operators on Hilbert spaces and fundamental properties

Measure what can be measured, and make measurable what can't be.

Galileo Galilei

The aim of this chapter, from the point of view of QM applications, is to introduce certain types of operators used to define *quantum states*. These operators, known in the literature as *operators of trace class*, or *nuclear operators*, are bounded operators on a Hilbert space that admit a trace. In order to introduce them it is necessary to define first *compact operators*, also known as *completely continuous* operators, that play an important role in several branches of mathematics and physical applications independent of quantum theories.

The first section will introduce the general notion of *compact operator* on a normed space, then briefly discuss general properties in normed and Banach spaces. We will prove the classical result on the non-compactness of the infinite-dimensional unit ball.

In section two we specialise to Hilbert spaces, with an eye to L^2 spaces on which compact operators (such as Hilbert–Schmidt operators) admit an integral representation. We will show that the set of compact operators determines a closed *-ideal in the C^*-algebra of bounded operators on a Hilbert space, hence, *a fortiori*, a C^*-subalgebra. We will prove the celebrated theorem due to Hilbert on the spectral expansion of compact operators, to be considered as a precursor of all spectral decomposition results of subsequent chapters.

The *-ideal of *Hilbert–Schmidt operators* and their elementary properties are the subject of section four. We will show, in particular, that Hilbert–Schmidt operators form a Hilbert space.

The penultimate section will be concerned with the *-ideal of operators of trace class, and the proofs of the basic (and most useful in physics) properties. In particular, the *ciclicity property of trace* will be proved.

The final section is devoted to a short introduction to *Fredholm's alternative theorem* for Fredholm's integral equations of the second kind.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_4,

4.1 Compact operators in normed and Banach spaces

This section deals with compact operators in normed spaces. It starts with recalling general results about compact subsets in normed spaces, especially infinite-dimensional ones. The next section will discuss the theory in Hilbert spaces.

4.1.1 Compact sets in (infinite-dimensional) normed spaces

In a completely general topological space X a *compact* set is defined by 1.19, which we recall below.

Definition. *Let* $(\mathsf{X},\mathscr{T})$ *be a topological space and* $K \subset \mathsf{X}$.
(a) *K is* **compact** *if any open covering of it admits a finite subcovering: if* $\{A_i\}_{i\in I} \subset \mathscr{T}$, $\cup_{i\in I} A_i \supset K$, *then* $\cup_{i\in J} A_i \supset K$ *for some finite* $J \subset I$.
(b) *K is* **relatively compact** *if* $\overline{K}$ *is compact.*
(c) X *is* **locally compact** *if every point admits a relatively compact open neighbourhood.*

Related to this is the notion of *sequential compactness*.

Definition 4.1. *A subset K in a topological space is* **sequentially compact** *if any sequence* $\{x_n\}_{n\in\mathbb{N}} \subset K$ *has a subsequence* $\{x_{n_p}\}_{p\in\mathbb{N}}$ *that converges in K.*

Remark 4.2. Let us list below a few general features of compact sets that should be known from basic topology courses [Ser94II]. We shall make use of them later.

(1) Compactness is hereditary, in the sense that it is passed on to induced topologies.
(2) Closed subsets in compact sets are compact, and in Hausdorff spaces (like normed vector spaces, Hilbert spaces), compact sets are closed.
(3) In metrisable spaces (in particular normed vector spaces, Hilbert spaces), compactness is *equivalent* to sequential compactness. ■

We prove the next useful property, valid in metric spaces as well.

Proposition 4.3. *Let* $(\mathsf{X},||\ ||)$ *be a normed space and* $A \subset \mathsf{X}$. *If any sequence in A admits a converging subsequence (not in A necessarily), then A is relatively compact.*

Proof. The only thing to prove is, if $\{y_k\}_{k\in\mathbb{N}} \subset \overline{A}$, that there is a subsequence of $\{y_k\}_{k\in\mathbb{N}}$ that converges (in $\overline{A}$, being closed). Given $\{y_k\}_{k\in\mathbb{N}} \subset \overline{A}$, there will be sequences $\{x_n^{(k)}\}_{n\in\mathbb{N}} \subset A$, one for each k, with $x_n^{(k)} \to y_k$ as $n \to +\infty$. Fix k and take the corresponding n_k big enough. Then we can construct, term by term, a new sequence $\{z_k := x_{n_k}^{(k)}\}_{k\in\mathbb{N}} \subset A$ such that $||y_k - z_k|| < 1/k$. Under the assumptions made on A there will be a subsequence $\{z_{k_p}\}_{p\in\mathbb{N}}$ of $\{z_k\}_{k\in\mathbb{N}}$ converging to some $y \in \overline{A}$. Then

$$||y_{k_p} - y|| \le ||y_{k_p} - z_{k_p}|| + ||z_{k_p} - y|| .$$

Since $1/k_p \to 0$ for $p \to +\infty$, given $\varepsilon > 0$ there will be P such that, if $p > P$, $||z_{k_p} - y|| < \varepsilon/2$ and $1/k_p < \varepsilon/2$, so $||y_{k_p} - y|| < \varepsilon$. In other words $y_{k_p} \to y$ as $p \to +\infty$. □

Remarks 4.4. This proposition holds on metric spaces too, and the proof is the same with minor modifications. ■

Examples of compact sets in an infinite-dimensional normed space $(\mathsf{X}, ||\ ||)$ are easily obtained from finite-dimensional subspaces. As we know from Chapter 2.5, any finite-dimensional space S is homeomorphic to $\mathbb{C}^n$ (or $\mathbb{R}^n$ for real vector spaces); so, any closed and bounded set $K \subset \mathsf{S}$ (e.g., the closure of an open ball of finite radius) is compact in S by the Heine-Borel theorem. Since compactness is a hereditary property K is compact also in the topology of $(\mathsf{X}, ||\ ||)$.

The following is an important result, that discriminates between finite- and infinite-dimensional normed spaces. We leave to the reader the proof of the fact that the closure of an open ball in a normed space is nothing but the corresponding closed ball with the same centre and radius:

$$\overline{\{x \in \mathsf{X} \,|\, ||x - x_0|| < r\}} = \{x \in \mathsf{X} \,|\, ||x - x_0|| \leq r\}\,.$$

Proposition 4.5. *Let* $(\mathsf{X}, ||\ ||)$ *be a normed space of infinite dimension. The closure of the open unit ball* $\{x \in \mathsf{X} \,|\, ||x|| < 1\}$ *(that is the closed unit ball* $\{x \in \mathsf{X} \,|\, ||x|| \leq 1\}$*) cannot be compact.*
The same is true for any open ball, with arbitrary finite radius and centre.

Proof. Let us indicate by B the open unit ball centred at the origin, and suppose $\overline{B}$ is compact. Then we can cover $\overline{B}$, hence B, with $N > 0$ open balls B_k of radius $1/2$ centred at x_k, $k = 1, \dots, N$. Consider a subspace X_n in X, of finite dimension n, containing the vectors x_k. Since $\dim \mathsf{X}$ is infinite, we may choose $n > N$ as large as we want. Define further "balls" $P := B \cap \mathsf{X}_n$ of radius 1 and $P_k := B_k \cap X_n$, $k = 1, \dots, N$, all of radius $1/2$. Let us identify X_n with $\mathbb{R}^{2n}$ (or $\mathbb{R}^n$ if the field is $\mathbb{R}$) by choosing a basis of X_n, say $\{z_k\}_{k=1,\dots,n}$. Notice that a "ball" P_k does not necessarily have the shape of a Euclidean ball. If we normalise the Lebesgue measure m on $\mathbb{R}^{2n}$ by dividing by the volume of P (non-zero since P is open, non-empty by Proposition 2.103), then $m(P) = 1$. Let us show $m(P_k) = (1/2)^n$. Lebesgue's measure in translation-invariant, so we may limit ourselves to balls $B(r)$ centred at the origin of radius r. Since every norm is a homogeneous function, $B(\lambda r) = \{\lambda u \,|\, u \in B(r)\} =: \lambda B(r)$ for any $\lambda > 0$. The Lebesgue measure on $\mathbb{R}^{2n}$ satisfies $m(\lambda E) = \lambda^{2n} m(E)$, hence $m(P_k) = m((1/2)P) = (1/2)^n m(P) = (1/2)^n$. Eventually, as $B \subset \cup_{k=1}^N B_k$ and $P \subset \cup_{k=1}^N P_k$, we have $m(P) \leq \sum_{k=1}^N m(P_k)$ by sub-additivity, i.e. $1 \leq N(1/2)^{2n}$. But this is impossible if n is big enough (N is fixed). We prove it similarly for any open ball of finite radius and centred at any point in the normed space. □

The next result explains, once more, how compact sets acquire 'counter-intuitive' properties in passing from finitely to infinitely many dimensions. In the standard $\mathbb{C}^n$ or $\mathbb{R}^n$ there are compact sets with non-empty interior: it is enough to close any bounded open set. The Heine-Borel theorem warrants that the closure (still bounded) is compact and clearly has points in its interior.

The complete space $\mathbb{C}$ can be obtained alike, from the union of a countable collection of compact subsets: we can take open discs with rational centres and radii. In the infinite-dimensional case the picture changes dramatically.

Corollary 4.6. *Let* X *be an infinite-dimensional normed space.*
(a) *If* $K \subset \mathsf{X}$ *is compact, the interior of* K *is empty.*
(b) *If* X *is also complete (i.e. a Banach space),* X *cannot be obtained as a countable union of compact subsets.*

Proof. (a) Suppose the interior of K is not empty; then it contains an open ball B, since open balls form a basis for the topology. Compact subsets are closed because normed spaces are Hausdorff, so $\overline{B} \subset \overline{K} = K$. Closed subsets in compact sets are compact, hence $\overline{B}$ should be compact, contradicting the previous proposition.
(b) The claim follows from (a) and the last statement in Baire's Theorem 2.89, where X is our complete normed space. □

4.1.2 Compact operators in normed spaces

We are ready to introduce *compact operators*. Recall that a subset M in a normed space $(\mathsf{X}, || \, ||)$ is **bounded** (in norm $|| \, ||$) if there is an open ball $B_\delta(x_0)$, of finite radius $\delta > 0$ and centred at some $x_0 \in \mathsf{X}$, such that $M \subset B_\delta(x_0)$.

Clearly, M is bounded if and only if there is a metric ball of finite radius $\delta > 0$ and *centred at the origin* of X, containing M (just choose as radius $\delta + ||x_0||$).

Definition 4.7. *Let* X, Y *be normed spaces over* $\mathbb{R}$*, or* $\mathbb{C}$. $T \in \mathfrak{L}(\mathsf{X}, \mathsf{Y})$ *is a* **compact operator** *(or* **completely continuous** *operator) when either of the following equivalent conditions holds:*

(a) *For any bounded subset* $M \subset \mathsf{X}$*,* $T(M)$ *is relatively compact in* Y.
(b) *If* $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ *is bounded, there is a subsequence* $\{x_{n_k}\}_{k\in\mathbb{N}}$ *such that* $\{Tx_{n_k}\}_{k\in\mathbb{N}}$ *converges in* Y.

$\mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$ *denotes the subset of compact operators from* X *to* Y*, and* $\mathfrak{B}_\infty(\mathsf{X})$ *the subset of compact operators on* X.

Remark 4.8. (1) Clearly (a) $\Rightarrow$ (b). The opposite implication (b) $\Rightarrow$ (a) is an immediate consequence of Proposition 4.3.
(2) Any compact operator is certainly bounded. In fact, the unit closed ball centred at the origin is mapped to a set (containing the origin) with compact closure K. The latter can be covered by N open balls of radius $r > 0$, say $B_r(y_i)$. Then $K \subset \cup_{i=1}^N B_r(y_i) \subset B_{R+r}(0)$, where R is the maximum distance between the centres y_i and the origin. In particular $||T(x)|| \leq (R+r)$ for $||x|| = 1$, so $||T|| \leq r+R < +\infty$. ■

The sets $\mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$ and $\mathfrak{B}_\infty(\mathsf{X})$ are actually vector spaces with the usual linear combinations of operators, hence subspaces of $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$ and $\mathfrak{B}(\mathsf{X})$ respectively. Not only that, but they enjoy the next properties as well.

Proposition 4.9. *If* X*,* Y *are normed spaces,* $\mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$ *satisfies these two properties.*

(a) $\mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$ *is a vector subspace of* $\mathfrak{B}(\mathsf{X}, \mathsf{Y})$.
(b) *If* Z *is a normed space and* $A \in \mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$*:*

(i) $B \in \mathfrak{B}(\mathsf{Z},\mathsf{X})$ *implies* $AB \in \mathfrak{B}_\infty(\mathsf{Z},\mathsf{Y})$*;*
(ii) $B \in \mathfrak{B}(\mathsf{Y},\mathsf{Z})$ *implies* $BA \in \mathfrak{B}_\infty(\mathsf{X},\mathsf{Z})$.

(c) *If* Y *is a Banach space and* $\{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{B}_\infty(\mathsf{X},\mathsf{Y})$ *converges uniformly to* $A \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$*, then* $A \in \mathfrak{B}_\infty(\mathsf{X},\mathsf{Y})$*. I.e.* $\mathfrak{B}_\infty(\mathsf{X},\mathsf{Y})$ *is a closed subspace in the Banach space* $(\mathfrak{B}_\infty(\mathsf{X},\mathsf{Y}),||\ ||)$*, where* $||\ ||$ *is the operator norm.*

Proof. (a) Consider the operator $\alpha A+\beta B$, $\alpha,\beta \in \mathbb{C}$, $A,B \in \mathfrak{B}_\infty(\mathsf{X},\mathsf{Y})$. Let us prove it is compact by showing that any bounded sequence $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ has a subsequence $\{x_{n_r}\}_{r\in\mathbb{N}} \subset \mathsf{X}$ whose image $\{(\alpha A+\beta B)(x_{n_r})\}_{r\in\mathbb{N}} \subset \mathsf{Y}$ converges.

Let $\{x_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ be a bounded sequence. There is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ for which $Ax_{n_k} \subset \mathsf{Y}$ converges, as A is compact. The subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ is also bounded by assumption, so there is a sub-subsequence $\{x_{n_{k_m}}\}_{m\in\mathbb{N}}$ such that $Bx_{n_{k_m}} \in \mathsf{Y}$ converges. Now, by construction, $\{x_{n_{k_m}}\}_{m\in\mathbb{N}}$ is a subsequence of $\{x_{n_k}\}_{k\in\mathbb{N}}$ for which $\alpha Ax_{n_{k_m}}+\beta Bx_{n_{k_m}} \subset \mathsf{Y}$ converges.
(b) In case (i), if $\{z_k\}_{k\in\mathbb{N}} \subset \mathsf{Z}$ is bounded by $M>0$, the values Bz_k form a bounded set by $||B||M$, as B is bounded. But A is compact, so there is a subsequence $\{z_{n_k}\}_{k\in\mathbb{N}}$ for which ABz_{n_k} converges. Thus AB is compact. In case (ii), as A is compact, if $\{x_k\}_{k\in\mathbb{N}} \subset \mathsf{X}$ is bounded there is a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ such that the Ax_{n_k} converge. Since B is continuous, also BAx_{n_k} converge, and thus BA is compact.
(c) Let $\mathfrak{B}(\mathsf{X},\mathsf{Y}) \ni A = lim_{i\to+\infty}A_i$ with $A_i \in \mathfrak{B}_\infty(\mathsf{X},\mathsf{Y})$. Take $\{x_n\}_{n\in\mathbb{N}}$ a bounded sequence in X: $||x_n|| \leq C$ for any n. We want to prove the existence of a converging subsequence of $\{Ax_n\}$. Using a hopefully-clear notation, we build recursively a family of subsequences:

$$\{x_n\} \supset \{x_n^{(1)}\} \supset \{x_n^{(2)}\} \supset \cdots \tag{4.1}$$

such that, for any $i=1,2,\ldots$, $\{x_n^{(i+1)}\}$ is a subsequence of $\{x_n^{(i)}\}$ with $\{A_{i+1}x_n^{(i+1)}\}$ converging. This is always possible, because any $\{x_n^{(i)}\}$ is bounded by C as subsequence of $\{x_n\}$, and A_{i+1} is compact by assumption. The subsequence of $\{Ax_n\}$ that will converge is $\{Ax_i^{(i)}\}$. From the triangle inequality

$$||Ax_i^{(i)} - Ax_k^{(k)}|| \leq ||Ax_i^{(i)} - A_nx_i^{(i)}|| + ||A_nx_i^{(i)} - A_nx_k^{(k)}|| + ||A_nx_k^{(k)} - Ax_k^{(k)}|| \, .$$

With this estimate,

$$||Ax_i^{(i)} - Ax_k^{(k)}|| \leq ||A-A_n||(||x_i^{(i)}|| + ||x_k^{(k)}||) + ||A_nx_i^{(i)} - A_nx_k^{(k)}||$$

$$\leq 2C||A-A_n|| + ||A_nx_i^{(i)} - A_nx_k^{(k)}|| \, .$$

Given $\varepsilon>0$, if n is large enough $2C||A-A_n|| \leq \varepsilon/2$, since $A_n \to A$. Fix n and $r \geq n$. Then $\{A_n(x_p^{(r)})\}_p$ is a subsequence of the converging $\{A_n(x_p^{(n)})\}_p$. Consider the sequence $\{A_n(x_p^{(p)})\}_p$, for $p \geq n$, that picks up the "diagonal" terms of all those subsequences, *each of which is a subsequence of the preceding one by (4.1)*; it is still a subsequence of the converging $\{A_n(x_p^{(n)})\}_p$, so it, too, converges (to the same limit). We conclude that if $i,k \geq n$ are large enough, $||A_nx_i^{(i)} - A_nx_k^{(k)}|| \leq \varepsilon/2$. Hence if i,k

are big enough then $||Ax_i^{(i)} - Ax_k^{(k)}|| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. This finishes the proof, for we have obtained a Cauchy subsequence in the Banach space Y, which must converge in the space. □

Keeping in mind Proposition 2.70, a remarkable property of compact operators is spelt out by the next fact.

Proposition 4.10. *Let* X, Y *be normed spaces. If* $\mathsf{X} \ni x_n \to x \in \mathsf{X}$ *weakly and* $T \in \mathfrak{B}_\infty(\mathsf{X},\mathsf{Y})$, *then* $||T(x_n) - T(x)||_Y \to 0$, $n \to +\infty$. *In other terms compact operators map weakly convergent sequences to sequences that converge in norm.*

Proof. See Exercise 4.5. □

The last general property of compact operators in normed spaces concerns eigenvalues. We will give a proof before passing to compact operators in Hilbert spaces.

Theorem 4.11 (On the eigenvalues of compact operators in normed spaces). *Let* $T \in \mathfrak{B}_\infty(\mathsf{X})$, X *a normed space.*
(a) *For any* $\delta > 0$ *there exist a finite number of eigenspaces of* T *with eigenvalues* λ *such that* $|\lambda| > \delta$.
(b) *If* $\lambda \neq 0$ *is an eigenvalue of* T, *the corresponding eigenspace has finite dimension.*
(c) *The eigenvalues of* T, *in general complex numbers, form a bounded, at most countable, set; they can be ordered by decreasing modulus:*

$$|\lambda_1| \geq |\lambda_2| \geq \cdots 0,$$

(possibly) with 0 *as unique limit point.*

Proof. Let us give, first, a lemma to be used in cases (a), (b).

Lemma 4.12 (Banach). *Let* $x_1, x_2, \ldots$ *be a sequence (finite or infinite) of linearly independent vectors in the normed space* X, *and* $\mathsf{X}_n := <\{x_1, x_2, \ldots, x_n\}>$. *Then there exists a corresponding sequence* $y_1, y_2, \ldots \subset \mathsf{X}$ *such that:*

(i) $||y_n|| = 1$;
(ii) $y_n \in \mathsf{X}_n$;
(iii) $d(y_n, \mathsf{X}_{n-1}) > 1/2$,

for any $n = 1, 2, \ldots$, *where* $d(y_n, \mathsf{X}_{n-1})$ *is the distance of* y_n *from* X_{n-1}:

$$d(y_n, \mathsf{X}_{n-1}) = \inf_{x \in \mathsf{X}_{n-1}} ||x - y_n|| \,.$$

Proof of Lemma 4.12. Observe $d(y_n, \mathsf{X}_{n-1})$ exists and is finite, since it is the infimum of a non-empty real set that is bounded from below by 0. Choose $y_1 := x_1/||x_1||$ and build the sequence $\{y_n\}$ inductively as follows. The vectors $x_1, x_2, \ldots$ are linearly independent, so $x_n \notin \mathsf{X}_{n-1}$ and $d(x_n, \mathsf{X}_{n-1}) = \alpha > 0$. So let $x' \in \mathsf{X}_{n-1}$ be such that $\alpha < ||x_n - x'|| < 2\alpha$. As $\alpha = d(x_n, \mathsf{X}_{n-1}) = d(x_n - x', \mathsf{X}_{n-1})$, the vector

$$y_n := \frac{x_n - x'}{||x_n - x'||}$$

satisfies (i), (ii), (iii). □

Let us resume the proof of (a)–(b). If X is finite-dimensional the claims hold because eigenvectors with distinct eigenvalues are linearly independent. So consider X infinite-dimensional, where there can be infinitely many eigenvalues and eigenvectors. The proof of both (a) and (b) follows simultaneously from the existence, for any $\delta > 0$, of a finite number of linearly independent eigenvectors corresponding to eigenvalues λ with $|\lambda| > \delta$. Let us prove this, then. Let $\lambda_1, \lambda_2, \ldots$ be a sequence of eigenvalues of T, possibly repeated, such that $|\lambda_n| > \delta$. Assume, *by contradiction*, there is an infinite sequence $x_1, x_2, \ldots$, of corresponding linearly independent eigenvectors. We are claiming, by refuting the theorem, that there are *infinitely many* linearly independent eigenvectors with eigenvalues λ such that $|\lambda| > \delta$. Using the preliminary lemma, construct the sequence $y_1, y_2, \ldots$ fulfilling (i), (ii) and (iii), where X_n is spanned by $x_1, x_2, \ldots, x_n$. Since $|\lambda_n| > \delta$, the sequence $\{\frac{y_n}{\lambda_n}\}_{n=1,2,\ldots}$ is bounded. Now we show that we cannot extract a converging subsequence from the images $\{T\frac{y_n}{\lambda_n}\}_{n=1,2,\ldots}$. By construction, namely,

$$y_n := \sum_{k=1}^{n} \beta_k x_k \,,$$

so

$$T\frac{y_n}{\lambda_n} = \sum_{k=1}^{n-1} \frac{\beta_k \lambda_k}{\lambda_n} x_k + \beta_n x_n = y_n + z_n \,,$$

where

$$z_n := \sum_{k=1}^{n-1} \beta_k \left(\frac{\lambda_k}{\lambda_n} - 1 \right) x_k \in \mathsf{X}_{n-1} \,.$$

For any $i > j$, then:

$$\begin{aligned} \left|\left| T\left(\frac{y_i}{\lambda_i}\right) - T\left(\frac{y_j}{\lambda_j}\right) \right|\right| &= ||y_i + z_i - (y_j + z_j)|| \\ &= ||y_i - (y_j + z_j - z_i)|| > 1/2 \end{aligned}$$

as $y_j + z_j - z_i \in \mathsf{X}_{i-1}$. This is clearly incompatible with the compactness of T. Therefore we have to conclude that an infinite sequence of linearly independent eigenvectors $x_1, x_2, \ldots$ cannot exist. This ends the proof of (a) and (b).
(c) This part follows from (a) by picking a sequence of numbers $\delta > 0$ of the form $\delta_n = 1/n$, $n = 1, 2, 3, \ldots$. □

Remark 4.13. (1) One final property, that we shall not prove, establishes that *in the Banach setting* the conjugate (cf. Definition 2.42) to a compact operator is compact. We will prove it for the Hermitian adjoint to a compact operator in a Hilbert space.
(2) From Lemma 4.12 descends an alternative proof that the closed unit ball in an infinite-dimensional normed space is not compact (see Exercise 4.2). ■

4.2 Compact operators in Hilbert spaces

From now on we will consider compact operators in Hilbert spaces, even if certain properties are valid in less-structured spaces, like normed or Banach spaces.

4.2.1 General properties and examples

In the first theorem we prove about compact operators in Hilbert spaces, the completeness assumption is necessary only for the last statement.

Before, though, we need a preparatory proposition.

Proposition 4.14. *Let* H *be a Hilbert space. Then* $A \in \mathfrak{B}(\mathsf{H})$ *is compact iff* $|A|$ *is compact (see Definition 3.69).*

Proof. Assume A compact. Let $\{x_k\}_{k\in\mathbb{N}}$ be a bounded sequence in H and $\{Ax_{k_n}\}_{n\in\mathbb{N}}$ a subsequence of $\{Ax_k\}_{k\in\mathbb{N}}$ that converges, by virtue of compactness. Since the latter is a Cauchy subsequence by (3.54), the subsequence $\{|A|x_{k_n}\}_{n\in\mathbb{N}}$ of $\{|A|x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence and converges. Thus $|A|$ is compact. Swapping A and $|A|$ and repeating the proof proves the reverse implication. □

Theorem 4.15. *Let* H *be a Hilbert space.*
(a) $\mathfrak{B}_\infty(\mathsf{H})$ *is a subspace in* $\mathfrak{B}(\mathsf{H})$.
(b) $\mathfrak{B}_\infty(\mathsf{H})$ *is a two-sided* **-ideal of* $\mathfrak{B}(\mathsf{H})$*; i.e., beside being a subspace,* $\mathfrak{B}_\infty(\mathsf{H})$ *is such that* $TK, KT \in \mathfrak{B}_\infty(\mathsf{H})$ *for any* $T \in \mathfrak{B}_\infty(\mathsf{H})$, $K \in \mathfrak{B}(\mathsf{H})$, *and* $T^* \in \mathfrak{B}_\infty(\mathsf{H})$.
(c) $\mathfrak{B}_\infty(\mathsf{H})$ *is closed in the uniform topology, hence a* C^**-algebra (without unit if* $\dim \mathsf{H} = \infty$*), actually a* C^**-subalgebra of* $\mathfrak{B}(\mathsf{H})$.

Proof. (a) We proved this, in the more general setting of normed spaces, in Proposition 4.9(a).
(b) Proposition 4.9(b) shows that, in normed spaces, multiplying a compact operator on the left or right by a bounded operator produces a compact operator. To show closure under Hermitian conjugation, let us observe $|T|$ is compact if and only if T is, by Proposition 4.14. From the polar decomposition $T = U|T|$ of Theorem 3.71, we have $T^* = |T|U^*$, where we used $|T| \geq 0$, so $|T|$ is self-adjoint. The boundedness of U^* together with the compactness of $|T|$ force $T^* = |T|U^*$ to be compact by what we saw at the beginning.
(c) This part follows directly from Proposition 4.9(c) and the definition of C^*-algebra (recall H has infinite dimension, so the identity operator I is not compact, for otherwise the closed ball would be compact, and we know this cannot be). □

Examples 4.16. (1) If X, Y are normed spaces and $T \in \mathfrak{B}(\mathsf{X}, \mathsf{Y})$ has $\dim Ran(T)$ finite, T must be compact. Let us prove it. If $V \subset \mathsf{X}$ is bounded, i.e. $V \subset B_r(0)$ for a finite $r > 0$, then $||T(V)|| \leq r||T|| < +\infty$, whence $T(V)$ is bounded. $\overline{T(V)}$ is closed and bounded in a normed space of finite dimension that is homeomorphic to $\mathbb{C}^n$ (Proposition 2.103). By Heine-Borel $\overline{T(V)}$ is compact in the topology induced on the range of T. Thus T is compact, because compactness in the induced topology is the same as ambient compactness. As further subcase, if H is a Hilbert space consider an operator $T_x \in \mathfrak{L}(\mathsf{H})$ of the form

$$T_x : u \mapsto (x|u)y\,,$$

where $x, y \in \mathsf{H}$ are given vectors (possibly equal). This operator is compact, for it has finite-dimensional range.

(2) If $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ are orthogonal subsets in H, and if $T = \sum_{n\in\mathbb{N}}(x_n|\ )y_n$ is a bounded operator (interpreting the series in the uniform topology), then T is compact by (a) and (c) in Theorem 4.15.

(3) In $\ell^2(\mathbb{N})$ consider the operator $A : \{x_n\} \mapsto \{x_{n+1}/n\}$. It is compact because uniform limit of:

$$A_m : \{x_n\} \mapsto \{x_2/1, x_3/2, \ldots, x_{m+1}/m, 0, 0, \ldots\}$$

for $n = 1,2\ldots$ It is easy to prove (exercise)

$$||A - A_n|| \leq 1/(n+1)\,.$$

(4) Let X have measure μ on the σ-algebra Σ of X and let μ be σ-finite, so to define the product measure $\mu \otimes \mu$; we will use the simpler notation $L^2(\mu) := L^2(\mathsf{X},\mu)$ and $L^2(\mu\otimes\mu) := L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$. We consider $K \in L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$ and prove that

$$T_K : L^2(\mu) \ni f \mapsto \int_{\mathsf{X}} K(x,y) f(y) d\mu(y)$$

defines a compact operator $T_K \in \mathfrak{B}(L^2(\mathsf{X},\mu))$ in case μ is separable (cf. Examples 3.32(3)). Firts of all, irrespective of separability, if $f \in L^2(\mu)$:

$$\int_{\mathsf{X}} K(\cdot,y) f(y) d\mu(y) \in L^2(\mu)$$

and

$$\left|\left|\int_{\mathsf{X}} K(\cdot,y) f(y) d\mu(y)\right|\right|_{L^2(\mathsf{X},\mu)} \leq ||K||_{L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)} ||f||_{L^2(\mathsf{X},\mu)}\,,$$

which is to say:

$$||T_K|| \leq ||K||_{L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)}\,. \tag{4.2}$$

The proof of this is entirely based on the theorem of Fubini–Tonelli: if $K \in L^2(\mu\otimes\mu)$, by Fubini–Tonelli we have:

(1) $|K(x,\cdot)|^2 \in L^1(\mu)$, μ-almost everywhere;
(2) $\int_{\mathsf{X}} |K(\cdot,y)|^2 d\mu(y) \in L^1(\mu)\,.$

From (1) $K(x,\cdot) \in L^2(\mu)$ a.e., so $K(x,\cdot)f \in L^1(\mu)$ a.e. By the Cauchy-Schwarz inequality:

(3) $\int_{\mathsf{X}} |K(x,y)||f(y)| d\mu(y) \leq ||K(x,\cdot)||_{L^2} ||f||_{L^2}$.

Setting $F(x) := \int_{\mathsf{X}} K(x,y) f(y) d\mu(y)$, F is measurable, and by (3):

(4) $|F(x)|^2 \leq ||f||^2_{L^2} \int_{\mathsf{X}} |K(x,y)|^2 d\mu(y)$.

From (2) we have $|F|^2 \in L^2(\mu)$, so it is true that

$$\int_{\mathsf{X}} K(\cdot,y) f(y) d\mu(y) \in L^2(\mu)\,.$$

By (4) and Fubini–Tonelli, finally, we obtain

$$\left|\left|\int_{\mathsf{X}} K(\cdot,y)f(y)d\mu(y)\right|\right|_{L^2(\mu)} \leq ||K||_{L^2(\mu\otimes\mu)}||f||_{L^2(\mu)}\,,$$

hence (4.2).

In order to show T_K is compact, let us assume further μ is separable, so to make $L^2(\mathsf{X},\mu)$ separable (see Proposition 3.33). For instance, X could be an interval (or a Borel set) in $\mathbb{R}$ and μ the Lebesgue measure on $\mathbb{R}$. If $\{u_n\}_{n\in\mathbb{N}}$ is a basis of $L^2(\mathsf{X},\mu)$, $\{u_n\cdot\overline{u_m}\}_{n,m\in\mathbb{N}}$ is a basis of $L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)$ ($\cdot$ is the ordinary pointwise product of functions). Then, in the topology of $L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)$, we have

$$K = \sum_{n,m} k_{nm}u_n\cdot\overline{u_m}\,,$$

where the numbers $k_{nm}\in\mathbb{C}$ depend on K. So, setting

$$K_p := \sum_{n,m\leq p} k_{nm}u_n\cdot\overline{u_m}$$

we have $K_p\to K$ as $p\to+\infty$ in $L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)$. Applying (4.2) to $T_{K_p-K} = T_{K_p} - T_K$, where T_{K_p} is induced by the integral kernel K_p, we have

$$||T_K - T_{K_p}|| = \left|\left|\sum_{n,m>p} k_{nm}u_n\cdot\overline{u_m}\right|\right|_{L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)} \to 0\,,$$

as $p\to+\infty$. Thus T_K is compact, because the operators T_{K_p} are compact being finite sums of operators with finite-dimensional ranges, like those of example (1) above. Even without assuming μ separable, and demanding instead $K\in L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)$ and μ σ-finite, it is easy to see that

$$T_{\overline{K}} = T_K^*\,, \tag{4.3}$$

where $\overline{K}(x,y) := \overline{K(x,y)}$ for any $x,y\in\mathsf{X}$, and the bar is complex conjugation. The proof follows from Proposition 3.36 and Fubini–Tonelli. ∎

4.2.2 Spectral decomposition of compact operators on Hilbert spaces

Compact operators on Hilbert spaces enjoy ramarkable properties: concerning eigenvectors, eigenvalues and eigenspaces, in particular, the features of compact and self-adjoint operators generalise to infinite dimensions the properties of Hermitian matrices. The first two results clarify the matter.

Theorem 4.17 (Hilbert). *Let* H *be a Hilbert space,* $T\in\mathfrak{B}_\infty(\mathsf{H})$ *with* $T=T^*$.
(a) *Every eigenspace of* T *with eigenvalue* $\lambda\neq 0$ *has finite dimension.*
(b) *The set* $\sigma_p(T)$ *of eigenvalues of* T *is:*

(1) *non-empty;*
(2) *real;*
(3) *at most countable;*
(4) *it has one limit point at most, and this is* 0*;*
(5) *it satisfies:*

$$\sup\{|\lambda| \mid \lambda \in \sigma_p(T)\} = ||T|| .$$

More precisely, the least upper bound is the maximum $\Lambda \in \sigma_p(T)$*, where*

$$\Lambda = ||T|| \;\; \textit{if} \;\; \sup_{||x||=1}(x|Tx) = ||T|| , \tag{4.4}$$

or

$$\Lambda = -||T|| \;\; \textit{if} \;\; \inf_{||x||=1}(x|Tx) = -||T|| . \tag{4.5}$$

(6) *T coincides with the null operator iff* 0 *is the only eigenvalue.*

Partial proof. (a) Let H_λ be the eigenspace of T with eigenvalue $\lambda \neq 0$. If $B \subset \mathsf{H}_\lambda$ is the unit open ball at the origin, we can write $B = T(\lambda^{-1}B)$, and $\lambda^{-1}B$ is bounded by construction. Since T is compact, $\overline{B}$ is compact too. Hence in the Hilbert space H_λ the closure of the unit open ball is compact, and so $\dim \mathsf{H}_\lambda < +\infty$ by Proposition 4.5.
(b) We will prove but items (3) and (4), which will be part of the next theorem. If $\sigma_p(T)$ is not empty it must consist of real numbers by Proposition 3.54(c) part (ii), T being self-adjoint. By Proposition 3.54(a) $-||T|| \leq (x|Tx) \leq ||T||$ for any unit x, so only one of two possibilities can occur: either $\sup_{||x||=1}(x|Tx) = ||T||$ or $\inf_{||x||=1}(x|Tx) = -||T||$. Suppose the former is true, the other case being analogous by flipping the sign of T to $-T$. Assume $||T|| > 0$, otherwise the theorem is trivial. For any eigenvalue λ choose an eigenvector x with $||x|| = 1$, so $||T|| \geq |(x|Tx)| = |\lambda|(x|x) = |\lambda|$, and then $\sup|\sigma_p(T)| \leq ||T||$. To prove (5) it suffices to exhibit an eigenvector with eigenvalue $\Lambda = ||T||$. This also proves $\sigma_p(T) \neq \varnothing$ by the way. By assumption there is a sequence of unit points x_n such that $(x_n|Tx_n) \to ||T|| =: \Lambda > 0$. Using $||Tx_n|| \leq ||T|| ||x_n|| = ||T||$, we have

$$||Tx_n - \Lambda x_n||^2 = ||Tx_n||^2 - 2\Lambda\,(x_n|Tx_n) + \Lambda^2 \leq ||T||^2 + \Lambda^2 - 2\Lambda\,(x_n|Tx_n) .$$

As $||T|| = \Lambda$, taking the limit for $n \to +\infty$ in the inequality gives

$$Tx_n - \Lambda x_n \to 0 . \tag{4.6}$$

To conclude it would be enough to show either $\{x_n\}_{n\in\mathbb{N}}$ converges, or a subsequence does. As $||x_n|| = 1$, the sequence $\{x_n\}_{n\in\mathbb{N}}$ is bounded; but T is compact, so we may extract from $\{Tx_n\}_{n\in\mathbb{N}}$ a converging subsequence $\{Tx_{n_k}\}_{k\in\mathbb{N}}$.
Formula (4.6) implies

$$x_{n_k} = \frac{1}{\Lambda}[Tx_{n_k} - (Tx_{n_k} - \Lambda x_{n_k})]$$

converges to some $x \in \mathsf{H}$, $k \to +\infty$, as *linear combination of converging sequences*. Since T is continuous and $x_{n_k} \to x$, (4.6) forces

$$Tx = \Lambda x .$$

Observe that $x \neq 0$ because $||x|| = \lim_{k\to+\infty} ||x_{n_k}|| = 1$. So we have shown x is an eigenvector with eigenvalue Λ.
(6) is an immediate consequence of (5). $\square$

Let us move on to the celebrated theorem of Hilbert on the expansion of self-adjoint compact operators in terms of a basis made of eigenvectors.

Theorem 4.18 (Hilbert). *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* $T \in \mathfrak{B}_\infty(T)$ *with* $T = T^*$.

(a) *If* P_λ *is the orthogonal projector on the eigenspace with* $\lambda \in \sigma_p(T)$ *(eigenvalue set of* T*),*

$$T = \sum_{\lambda\in\sigma_p(T)} \lambda P_\lambda \,. \tag{4.7}$$

If $\sigma_p(T)$ *is infinite, the series (4.7) is understood in uniform topology, and the eigenvalues:* $\lambda_0, \lambda_1, \ldots$ *(*$\lambda_i \neq \lambda_j$*,* $i \neq j$*) are ordered so that* $|\lambda_0| \geq |\lambda_1| \geq |\lambda_2| \geq \ldots$.
(b) *If* B_λ *is a basis for the eigenspace of* T *associated to* $\lambda \in \sigma_p(T)$*, then* $\cup_{\lambda\in\sigma_p(T)} B_\lambda$ *is a basis for* H*; put equivalently,* H *admits a basis of eigenvectors of* T.

Remarks 4.19. Notice that there can only be at most two distinct non-trivial eigenvalues with equal absolute value (the eigenvalues are real). As

$$|\lambda_0| \geq |\lambda_1| \geq |\lambda_2| \geq \ldots,$$

the ambiguity in ordering the terms of the series regards pairs λ, λ' with $|\lambda| = |\lambda'|$. As we shall see after the proof, the sum of the series does not depend on this choice. ■

Proof (including parts (3), (4) of Theorem 4.17). (a) Let λ be an eigenvalue with eigenspace H_λ. Call P_λ the orthogonal projector on H_λ and $Q_\lambda := I - P_\lambda$ the orthogonal projector on $\mathsf{H}_\lambda^\perp$. Then

$$TP_\lambda = P_\lambda T = \lambda P_\lambda \,. \tag{4.8}$$

In fact, if $x \in \mathsf{H}$, $P_\lambda x \in \mathsf{H}_\lambda$ then $TP_\lambda x = \lambda P_\lambda x$, so $TP_\lambda = \lambda P_\lambda$. Taking adjoints and recalling $\lambda \in \mathbb{R}$, $T = T^*$, $P_\lambda = P_\lambda^*$, we find $P_\lambda T = \lambda P_\lambda = TP_\lambda$. A further consequence is, directly by definition of Q_λ and the above, that

$$Q_\lambda T = TQ_\lambda \,. \tag{4.9}$$

Observe that from $I = P_\lambda + Q_\lambda$ we infer $T = P_\lambda T + Q_\lambda T$, i.e.

$$T = \lambda P_\lambda + Q_\lambda T \,. \tag{4.10}$$

The operator $Q_\lambda T$:

(i) is self-adjoint, for $(Q_\lambda T)^* = T^* Q_\lambda^* = TQ_\lambda = Q_\lambda T$;
(ii) is compact by Theorem 4.15(b);
(iii) satisfies, by construction, $P_\lambda(Q_\lambda T) = (Q_\lambda T)P_\lambda = 0$ since $P_\lambda Q_\lambda = Q_\lambda P_\lambda = 0$.

In the rest of the proof these identities will be used without further mention, and we shall write P_n, Q_n, H_n instead of P_{λ_n}, Q_{λ_n}, H_{λ_n}.

Let us begin by choosing an eigenvalue $\lambda = \lambda_0$ with highest absolute value: there are, at most, two such eigenvalues differing by a sign, in which case we choose one. If $T_1 := Q_0 T$ then

$$T = \lambda_0 P_0 + T_1$$

where T_1 satisfies the above (i), (ii) and (iii). If $T_1 = 0$ the proof ends; if not, we know T_1 is self-adjoint and compact, so we can iterate the procedure using T_1 in place of T and finding, for $T_2 := Q_1 T_1$,

$$T = \lambda_0 P_0 + \lambda_1 P_1 + T_2 \,.$$

λ_1 is an eigenvalue of T_1, non-null and of highest absolute value (if a maximal eigenvalue were zero, then $T_1 = 0$ by Theorem 4.17(b, part 6)), and P_1 is the orthogonal projector on the eigenspace of T_1 relative to λ_1.

Observe that any eigenvalue λ_1 of T_1 is also an eigenvalue of T, because, if $T_1 u_1 = \lambda_1 u_1$,

$$Tu_1 = (\lambda_0 P_0 + T_1)u_1 = \lambda_0 P_0 T_1 \frac{1}{\lambda_1} u_1 + T_1 u_1 = \lambda_0 P_0 Q_0 T \frac{1}{\lambda_1} u_1 + T_1 u_1$$

$$= \lambda_0 \cdot 0 \cdot T \frac{1}{\lambda_1} u_1 + \lambda_1 u_1 = \lambda_1 u_1 \,.$$

What is more, $\lambda_1 \neq \lambda_0$ since $u_1 \in RanT_1 = Ran(Q_0 T) \subset \mathsf{H}_0^{\perp}$. At last, every eigenvector u of T with eigenvalue λ_1 is an eigenvector for T_1 relative to λ_1. In fact, using $T_1 = Q_0 T = (I - P_0)T$ we have, with $Tu = \lambda_1 u$,

$$T_1 u = \lambda_1 u - \lambda_1 P_0 u = \lambda_1 u + 0 = \lambda_1 u \,,$$

also using $P_0 u = 0$ (because eigenspaces with distinct eigenvalues are orthogonal for self-adjoint operators, like T). Overall, the eigenspace $\mathsf{H}_1^{(T_1)}$ relative to λ_1 coincides with the eigenspace H_1 of T with eigenvalue λ_1. Thus P_1 is the orthogonal projector in H on such eigenspace for T and T_1.
Since $|\lambda_0|$ is the maximum,

$$|\lambda_1| \leq |\lambda_0|.$$

There is an important consequence to this. Since $||T|| = |\lambda_0|$ and $||T_1|| = \lambda_1$ by the previous theorem,

$$||T_1|| \leq ||T|| \,.$$

If $T_2 = 0$ the proof ends, otherwise we proceed alike, finding

$$T - \sum_{k=0}^{n} \lambda_k P_k = T_n \,, \tag{4.11}$$

where

$$|\lambda_0| \geq |\lambda_1| \geq \cdots \geq |\lambda_k| \geq \ldots$$

and

$$||T_k|| = |\lambda_k| \,. \tag{4.12}$$

If none T_k is null, the process never stops. In such a case we claim that the decreasing sequence of positive numbers $|\lambda_k|$ must tend to 0 (there cannot be a larger limit point). Suppose $|\lambda_0| \geq |\lambda_1| \geq \cdots \geq |\lambda_k| \geq \ldots \geq \varepsilon > 0$ and pick a unit vector $x_n \in \mathsf{H}_n$ for any n. The sequence of the x_n is bounded, so the sequence of Tx_n, or a subsequence of it, must converge as T is compact. But this is impossible: expanding the squared norm of $||\lambda_n x_n - \lambda_m x_m||$, and because x_n and x_m are perpendicular (eigenvectors with distinct eigenvalues for a self-adjoint operator, cf. Proposition 3.54(b, part ii)), we have

$$||Tx_n - Tx_m||^2 = ||\lambda_n x_n - \lambda_m x_m||^2 = |\lambda_n|^2 + |\lambda_m|^2 \geq 2\varepsilon \,,$$

for any n, m. Thus neither the Tx_n, nor a subsequence, can converge, for they cannot be Cauchy. This is a contradiction, and therefore the sequence of λ_n (if there really are infinitely many thereof) converges to 0. By (4.11) and (4.12), what we have proved implies

$$T = \sum_{k=0}^{+\infty} \lambda_k P_k \tag{4.13}$$

in uniform topology. By construction the result does not depend on how we decide to order pairs of eigenvalues with equal absolute value. Now we prove that (4.13) coincides with (4.7), because the sequence of eigenvalues $\{\lambda_k\}$ found exhausts the set of eigenvalues of T except, possibly, for 0 (which at any rate interferes with neither of (4.13), (4.7)). Let $\lambda \neq \lambda_n$ for any n be an eigenvalue of T, and P_λ its orthogonal projector. Given that $P_n P_\lambda = 0$ for any n (again, by Proposition 3.54(b, part ii)), (4.13) implies

$$TP_\lambda = \sum_{k=0}^{+\infty} \lambda_k P_k P_\lambda = 0 \,,$$

whence, if $u \in \mathsf{H}_\lambda$,

$$Tu = TP_\lambda u = 0 \,.$$

This means $\lambda = 0$.

The proof of part (a) ends here, and in due course we have also justified the remaining part of Theorem 4.17.

(b) The bases B_λ always exist by Theorem 3.27, eigenspaces of T being closed (exercise) in H and hence Hilbert spaces themselves. Call $B := \cup_{\lambda \in \sigma_p(T)} B_\lambda$. We assert that if $u \in B^\perp$ then $u = 0$; since B is orthonormal, Definition 3.22 holds and the proof would be over. So let $u \in B^\perp$, so $u \perp B_\lambda$ for any $\lambda \in \sigma_p(T)$, and hence $P_\lambda u = 0$ for any $\lambda \in \sigma_p(T)$. Using decomposition (4.7) for T we find $Tu = 0$; hence u belongs to the eigenspace with zero eigenvalue H_0. But being u orthogonal to *every* eigenspace of T by construction, we must have $u \in \mathsf{H}_0$ and $u \in \mathsf{H}_0^\perp$, i.e. $u = 0$ as claimed. □

Hilbert's theorem, together with the polar decomposition Theorem 3.71, allows to generalise formula (4.7) for the expansion of self-adjoint compact operators to operators that are not self-adjoint. First, let us see a definition useful for the sequel.

Definition 4.20. *Let* H *be a Hilbert space and* $A \in \mathfrak{B}_\infty(\mathsf{H})$. *Non-zero eigenvalues* λ *of* $|A|$ *are called* **singular values** *of* A, *and their set is denoted* $sing(A)$. *The (finite) dimension* m_λ *of the eigenspace of* $\lambda \in sing(A)$ *is called* **multiplicity** *of* λ.

Theorem 4.21. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* $A \in \mathfrak{B}_\infty(\mathsf{H})$, $A \neq 0$. *If* $sing(A)$ *is ordered decreasingly,* $\lambda_0 > \lambda_1 > \lambda_2 > \ldots > 0$,

$$A = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda\, (u_{\lambda,i}|\ \ )\, v_{\lambda,i}\,, \tag{4.14}$$

$$A^* = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda\, (v_{\lambda,i}|\ \ )\, u_{\lambda,i}\,, \tag{4.15}$$

where the sums, if infinite, are meant in uniform topology, and for any $\lambda \in sing(A)$ *the set of* $u_{\lambda,1}, \ldots, u_{\lambda,m_\lambda}$ *is an orthonormal basis for the* λ*-eigenspace of* $|A|$. *Moreover, for any* $\lambda \in sing(A)$, $i = 1, 2, \ldots, m_\lambda$, *the vectors*

$$v_{\lambda,i} := \frac{1}{\lambda} A u_{\lambda,i}\,, \tag{4.16}$$

form an orthonormal system and

$$v_{\lambda,i} = U u_{\lambda,i}\,, \tag{4.17}$$

where U *is defined by the polar decomposition* $A = U|A|$.

Proof. $|A|$ is self-adjoint, positive and compact. Its eigenvalues are real, positive and satisfy conditions (a) and (b) in Theorem 4.17. We examine the case where the eigenvalue set is infinite (countable), leaving the finite case to the reader. Theorem 4.18 allows to expand $|A|$:

$$|A| = \sum_{\lambda \in \sigma_p(|A|)} \lambda P_\lambda\,,$$

where the convergence is uniform. It is clear we can reduce to non-zero eigenvalues since 0 does not give contributions to the the series

$$|A| = \sum_{\lambda \in sing(A)} \lambda P_\lambda\,.$$

If U is bounded and $\mathfrak{B}(\mathsf{H}) \ni T_n \to T \in \mathfrak{B}(\mathsf{H})$ uniformly, $UT_n \to UT$ in the uniform topology. Since U (from $A = U|A|$) is bounded, in the uniform topology we have:

$$A = U|A| = \sum_{\lambda \in sing(A)} \lambda U P_\lambda\,. \tag{4.18}$$

Theorem 4.17(a) says the closed projection space of each P_λ ($\lambda > 0$) has finite dimension m_λ. We can find an orthonormal basis for it: $\{u_{\lambda,i}\}_{i=1,\ldots,m_\lambda}$. Note $(u_{\lambda,i}|u_{\lambda',j}) = \delta_{\lambda\lambda'}\delta_{ij}$ by construction, as eigenvectors with distinct eigenvalues are orthogonal ($|A|$ is normal because positive) by virtue of Proposition 3.54(b, ii). From $u_{\lambda,i} =$

$|A|(u_{\lambda,i}/\lambda)$ we have $u_{\lambda,i} \in Ran(|A|)$. Thus U acts on $u_{\lambda,i}$ isometrically, and the vectors on the right in (4.17) are still orthonormal. Equation (4.17) is equivalent to (4.16) by polar decomposition:

$$Au_{\lambda,i} = U|A|u_{\lambda,i} = U\lambda u_{\lambda,i} = \lambda v_{\lambda,i}\,.$$

It is an easy exercise to show that the orthogonal projector P_λ $(\lambda > 0)$ can be written

$$P_\lambda = \sum_{i=1}^{m_\lambda} (u_{\lambda,i}|\ \ )\, u_{\lambda,i}\,.$$

Consequently,

$$UP_\lambda = \sum_{i=1}^{m_\lambda} (u_{\lambda,i}|\ \ )\, Uu_{\lambda,i} = \sum_{i=1}^{m_\lambda} (u_{\lambda,i}|\ \ )\, v_{\lambda,i}\,.$$

Substituting in (4.18) gives (4.14). Equation (4.15) arises from (4.14) if we consider the following two facts: (i) Hermitian conjugation $A \mapsto A^*$ is antilinear, it transforms linear combinations of operators into linear combinations of the adjoints, and the coefficients get conjugated; (ii) Hermitian conjugation is continuous in the uniform topology of $\mathfrak{B}(\mathsf{H})$ because Proposition 3.38(a) implies $||A^*|| = ||A||$.
From these two facts, (4.14) gives (recall $\lambda \in \mathbb{R}$):

$$A^* = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda\, [(u_{\lambda,i}|\ \ )\, v_{\lambda,i}]^*\,,$$

where the series converges in uniform topology. An easy exercise shows that

$$[(u_{\lambda,i}|\ \ )\, v_{\lambda,i}]^* = (v_{\lambda,i}|\ \ )\, u_{\lambda,i}\,,$$

from which (4.15) is immediate. □

The theorem just proved allows us to introduce Hilbert–Schmidt operators and operators of trace class, which we will describe in the ensuing sections.

4.3 Hilbert–Schmidt operators

A particular class of compact operators is that of Hilbert–Schmidt operators. They have several applications in wide branches of mathematical physics, beside QM. This section is devoted to their study and their main properties.

4.3.1 Main properties and examples

Warning. In this section, and sometimes also elsewhere, the operator norm $||\ \ ||_2$ will denote the *Hilbert–Schmidt* norm (see later) and not the usual L^2 norm. This should not cause ambiguity, for the correct meaning should be clear from the context. ■

Definition 4.22. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* $||\ ||$ *the inner product norm.* $A \in \mathfrak{B}(\mathsf{H})$ *is a* **Hilbert–Schmidt operator** *(HS) if there is a basis U such that* $\sum_{u\in U} ||Au||^2 < +\infty$ *in the sense of Definition 3.19. The class of Hilbert–Schmidt operators on* H *will be indicated with* $\mathfrak{B}_2(\mathsf{H})$*. If* $A \in \mathfrak{B}_2(\mathsf{H})$*,*

$$||A||_2 := \sqrt{\sum_{u\in U} ||Au||^2}\,, \tag{4.19}$$

where U is the above basis.

As first thing let us prove that the particular basis chosen in the definition is not important, and that also $||A||_2$ does not depend on it.

Proposition 4.23. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space with norm* $||\ ||$ *induced by the inner product, U and V bases (possibly coinciding) and* $A \in \mathfrak{B}(\mathsf{H})$*.*

(a) $\{||Au||^2\}_{u\in U}$ *has finite sum iff* $\{||Av||^2\}_{v\in V}$ *has finite sum. In that case:*

$$\sum_{u\in U} ||Au||^2 = \sum_{v\in V} ||Av||^2\,. \tag{4.20}$$

(b) $\{||Au||^2\}_{u\in U}$ *has finite sum iff* $\{||A^*v||^2\}_{v\in V}$ *has finite sum. If so:*

$$\sum_{u\in U} ||Au||^2 = \sum_{v\in V} ||A^*v||^2\,. \tag{4.21}$$

Proof. In the light of Theorem 3.26(d),

$$||Au||^2 = \sum_{v\in V} |(v|Au)|^2 < +\infty\,,$$

so, given u, only a countable number of coefficients $|(v|Au)|$, at most, is non-zero by Proposition 3.21(b). This gives at most a countable set $V(u) \subset V$ such that

$$\sum_{u\in U} ||Au||^2 = \sum_{u\in U} \sum_{v\in V(u)} |(v|Au)|^2 < +\infty\,. \tag{4.22}$$

In particular, using Proposition 3.21(b) again, it means that a countable (at most) set of $u \in U$ gives non-zero sum $\sum_{v\in V(u)} |(v|Au)|^2$. Therefore the coefficients $(v|Au)$ do not vanish only for a countable (at most) set Z of pairs $(u,v) \in U \times V$. Define sets (at most countable):

$$U_0 := \{u \in U \mid \text{there exists } v \in V \text{ with } (v|Au) \neq 0\}\,,$$

$$V_0 := \{v \in V \mid \text{there exists } u \in U \text{ with } (v|Au) \neq 0\}\,.$$

Thus $Z \subset U_0 \times V_0$. Endow U_0 and V_0 with counting measures μ and ν, and write the above series using integrals and these measures (Proposition 3.21(c)). In particular (4.22) becomes:

$$\sum_{u\in U} ||Au||^2 = \sum_{u\in U} \sum_{v\in V(u)} |(v|Au)|^2 = \int_{U_0} d\mu(u) \int_{V_0} d\nu(v) |(v|Au)|^2 < +\infty\,. \tag{4.23}$$

μ and ν are σ-finite because U_0 and V_0 are at most countable, so we can define the product $\mu \otimes \nu$ and use Fubini–Tonelli. Concerning the last part of (4.23), this theorem ensures that $(v,u) \mapsto |(v|Au)|^2$ is integrable in the product measure and we can swap integrals:

$$\sum_{u\in U} ||Au||^2 = \int_{U_0\times V_0} |(v|Au)|^2 d\mu(u) \otimes d\nu(v) = \int_{V_0} d\nu(v) \int_{U_0} d\mu(u) |(v|Au)|^2 < +\infty\,.$$

Note $(v|Au) = (A^*v|u)$, so just countably many, at most, products $(A^*v|u)$ (with $(u,v) \in U \times V$) will be different from zero, and in particular:

$$\sum_{v\in V}\sum_{u\in U} |(A^*v|u)|^2 = \int_{V_0} d\nu(v) \int_{U_0} d\mu(u) |(A^*v|u)|^2 = \sum_{u\in U} ||Au||^2 < +\infty\,.$$

But the left-hand side is precisely $\sum_{v\in V} ||A^*v||^2$. Thus we have proved this part of assertion (b): *if* $\{||Au||^2\}_{u\in U}$ *has finite sum, so does* $\{||A^*v||^2\}_{v\in V}$*, and the sums coincide.* Now we use the same proof, just exchanging bases and starting from A^*: recalling that $(A^*)^* = A$ for bounded operators, we can prove the remaining part of (b): *if* $\{||A^*v||^2\}_{v\in V}$ *has finite sum, then also* $\{||Au||^2\}_{u\in U}$ *does, and then (4.21) holds.*

The proof of (a) is straighforward from (b) because the bases used are arbitrary. □

With that settled we can discuss some of the many and interesting mathematical properties of HS operators. The most fascinating from a mathematical viewpoint is (b) in the next theorem: HS operators A form a Hilbert space whose inner product induces precisely the norm we called $||A||_2$. Another important fact is that HS operators are compact and their space is an ideal inside bounded operators, closed under Hermitian conjugation.

Theorem 4.24. *Hilbert–Schmidt operators on a Hilbert space* H *enjoy the following properties.*

(a) $\mathfrak{B}_2(\mathsf{H})$ *is a subspace in* $\mathfrak{B}(\mathsf{H})$ *and, actually, a two-sided* **-ideal in* $\mathfrak{B}(\mathsf{H})$*; moreover:*

(i) $||A||_2 = ||A^*||_2$ *for any* $A \in \mathfrak{B}_2(\mathsf{H})$*;*
(ii) $||AB||_2 \leq ||B||\,||A||_2$ *and* $||BA||_2 \leq ||B||\,||A||_2$ *for any* $A \in \mathfrak{B}_2(\mathsf{H})$, $B \in \mathfrak{B}(\mathsf{H})$*;*
(iii) $||A|| \leq ||A||_2$ *for any* $A \in \mathfrak{B}_2(\mathsf{H})$.

(b) *If* $A, B \in \mathfrak{B}_2(\mathsf{H})$ *and if* N *is a basis in* H*, define:*

$$(A|B)_2 := \sum_{x\in N} (Ax|Bx)\,. \tag{4.24}$$

The map

$$(\ |\ )_2 : \mathfrak{B}_2(\mathsf{H}) \times \mathfrak{B}_2(\mathsf{H}) \to \mathbb{C}$$

is well defined (the sum always reduces to an absolutely convergent series and does not depend on the basis) and determines an inner product on $\mathfrak{B}_2(\mathsf{H})$ *such that*

$$(A|A)_2 = ||A||_2^2 \tag{4.25}$$

for any $A \in \mathfrak{B}_2(\mathsf{H})$.

(c) $(\mathfrak{B}_2(\mathsf{H}), (\ |\ )_2)$ *is a Hilbert space.*
(d) $\mathfrak{B}_2(\mathsf{H}) \subset \mathfrak{B}_\infty(\mathsf{H})$. *More precisely,* $A \in \mathfrak{B}_2(\mathsf{H})$ *iff* A *is compact and the set of positive numbers* $\{m_\lambda \lambda^2\}_{\lambda \in sing(A)}$ *(*m_λ *is the multiplicity of* λ*) has finite sum. In this case:*

$$||A||_2 = \sqrt{\sum_{\lambda \in sing(A)} m_\lambda \lambda^2}\,. \tag{4.26}$$

Proof. (a) Obviously $\mathfrak{B}_2(\mathsf{H})$ is closed under multiplication by a scalar. Let us show it is closed under sums. If $A, B \in \mathfrak{B}_2(\mathsf{H})$ and N is any basis:

$$\sum_{u \in N} ||(A+B)u||^2 \leq \sum_{u \in N} (||Au|| + ||Bu||)^2 \leq 2\left[\sum_{u \in N} ||Au||^2 + \sum_{u \in N} ||Bu||^2\right].$$

Thus $\mathfrak{B}_2(\mathsf{H})$ is a subspace in $\mathfrak{B}(\mathsf{H})$. A consequence of Proposition 4.23(b) is the closure under Hermitian conjugation, which proves (i). We prove (ii) and at the same time that $\mathfrak{B}_2(\mathsf{H})$ is closed under left and right composites with bounded operators. If $A \in \mathfrak{B}_2(\mathsf{H})$ and $B \in \mathfrak{B}(\mathsf{H})$:

$$\sum_{u \in N} ||BAu||^2 \leq \sum_{u \in N} ||B||^2 ||Au||^2 = ||B||^2 \sum_{u \in N} ||Au||^2 .$$

This shows $\mathfrak{B}_2(\mathsf{H})$ is closed under left composites and the second inequality in (ii). Closure under right composition follows from closure under Hermitian conjugation and left composites: $B^*A^* \in \mathfrak{B}_2(\mathsf{H})$ if $A \in \mathfrak{B}_2(\mathsf{H})$ and $B \in \mathfrak{B}(\mathsf{H})$, so $(B^*A^*)^* \in \mathfrak{B}_2(\mathsf{H})$, i.e. $AB \in \mathfrak{B}_2(\mathsf{H})$. From (i) we find that if $A \in \mathfrak{B}_2(\mathsf{H})$ and $B \in \mathfrak{B}(\mathsf{H})$, then $||AB||_2 = ||(AB)^*||_2 = ||B^*A^*||_2 \leq ||B^*||\,||A^*||_2 = ||B||\,||A||_2$, finishing part (ii). As for (iii) observe:

$$||A|| = \sup_{||x||=1} ||Ax|| = \sup_{||x||=1} (||Ax||^2)^{1/2} = \sup_{||x||=1} \left(\sum_{u \in N} |(u|Ax)|^2\right)^{1/2}$$

$$= \sup_{||x||=1} \left(\sum_{u \in N} |(A^*u|x)|^2\right)^{1/2},$$

where Theorem 3.26(d) was used for the basis N. Using Cauchy-Schwarz on the last term above gives:

$$||A|| \leq \sup_{||x||=1} \left(\sum_{u \in N} ||A^*u||^2 ||x||^2\right)^{1/2} = \left(\sum_{u \in N} ||A^*u||^2\right)^{1/2} = ||A^*||_2 = ||A||_2 .$$

(b) If $A, B \in \mathfrak{B}_2(\mathsf{H})$ and N is a basis, the number of non-zero Au and Bu, for $u \in N$, is at most countable by Definition 4.22 and Proposition 3.21(b). Since

$$|(Au|Bu)| \leq ||Au||\,||Bu|| \leq \frac{1}{2}(||Au||^2 + ||Bu||^2),$$

the number of non-vanishing $(Au|Bu)$, $u \in N$, is also countable at most, and the series of non-null $(Au|Bu)$ is absolutely convergent, so the order in (4.24) is irrelevant. In a moment we will show that the choice of basis is not important. First, though, notice that (4.25) holds trivially and $(\ |\ )_2$ satisfies the axioms of a semi-inner product, as is easy to check. Positive definiteness (axiom **H3**) follows directly from (iii), so $(\,|\,)_2$ is an inner product inducing $||\,||_2$. Therefore we have polarisation, for this formula holds for any scalar product:

$$4(A|B)_2 = ||A+B||_2^2 + ||A-B||_2^2 - i||A+iB||_2^2 + i||A-iB||_2^2 .$$

Since, by Proposition 4.23, the number on the right does not depend on any basis, neither will the left-hand side.
(c) We need only prove the space is complete. Take N a basis of H and $\{A_n\}_{n\in\mathbb{N}}$ a Cauchy sequence of HS operators with respect to $||\,||_2$. From part (iii) in (a) it is a Cauchy sequence also in the uniform topology, and since $\mathfrak{B}(\mathsf{H})$ is complete by Theorem 2.41, there will be $A \in \mathfrak{B}(\mathsf{H})$ with $||A-A_n|| \to 0$, $n \to +\infty$. Cauchy's property asserts that however we take $\varepsilon > 0$ there is N_ε such that $||A_n - A_m||_2^2 \leq \varepsilon^2$ if $n,m > N_\varepsilon$. By definition of $||\,||_2$, for the same ε we will also have that for any finite subset $I \subset N$:

$$\sum_{u\in I} ||(A_n - A_m)u||^2 \leq ||A_n - A_m||_2^2 \leq \varepsilon^2$$

whenever $n,m > N_\varepsilon$. Passing to the limit as $m \to +\infty$, we find

$$\sum_{u\in I} ||(A_n - A)u||^2 \leq \varepsilon^2 ,$$

for any finite $I \subset N$ if $n > N_\varepsilon$. Overall, given that I is arbitrary,

$$||A - A_n||_2 \leq \varepsilon \quad \text{se } n > N_\varepsilon . \tag{4.27}$$

In particular, then, $A - A_n \in \mathfrak{B}_2(\mathsf{H})$, and so:

$$A = A_n + (A - A_n) \in \mathfrak{B}_2(\mathsf{H}) .$$

Furthermore, $\varepsilon > 0$ was also arbitrary in (4.27), so A_n tends to A with respect to $||\ ||_2$. Therefore every Cauchy sequence for $||\,||_2$ converges inside $\mathfrak{B}_2(\mathsf{H})$, making the latter complete.
(d) Let $A \in \mathfrak{B}_2(\mathsf{H})$: we claim it is compact and it fulfills (4.26). Take a basis N. Then $\sum_{u\in N} ||Au||^2 < +\infty$, where at most countably many summands do not vanish, and the sum can be written as series or finite sum by taking only the u_n for which $||Au_n||^2 > 0$. Therefore, for any $\varepsilon > 0$ there exists N_ε such that

$$\sum_{n=N_\varepsilon}^{+\infty} ||Au_n||^2 \leq \varepsilon^2 .$$

The same property can be expressed in terms of N: for any $\varepsilon > 0$ there is a finite subset $I_\varepsilon \subset N$ such that

$$\sum_{u\in N\setminus I_\varepsilon} ||Au||^2 \leq \varepsilon^2 .$$

Let then A_{I_ε} be the operator defined by: $A_{I_\varepsilon}u := Au$ if $u \in I_\varepsilon$ and $A_{I_\varepsilon}u := 0$ if $u \in N \setminus I_\varepsilon$. The range of A_{I_ε} is spanned by Au with $u \in I_\varepsilon$, because these are finite in number, and A_{I_ε} is bounded and compact (Example 4.16(1)). By construction $||A - A_{I_\varepsilon}||_2$ exists and equals:

$$||A - A_{I_\varepsilon}||_2^2 = \sum_{u \in N} ||(A - A_{I_\varepsilon})u||^2 = \sum_{u \in N \setminus I_\varepsilon} ||Au||^2 ;$$

by (a)(iii) therefore,

$$||A - A_{I_\varepsilon}|| \leq ||A - A_{I_\varepsilon}||_2 = \left(\sum_{u \in N \setminus I_\varepsilon} ||Au||^2 \right)^{1/2} \leq \varepsilon .$$

Thus A is a limit point in the space of compact operators in the uniform topology. As the ideal of compact operators is uniformly closed (Theorem 4.15(c)), the proven result shows A is compact. Now let us prove (4.26): consider the positive compact operator $|A| = \sqrt{A^*A}$ and let $\{u_{\lambda,i}\}_{\lambda \in sing(A), i=\{1,2,\ldots,m_\lambda\}}$ be a basis of $Ker(|A|)^\perp$, built as in Theorem 4.21. We may complete it to a basis of the Hilbert space by adding a basis for $Ker(|A|)$; the latter is $Ker(A)$ by Remark 3.70. (Using the orthogonal splitting $\mathsf{H} = Ker(|A|) \oplus Ker(|A|)^\perp$, if $\{u_i\}$ is a basis for the closed subspace $Ker(|A|)$ and $\{v_j\}$ a basis for the closed $Ker(|A|)^\perp$, the orthonormal system $N := \{u_i\} \cup \{v_j\}$ is a basis of H, since $x \in \mathsf{H}$ orthogonal to N implies $x = 0$.) Using that basis to write $||A||_2$:

$$||A||_2^2 = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} (Au_{\lambda,i}|Au_{\lambda,i}) = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} (A^*Au_{\lambda,i}|u_{\lambda,i}) = \sum_{\lambda \in sing(A)} m_\lambda \lambda^2 , \tag{4.28}$$

where the basis of $Ker(A)$, by construction, does not contribute, $|A|u_{\lambda,i} = \sqrt{A^*A}u_{\lambda,i} = \lambda u_{\lambda,i}$ and $(u_{\lambda,i}|u_{\lambda',j}) = \delta_{\lambda\lambda'}\delta_{ij}$.

If, conversely, A is compact and $\{m_\lambda \lambda^2\}_{\lambda \in sing(A)}$ has finite sum, then (4.28) implies $||A||_2 < +\infty$, so $A \in \mathfrak{B}_2(\mathsf{H})$. □

Examples 4.25. (1) Let us go back to example (4) in 4.16. Consider the operators:

$$T_K : L^2(\mathsf{X}, \mu) \to L^2(\mathsf{X}, \mu)$$

induced by integral kernels $K \in L^2(\mathsf{X} \times \mathsf{X}, \mu \otimes \mu)$ (μ a σ-finite separable measure),

$$(T_K f)(x) := \int_{\mathsf{X}} K(x,y) f(y) d\mu(y) \quad \text{for any } f \in L^2(\mathsf{X}, \mu) .$$

We did prove these operators are compact. Now we show they are Hilbert–Schmidt operators.

Using the same definition of the mentioned example, if $f \in L^2(\mathsf{X}, \mu)$ we saw (cf. part (3) in example (4)) that for any $x \in \mathsf{X}$

$$F(x) = \int_{\mathsf{X}} |K(x,y)| \, |f(y)| d\mu(y) < +\infty .$$

Since $F \in L^2(\mathsf{X}, \mu)$, for any $g \in L^2(\mathsf{X}, \mu)$ the map $x \mapsto g(x)F(x)$ is integrable (so we can define the inner product of g and F). The Fubini–Tonelli theorem guarantees the map $(x,y) \mapsto g(x)K(x,y)f(y)$ is in $L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$ and that

$$\int_{\mathsf{X}\times\mathsf{X}} \overline{g(x)}K(x,y)f(y)\,d\mu(x)\otimes d\mu(y) = \int_{\mathsf{X}} d\mu(x)\overline{g(x)}\int_{\mathsf{X}} K(x,y)f(y)d\mu(y) = (g|T_k f)\,. \tag{4.29}$$

So let us consider a basis of $L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$ of type $\{u_i\cdot\overline{u_j}\}_{i,j}$, where $\{u_k\}_k$ is a basis for $L^2(\mathsf{X}, \mu)$ (and so for $\{\overline{u_k}\}_k$ as well, as is easy to prove). As $K \in L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$, we have an expansion:

$$K = \sum_{i,j}\alpha_{ij}u_i\cdot\overline{u_j}\,, \tag{4.30}$$

and then

$$||K||^2_{L^2} = \sum_{i,j}|\alpha_{ij}|^2 < +\infty\,. \tag{4.31}$$

On the other hand, (4.29) and (4.30) imply

$$(u_i|T_K u_j) = \int_{\mathsf{X}\times\mathsf{X}} \overline{u_i(x)}u_j(y)K(x,y)\,d\mu(x)\otimes d\mu(y) = (u_i\cdot\overline{u_j}|K) = \alpha_{ij}\,,$$

hence (4.31) rephrases as:

$$||K||^2_{L^2} = \sum_{i,j}|(u_i|T_K u_j)|^2 < +\infty\,.$$

By definition T_K is thus a HS operator, and

$$||T_K||_2 = ||K||_{L^2}\,. \tag{4.32}$$

(2) It is not so hard to prove that if $\mathsf{H} = L^2(\mathsf{X}, \mu)$ with μ σ-finite and separable, $\mathfrak{B}_2(\mathsf{H})$ consists precisely of the operators T_K with $K \in L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$, so that the map identifying K with T_K is a Hilbert space isomorphism between $L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$ and $\mathfrak{B}_2(\mathsf{H})$. To see that we take $T \in \mathfrak{B}_2(\mathsf{H})$ and will exhibit $K \in L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$ for which $T = T_K$. Givan any basis $\{u_n\}_{n\in\mathbb{N}}$ of $L^2(\mathsf{X}, \mu)$ we have $\sum_{n\in\mathbb{N}}||Tu_n||^2 < +\infty$. Consequently, by expanding Tu_n in $\{u_n\}_{n\in\mathbb{N}}$ we obtain:

$$+\infty > \sum_{n\in\mathbb{N}}||Tu_n||^2 = \sum_{m\in\mathbb{N}}\sum_{n\in\mathbb{N}}|(u_m|Tu_n)|^2\,.$$

Interpreting the series as integrals on $\{u_n\}_{n\in\mathbb{N}}$, and applying Fubini–Tonelli, we conclude $\sum_{(n,m)\in\mathbb{N}^2}|(u_m|Tu_n)|^2 < +\infty$, so the integral HS operator T_K with integral kernel $K \in L^2(\mathsf{X}\times\mathsf{X}, \mu\otimes\mu)$:

$$K := \sum_{(n,m)\in\mathbb{N}^2}(u_m|Tu_n)u_m\cdot\overline{u_n}$$

is well defined. At the same time, the results of the previous example tell that, by construction:

$$(u_m|T_K u_n) = \int_{\mathsf{X}} d\mu(x)\overline{u_m(x)}\int_{\mathsf{X}} d\mu(y)K(x,y)u_n(y) = (u_m|Tu_n)$$

and so $T_K u_n = Tu_n$ for any $n \in \mathbb{N}$. By continuity $T_K = T$ follows immediately.

(3) Consider the **Volterra equation** in the unknown function $f \in L^2([0,1],dx)$:

$$f(x) = \rho \int_0^x f(y)dy + g(x)\,, \quad \text{with given } g \in L^2([0,1],dx) \text{ and } \rho \in \mathbb{C}\setminus\{0\}. \quad (4.33)$$

Above, dx is Lebesgue's measure, and the integral exists because, for any given $x \in [0,1]$, we can view it as inner product of f and the map $[0,1] \ni y \mapsto \theta(x-y)$. As such we write it as follows:

$$\int_0^x f(y)dy = \int_0^1 \theta(x-y)f(y)dy\,,$$

where $\theta(u) = 1$ if $u \geq 0$ and $\theta(u) = 0$ if $u < 0$. Clearly $(x,y) \mapsto \theta(x-y)$ is also in $L^2([0,1]^2, dx \otimes dy)$, so the equation can be rephrased using a Hilbert–Schmidt operator T:

$$f = \rho T f + g\,, \quad \text{with given } g \in L^2([0,1],dx) \text{ and } \rho \in \mathbb{C}\setminus\{0\} \quad (4.34)$$

where we defined the **Volterra operator**:

$$(Tg)(x) := \int_0^x g(y)dy \quad g \in L^2([0,1],dx)\,. \quad (4.35)$$

Volterra operators, and the associated equations, are more generally defined as:

$$(T_V f)(x) := \int_0^x V(x,y)g(y)dy\,,$$

for some suitably regular $V : [0,1]^2 \to \mathbb{R}$. We will study the simplest situation, given by (4.35). If the operator $(I - \rho T)$ is invertible, the solution to (4.34) reads:

$$f = (I - \rho T)^{-1} g\,. \quad (4.36)$$

Formally, using the geometric series we see that the (left and right) inverse to $I - \rho T$ is the sum of:

$$(I - \rho T)^{-1} = I + \sum_{n=0}^{+\infty} \rho^{n+1} T^{n+1}\,, \quad (4.37)$$

where the convergence is in the uniform topology. A sufficient condition for convergence is $||\rho T|| < 1$, proved in analogy to the geometric series. Yet we will look for a finer estimate. Use the norm of $\mathfrak{B}_2(L^2([0,1],dx))$ and recall part (iii) in Theorem 4.24(a). Moreover, if $||A_n|| \leq a_n \geq 0$ for any $A_n \in \mathfrak{B}(L^2([0,1],dx))$ where $\sum_{n=0}^{+\infty} a_n$ converges, then also $\sum_{n=0}^{+\infty} A_n$ converges in $\mathfrak{B}(L^2([0,1],dx))$. The proof of the latter fact is similar to that of the Weierstrass M-test in elementary calculus. A direct computation with (4.35) shows that if $n \geq 1$:

$$(T^{n+1}g)(x) = \int_0^x \frac{(x-y)^n}{n!} g(y)dy\,,$$

so $T^n \in \mathfrak{B}_2(L^2([0,1],dx))$ and:

$$||T^n|| \leq ||T^n||_2 = \sqrt{\int_{[0,1]^2} |\theta(x-y)|^2 \left|\frac{(x-y)^{n-1}}{(n-1)!}\right|^2 dx \otimes dy} \leq \frac{2^{n-1}}{(n-1)!}.$$

Since the series of terms $\frac{\rho^n 2^n}{n!}$ converges, for any $\rho \neq 0$ the operator $(I-\rho T)^{-1}$ exists in $\mathfrak{B}(L^2([0,1],dx))$ and is given by the sum in the right-hand side of (4.37). Therefore (4.36) solves the initial Volterra equation. It is possible to make $(I-\rho T)^{-1}$ explicit

$$\begin{aligned}\left((I-\rho T)^{-1}g\right)(x) &= g(x) + \sum_{n=0}^{+\infty} \rho^{n+1}\left(T^{n+1}g\right)(x) \\ &= g(x) + \rho\sum_{n=0}^{+\infty}\int_0^x \frac{(\rho(x-y))^n}{n!} g(y)dy.\end{aligned}$$

The theorem of dominated convergence warrants we may swap sum and integral, so that Volterra's solution reads:

$$f(x) = \left((I-\rho T)^{-1}g\right)(x) = g(x) + \rho\int_0^x e^{\rho(x-y)} g(y)dy.$$

For these operations we used a notion of pointwise convergence other than the uniform operator convergence. That the above expression is indeed the explicit inverse to $I-\rho T$ can be checked by a direct computation, using (4.35) and integrating by parts with $g \in C([0,1])$. The result extends to $L^2([0,1],dx)$ because the operator with integral kernel $\theta(x-y)e^{\rho(x-y)}$ is bounded (HS) and $C([0,1])$ is dense in $L^2([0,1],dx)$. The inverse's uniqueness ends the proof.

(4) Take $L^2(\mathsf{X},\mu)$ with μ separable. An integral operator $T_K : L^2(\mathsf{X},\mu) \to L^2(\mathsf{X},\mu)$ given by the kernel:

$$K(x,y) = \sum_{k=1}^{N} p_k(x)q_k(y),$$

where $p_k, q_k \in L^2(\mathsf{X},\mu)$, $k = 1,2,3,\ldots,N$ are arbitrary maps and $N \in \mathbb{N}$ is chosen at random, is called **degenerate operator**. Degenerate operators forms a two-sided *-ideal $\mathfrak{B}_D(L^2(\mathsf{X},\mu))$ in $\mathfrak{B}(L^2(\mathsf{X},\mu))$ that is a subspace of $\mathfrak{B}_\infty(L^2(\mathsf{X},\mu))$ and $\mathfrak{B}_2(L^2(\mathsf{X},\mu))$. It can be proved rather easily that $\mathfrak{B}_D(L^2(\mathsf{X},\mu))$ is dense in $\mathfrak{B}_2(L^2(\mathsf{X},\mu))$ in the latter's norm topology. ∎

4.3.2 Integral kernels and Mercer's theorem

The content of Examples 4.25(1) and (2) can be subsumed in a theorem. The final assertion is easy and left as exercise.

Theorem 4.26. *If μ is a positive σ-additive and separable measure on X, the space $\mathfrak{B}_2(L^2(\mathsf{X},\mu))$ consists of the operators T_K:*

$$(T_K f)(x) := \int_{\mathsf{X}} K(x,y)f(y)dy, \quad \textit{for any } f \in L^2(\mathsf{X},\mu), \tag{4.38}$$

where $K \in L^2(\mathsf{X} \times \mathsf{X}, \mu \otimes \mu)$. Moreover:

$$||T_K||_2 = ||K||_{L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)}\,.$$

In particular, if $T \in \mathfrak{B}_2(L^2(\mathsf{X}, \mu))$ and U is a basis of $L^2(\mathsf{X}, \mu)$, then $T = T_K$ for the kernel

$$K = \sum_{u,v \in U \times U} (u|Tv)u \cdot \overline{v} \tag{4.39}$$

and the convergence is in $L^2(\mathsf{X} \times \mathsf{X}, \mu \otimes \mu)$.

The map $L^2(\mathsf{X} \times \mathsf{X}, \mu \otimes \mu) \ni K \mapsto T_K \in \mathfrak{B}_2(L^2(\mathsf{X}, \mu))$ is an isomorphism of Hilbert spaces.

Mercer's theorem [RiNa53] is sometimes useful when dealing with $\mathsf{X} \subset \mathbb{R}^n$ compact and Lebesgue's measure μ. In such a case, if K is continuous and T_K is positive the convergence of (4.39) is in $||\ ||_\infty$, provided one uses a basis of eigenvectors for T_K. We state and prove the theorem in a slightly more general situation, so to include Lebesgue's measures on compact sets in $\mathbb{R}^n$.

Theorem 4.27 (Mercer). *Let μ be a positive, separable Borel measure on a compact Haudorff space X such that $\mu(\mathsf{X}) < +\infty$ and $\mu(A) > 0$ for any open set $A \neq \varnothing$. Assume $K : \mathsf{X} \times \mathsf{X} \to \mathbb{C}$ is continuous. If T_K in (4.38) is positive, i.e. $(f|T_K f) \geq 0$ for $f \in L^2(\mathsf{X}, \mu)$, then*

$$K(x,y) = \sum_{\lambda \in \sigma(T_K)} \sum_{i=1}^{m_\lambda} \lambda u_{\lambda,i}(x)\overline{u_{\lambda,i}(y)}\,, \tag{4.40}$$

where the series converges on $\mathsf{X} \times \mathsf{X}$ in norm $||\ ||_\infty$. The number m_λ indicates the dimension (finite if $\lambda \neq 0$) of the λ-eigenspace of T_K, and $\{u_{\lambda,i}\}_{\lambda \in \sigma_p(T_k), i=1,\ldots,m_\lambda}$ is a basis of (continuous if $\lambda \neq 0$) eigenvectors of T_K.

Proof. For simplicity let us relabel eigenvectors as u_j with $j \in \mathbb{N}$, and call λ_j the corresponding eigenvalues (it may happen that $\lambda_j = \lambda_k$ if $j \neq k$). To begin with, notice the eigenvectors with $\lambda \neq 0$ are continuous, by the Cauchy-Schwarz inequality:

$$|u_j(x) - u_j(x')|^2 \leq \int_{\mathsf{X}} |K(x,y) - K(x',y)|^2 d\mu(y) \int_{\mathsf{X}} |u_j(y)|^2 d\mu(y) \to 0 \quad \text{as } x \to x'.$$

We used dominated convergence for the first integral on the right, since K is integrable on X, as $\mu(\mathsf{X})$ is finite, and also continuous on the compact set $\mathsf{X} \times \mathsf{X}$ and thus $|K(x,\cdot) - K(x',\cdot)|^2$ is bounded, uniformly in x, x', by some constant map $C \geq 0$. So take the continuous maps

$$K_n(x,y) := K(x,y) - \sum_{j=0}^{n} \lambda_j u_j(x)\overline{u_j(y)} = \sum_{j=n+1}^{+\infty} \lambda_j u_j(x)\overline{u_j(y)}\,,$$

where the last series converges in $L^2(\mathsf{X} \times \mathsf{X}, \mu(x) \otimes \mu(y))$ by Theorem 4.26. Note $\lambda_j \geq 0$, because $0 \leq (u_j|T_K u_j) = \lambda_j$. In the topology of $L^2(\mathsf{X} \times \mathsf{X}, \mu(x) \otimes \mu(y))$ we have:

$$K_n(x,y) = \sum_{j=n+1}^{+\infty} \lambda_j u_j(x)\overline{u_j(y)}\,,$$

so if $f \in L^2(\mathsf{X},\mu)$

$$\int_{\mathsf{X}}\int_{\mathsf{X}} K_n(x,y)f(y)\overline{f(x)}\mu(x)\mu(y) = \sum_{j=n+1}^{+\infty} \lambda_j (u_j|f)(f|u_j) \geq 0\,.$$

We claim this fact implies $K_n(x,x) \geq 0$. If there were $x_0 \in \mathsf{X}$ with $K_n(x_0,x_0) < 0$, as K_n is continuous we would be able to find an open neighbourhood (x_0,x_0) where $K_n(x,y) \leq K_n(x_0,x_0)+\varepsilon < 0$. Since $\mathsf{X}\times\mathsf{X}$ has the product topology, we could choose the neighbourhood to be $B_{x_0}\times B_{x_0}$ where B_{x_0} is an open neighbourhood of x_0. By Urysohn's lemma (Theorem 1.24) we could find a continuous map f with support in B_{x_0} such that $0 \leq f \leq 1$ and $f(x_0) = 1$ (note $\{x_0\}$ is compact because closed inside a compact space). Then a contradiction would ensue, since $\mu(B_{x_0}) > 0$ by assumption:

$$\begin{aligned}
(f|T_K f) &= \int_{\mathsf{X}}\int_{\mathsf{X}} K_n(x,y)f(y)\overline{f(x)}d\mu(x)d\mu(y)\\
&= \int_{B_{x_0}}\int_{B_{x_0}} K_n(x,y)f(y)\overline{f(x)}d\mu(x)d\mu(y)\\
&\leq \left(\int_{B_{x_0}} f(x)d\mu(x)\right)^2 (K_n(x_0,x_0)+\varepsilon) \leq \left(\int_{B_{x_0}} 1d\mu(x)\right)^2 (K_n(x_0,x_0)+\varepsilon)\\
&= \mu(B_{x_0})(K_n(x_0,x_0)+\varepsilon) < 0\,.
\end{aligned}$$

Thus if $n = 0,1,2,\ldots$

$$0 \leq K_n(x,x) = K(x,y) - \sum_{j=0}^{n} \lambda_j u_j(x)\overline{u_j(x)}\,,$$

and the positive-term series $\sum_{j=0}^{+\infty}\lambda_j u_j(x)\overline{u_j(x)}$ converges, with sum bounded by $K(x,x)$. Hence the series $\sum_{j=0}^{+\infty}\lambda_j u_j(x)\overline{u_j(y)}$ converges for any x, uniformly in y. In fact if $M = \max_{x\in\mathsf{X}} K(x,x)$, from Cauchy-Schwarz:

$$\left|\sum_{j=m}^{n}\lambda_j u_j(x)\overline{u_j(y)}\right|^2 \leq \sum_{j=m}^{n}\lambda_j|u_j(x)|^2\sum_{j=m}^{n}\lambda_j|u_j(y)|^2 \leq M\sum_{j=m}^{n}\lambda_j|u_j(x)|^2\,.$$

Therefore $B(x,y) := \sum_{j=0}^{+\infty}\lambda_j u_j(x)\overline{u_j(y)}$ is continuous in y for any given x. By dominated convergence, for any continuous $f : \mathsf{X} \to \mathbb{C}$ and any given x

$$\int_{\mathsf{X}} B(x,y)f(y)d\mu(y) = \sum_{j=0}^{+\infty}\lambda_j u_j(x)\int_{\mathsf{X}}\overline{u_j(y)}f(y)d\mu(y) \tag{4.41}$$

by virtue of the above series' uniform convergence on X (of finite measure). But we also know the series on the right converges to $T_K f$ in $L^2(\mathsf{X}, d\mu(x))$: we claim it converges to $(T_K f)(x)$ *also pointwise* at x. Since K is continuous on the compact X of finite measure, an obvious consequence of dominated convergence shows $\mathsf{X} \ni x \mapsto \int_{\mathsf{X}}|K(x,y)|^2 d\mu(y)$ is continuous, so there is a constant C such that:

$$\int_{\mathsf{X}}|K(x,y)|^2 d\mu(y) < C^2 \quad \text{for any } x \in \mathsf{X}\,.$$

Hence if $f_n \to f$ in $L^2(\mathsf{X},\mu)$, then $T_K f_n \to T_K f$ in norm $|| \ ||_\infty$:

$$|(Tf)(x) - (Tf_n)(x)|^2 = \left| \int_{\mathsf{X}} K(x,y)(f(y) - f_n(y)) d\mu(y) \right|^2$$
$$\leq \int_{\mathsf{X}} |K(x,y)|^2 d\mu(y) \int_{\mathsf{X}} |f(y) - f_n(y)|^2 d\mu(y) \leq C^2 \mu(\mathsf{X}) ||f - f_n||_\infty^2 .$$

Now if we decompose f in (4.41) using the eigenvector basis of T_K,

$$f(x) = \sum_{j=0}^{+\infty} u_j(x) \int_{\mathsf{X}} \overline{u_j(y)} f(y) d\mu(y),$$

we obtain a converging expansion in $L^2(\mathsf{X}, d\mu(x))$, so applying T_K must give a uniformly converging series at $x \in \mathsf{X}$. Therefore the last series in (4.41) converges pointwise (and uniformly) to $(T_K f)(x)$ for any $x \in \mathsf{X}$. Comparing with the left-hand side of (4.41) and recalling that $(T_K f)(x) = \int K(x,y) f(y) d\mu(y)$ we finally get

$$\int_{\mathsf{X}} (B(x,y) - K(x,y)) f(y) d\mu(y) = 0 .$$

Choosing $f(y) := \overline{B(x,y) - K(x,y)}$, for any given $x \in \mathsf{X}$, allows to conclude $B(x,y) = K(x,y)$ by using Proposition 1.70 suitably. So

$$K(x,x) = B(x,x) = \sum_{j=0}^{+\infty} \lambda_j |u_j(x)|^2 .$$

The terms are continuous, non-negative and the sum is a continuous map, so Dini's Theorem 2.19 forces uniform convergence. From Cauchy-Schwarz we deduce that $\sum_{j=0}^{+\infty} \lambda_j u_j(x) \overline{u_j(y)}$ converges uniformly, jointly in x, y, to some $K'(x,y)$. $\mathsf{X} \times \mathsf{X}$ has finite measure, so the convergence is in $L^2(\mathsf{X} \times \mathsf{X}, \mu \otimes \mu)$. But we know the series converges to $K(x,y)$ in that topology. Hence the series converges uniformly to $K(x,y)$ and the proof ends. □

Remarks 4.28. The theorem is still valid if T_K has a finite number of negative eigenvalues, as is easy to prove. ■

4.4 Trace-class (or nuclear) operators

In this part we introduce operators of *trace class*, also known as *nuclear operators*. We shall follow the approach of [Mar82] essentially (a different one would be that of [Pru81]).

4.4.1 General properties

Proposition 4.29. *Let* H *be a Hilbert space and* $A \in \mathfrak{B}(\mathsf{H})$. *The following three facts are equivalent.*

(a) *There exists a basis N of H such that $\{(u||A|u)\}_{u\in N}$ has finite sum:*

$$\sum_{u\in N}(u||A|u) < +\infty\,.$$

(a)' *$\sqrt{|A|}$ is a Hilbert–Schmidt operator.*
(b) *A is compact and the indexed set $\{m_\lambda \lambda\}_{\lambda\in sing(A)}$, where m_λ is the multiplicity of λ, has finite sum.*

Proof. Statement (a)' is a mere translation of (a), for $\sqrt{|A|}\sqrt{|A|} = |A|$, so (a) and (a)' are equivalent. We show (a) implies (b). Any HS operator, in particular $\sqrt{|A|}$, is compact (Theorem 4.24(d)), the product of compact operators, e.g. $|A| = (\sqrt{|A|})^2$, is compact (Theorem 4.15(b)), and $|A| = (\sqrt{|A|})^2$ is compact if and only if A is (Proposition 4.14); as a consequence of all this, A is compact. Let us take a basis of H made of eigenvectors of $|A|$: $u_{\lambda,i}$, $i = 1,\ldots,m_\lambda$ ($m_\lambda = +\infty$ possibly, for $\lambda = 0$ only) and $|A|u_{\lambda,i} = \lambda u_{\lambda,i}$. In such basis:

$$\left|\left|\sqrt{|A|}\right|\right|_2^2 = \sum_{\lambda,i}\left(\sqrt{|A|}u_{\lambda,i}\middle|\sqrt{|A|}u_{\lambda,i}\right) = \sum_{\lambda,i}\left(u_{\lambda,i}\middle|(\sqrt{|A|})^2 u_{\lambda,i}\right)$$
$$= \sum_{\lambda,i}(u_{\lambda,i}||A|u_{\lambda,i}) = \sum_\lambda m_\lambda \lambda\,.$$

So $\{m_\lambda \lambda\}_{\lambda\in sing(A)}$ has finite sum because $\left|\left|\sqrt{|A|}\right|\right|_2^2 < +\infty$ by assumption. Conversely, it is obvious that (b) implies (a)' by proceeding backwards in the argument and computing $\left|\left|\sqrt{|A|}\right|\right|_2^2$ in a basis of eigenvectors of $|A|$. $\square$

Definition 4.30. *Let H be a Hilbert space. $A \in \mathfrak{B}(\mathsf{H})$ is called operator of* **trace class**, *or equivalently* **nuclear operator**, *if it satisfies either of (a), (a)' or (b) in Proposition 4.29. The set of trace-class operators on H will be denoted by $\mathfrak{B}_1(\mathsf{H})$. In the notation of Proposition 4.29, if $A \in \mathfrak{B}_1(\mathsf{H})$,*

$$||A||_1 := \left|\left|\sqrt{|A|}\right|\right|_2^2 = \sum_{\lambda\in sing(A)} m_\lambda \lambda\,. \tag{4.42}$$

Remark 4.31. (1) The name "trace class" evidently has its origin in the following observation. For an operator A of trace class, the real number $||A||_1$ generalises to infinite dimensions the notion of trace of the matrix corresponding to $|A|$ (not A). As a matter of fact, the analogies do not end here, as we shall briefly see.
(2) The following inclusions hold:

$$\mathfrak{B}_1(\mathsf{H}) \subset \mathfrak{B}_2(\mathsf{H}) \subset \mathfrak{B}_\infty(\mathsf{H}) \subset \mathfrak{B}(\mathsf{H})\,.$$

The only relation we have not yet proved is the first one. To this end, if $A \in \mathfrak{B}_1(\mathsf{H})$, by definition $\sqrt{|A|} \in \mathfrak{B}_2(\mathsf{H})$, so $|A| = \sqrt{|A|}\sqrt{|A|}$ is HS by Theorem 4.24(a); from the polar decomposition $A = U|A|$, $U \in \mathfrak{B}(\mathsf{H})$, we have $A \in \mathfrak{B}_2(\mathsf{H})$ by Theorem 4.24(a).

(3) Each of the above sets is a subspace in the vector space of bounded operators, and also a two-sided *-ideal (for trace-class operators we will prove it in a moment); at last, each subspace is a Hilbert or Banach space with respect to a natural structure: compact operators are closed in $\mathfrak{B}(\mathsf{H})$ in uniform topology, so they form a Banach space for the operator norm, HS operators form a Hilbert space with Hilbert–Schmidt inner product, and trace-class operators form a Banach space with norm $|| \ ||_1$, as we will explain later. ■

Before we extend the notion of trace to the infinite-dimensional case, let us review the key features of nuclear, or trace-class, operators.

Theorem 4.32. *Let* H *be a Hilbert space. Nuclear operators in* H *enjoy the following properties.*

(a) *If* $A \in \mathfrak{B}_1(\mathsf{H})$ *there exist two operators* $B, C \in \mathfrak{B}_2(\mathsf{H})$ *such that* $A = BC$. *Conversely, if* $B, C \in \mathfrak{B}_2(\mathsf{H})$ *then* $BC \in \mathfrak{B}_1(\mathsf{H})$ *and:*

$$||BC||_1 \leq ||B||_2 \, ||C||_2 \, . \tag{4.43}$$

(b) $\mathfrak{B}_1(\mathsf{H})$ *is a subspace of* $\mathfrak{B}(\mathsf{H})$, *and actually a two-sided* **-ideal. Moreover:*
 (i) $||AB||_1 \leq ||B|| \, ||A||_1$ *and* $||BA||_1 \leq ||B|| \, ||A||_1$ *for any* $A \in \mathfrak{B}_1(\mathsf{H})$ *and* $B \in \mathfrak{B}(\mathsf{H})$*;*
 (ii) $||A||_1 = ||A^*||_1$ *for any* $A \in \mathfrak{B}_1(\mathsf{H})$*;*
(c) $|| \ ||_1$ *is a norm on* $\mathfrak{B}_1(\mathsf{H})$.

Remarks 4.33. It can be proved $\mathfrak{B}_1(\mathsf{H})$ is a Banach space with norm $|| \ ||_1$ [Sch60, BiSo87]. ■

Proof of Theorem 4.32 (part (b)(ii) is deferred to Proposition 4.36. (a) If A is of trace class, the polar decomposition $A = U|A|$ tells B and C can be taken to be $B := U\sqrt{|A|}$ and $C := \sqrt{|A|}$. By definition of trace-class operator $\sqrt{|A|}$ is a HS operator, so C is HS. Such is also B, for $U \in \mathfrak{B}(\mathsf{H})$ and Hilbert–Schmidt operators form a two-sided ideal in $\mathfrak{B}(\mathsf{H})$, by Theorem 4.24. Let now B, C be HS operators, and let us show $A := BC$ is of trace class. A is compact by Theorems 4.24(d) and 4.15(b); thus we need only show $\sum_{\lambda \in sing(A)} m_\lambda \lambda < +\infty$. If $BC = 0$, the claim is obvious. Assume $BC \neq 0$ and expand the compact operator BC as in Theorem 4.21:

$$A = BC = \sum_{\lambda \in sing(A)} \sum_{i=1}^{m_\lambda} \lambda (u_{\lambda,i}| \ ) v_{\lambda,i} \, .$$

Lest the notation become too heavy, set

$$\Gamma := \{(\lambda, i) | \lambda \in sing(A), \ i = 1, 2, \ldots, m_\lambda\}$$

and suppose λ_j is the first element in the pair $j = (\lambda, i)$. Then, clearly,

$$\sum_{j \in \Gamma} \lambda_j = \sum_{\lambda \in singA} m_\lambda \lambda \, .$$

From the polar decomposition theorem $A = U|A|$, with $U^*U = I$ on the range of $|A|$; since $v_j = Uu_j$ implies $U^*v_j = u_j$, we have:

$$(v_j|BCu_j) = (v_j|Au_j) = (v_j|U|A|u_j) = \lambda_j(v_j|Uu_j) = \lambda_j(U^*v_j|u_j) = \lambda_j(u_j|u_j) = \lambda_j\,.$$

If $S \subset \Gamma$ is finite:

$$\sum_{j\in S}\lambda_j = \sum_{j\in S}(v_j|BCu_j) = \sum_{j\in S}(B^*v_j|Cu_j)$$
$$\leq \sum_{j\in S}||B^*v_j||\,||Cu_j|| \leq \sqrt{\sum_{j\in S}||B^*v_j||^2}\,\sqrt{\sum_{j\in S}||Cu_j||^2}\,.$$

As the orthonormal systems $u_j = u_{\lambda,i}$ and $v_j = v_{\lambda,i}$ can be both completed to give bases of H, the final term in the chain of inequalities above is smaller than

$$||B^*||_2\,||C||_2 = ||B||_2\,||C||_2\,.$$

Taking the supremum over all finite S we conclude

$$||BC||_1 = \sum_{\lambda\in sing(A)} m_\lambda \lambda \leq ||B||_2\,||C||_2$$

and in particular $A = BC \in \mathfrak{B}_1(\mathsf{H})$.
(b)–(c) The closure of $\mathfrak{B}_1(\mathsf{H})$ under inner product is immediate from the definition itself. Let us show $\mathfrak{B}_1(\mathsf{H})$ is closed under sums. Take $A, B \in \mathfrak{B}_1(\mathsf{H})$. If $A+B = 0$, $A+B$ is clearly nuclear. So assume $A+B \neq 0$ (compact anyway). Decompose polarly: $A = U|A|$, $B = V|B|$, $A+B = W|A+B|$. Using the usual decomposition on singular values like in part (a), we find:

$$A+B = \sum_{\beta\in sing(A+B)}\sum_{i=1}^{m_\beta}\beta(u_{\beta,i}|\ )v_{\beta,i}\,.$$

If $\Gamma := \{(\beta,i)|\beta \in sing(A+B),\ i = 1,2,\ldots,m_\beta\}$, $S \subset \Gamma$ is finite and β_j is the first element in the pair $j \in \Gamma$, we have:

$$\sum_{j\in S}\beta_j = \sum_{j\in S}(v_j|(A+B)u_j) = \sum_{j\in S}(v_j|Au_j) + \sum_{j\in S}(v_j|Bu_j)\,.$$

We can rewrite this as follows:

$$\sum_{j\in S}\beta_j = \sum_{j\in S}(\sqrt{|A|}U^*v_j|\sqrt{|A|}u_j) + \sum_{j\in S}(\sqrt{|B|}V^*v_j|\sqrt{|B|}u_j)\,.$$

Proceeding as in (a) gives:

$$\sum_{j\in S}\beta_j \leq ||\sqrt{|A|}U^*||_2\,||\sqrt{|A|}||_2 + ||\sqrt{|B|}V^*||_2\,||\sqrt{|B|}||_2 \leq ||\sqrt{|A|}||_2^2 + ||\sqrt{|B|}||_2^2$$

(in the final passage we use the inequality $||\sqrt{|A|}U^*||_2 \leq ||\sqrt{|A|}||_2\,||U^*||$ (Theorem 4.24, part (a) (ii)), since $\sqrt{|A|}$ is Hilbert–Schmidt; furthermore, it is easy to

see $||U^*|| \leq 1$, because it is isometric on $Ker(|A|)^\perp$ and vanishes on $Ker(|A|)$).
Eventually note:

$$|| \sqrt{|A|}||_2^2 + || \sqrt{|B|}||_2^2 = ||A||_1 + ||B||_1 .$$

What we have proved is $A+B \in \mathfrak{B}_1(\mathsf{H})$ and that the triangle inequality

$$||A+B||_1 \leq ||A||_1 + ||B||_1$$

holds on $\mathfrak{B}_1(\mathsf{H})$. This turns $|| \ ||_1$ into a seminorm. It is indeed a norm, for $||A||_1 = 0$ implies the eigenvalues of $|A|$ are all zero; by compactness $|A| = 0$, from (b) in Theorem 4.17(6). The polar decomposition of $A = U|A|$ forces $A = 0$. At this stage we have proved $\mathfrak{B}_1(\mathsf{H})$ is a subspace of $\mathfrak{B}(\mathsf{H})$ and $|| \ ||_1$ is a norm. Let us explain that $\mathfrak{B}_1(\mathsf{H})$ is closed under composition, on the left and on the right, with bounded operators. Take $A \in \mathfrak{B}_1(\mathsf{H})$, $B \in \mathfrak{B}(\mathsf{H})$ and write $A = U|A|$. Then $BA = (BU\sqrt{|A|})\sqrt{|A|}$, where the factors are HS operators, so by part (a) $BA \in \mathfrak{B}_1(\mathsf{H})$. Using Theorem 4.24(a)(ii), equation (4.42) and part (a):

$$\begin{aligned} ||BA||_1 \leq ||BU\sqrt{|A|}||_2 \, ||\sqrt{|A|}||_2 \leq ||BU|| \, ||\sqrt{A}||_2 \, ||\sqrt{A}||_2 \leq ||B|| \, ||\sqrt{A}||_2^2 \\ = ||B|| \, ||A||_1 . \end{aligned}$$

Moreover $AB = (U\sqrt{|A|})\sqrt{|A|}B \in \mathfrak{B}_1(\mathsf{H})$ because both factors are HS and (a) holds. In a manner similar to part (a) we prove $||AB||_1 \leq ||B|| \, ||A||_1$. Statement (ii) in part (b) will be justified in the proof of Proposition 4.36. □

4.4.2 The notion of trace

To conclude we introduce the notion of trace of a nuclear operator and we show how the trace has the same formal properties of the trace of a matrix.

Proposition 4.34. *If* $(\mathsf{H}, (\ |\ ))$ *is a Hilbert space,* $A \in \mathfrak{B}_1(\mathsf{H})$ *and* N *is a basis of* H*, then*

$$trA := \sum_{u \in N} (u|Au) \tag{4.44}$$

is well defined, since the series on the right is finite or absolutely convergent. Moreover:

(a) trA *does not depend on the chosen basis.*
(b) *For any pair* (B, C) *of Hilbert–Schmidt operators such that* $A = BC$*:*

$$trA = (B^*|C)_2 . \tag{4.45}$$

(c) $|A| \in \mathfrak{B}_1(\mathsf{H})$ *and:*

$$||A||_1 = tr|A| . \tag{4.46}$$

Proof. (a) and (b) Any trace-class operator can be decomposed in the product of two HS operators as we saw in Theorem 4.32(a). We begin by noticing that if $A = BC$, B, C Hilbert–Schmidt, then

$$(B^*|C)_2 = \sum_{u \in N} (B^*u|Cu) = \sum_{u \in N} (u|BCu) = \sum_{u \in N} (u|Au) = trA .$$

This explains, apart from (4.45), that trA is well defined, being a Hilbert–Schmidt inner product, that in the infinite sum of (4.44) only countably many summands, at most, are non-zero, and that the sum reduces to a finite sum or to an absolutely convergent series, since $\sum_{u\in N}|(B^*u|Cu)| < +\infty$ by definition of Hilbert–Schmidt inner product. (This also shows $(B^*|C)_2 = (B'^*|C')_2$ if $BC = B'C'$, for B, B', C, C' are HS operators.) The result eventually proves the invariance of trA under changes of basis, because $(\ |\ )_2$ does not depend on the basis chosen.
(c) Firstly, by uniqueness of the square root, $|\,(|A|)\,| = |A|$. In fact $|\,(|A|)\,|$ is the only positive bounded operator whose square is $|A|^*|A| = |A|^2$, and $|A|$ is bounded, positive and squaring to $|A|^2$. Therefore, A being of trace class:

$$+\infty > \sum_{u\in N}(u||A|u) = \sum_{u\in N}(u||\,(|A|)\,|u)\,,$$

so Definition 4.30 implies $|A|$ itself is of trace class. Choosing a basis $\{u_{\lambda,i}\}$ of eigenvectors for $|A|$ we have:

$$tr|A| = \sum_{\lambda\in sing(A)}\sum_{i=1}^{m_\lambda}(u_{\lambda,i}|\ |A|u_{\lambda,i}) = \sum_{\lambda\in sing(A)}\sum_{i=1}^{m_\lambda}\lambda = \sum_{\lambda\in sing(A)} m_\lambda\lambda = ||A||_1\,.$$

The proof is thus finished. □

Definition 4.35. *Let* H *be a Hilbert space and* $A \in \mathfrak{B}_1(\mathsf{H})$*. The number* $trA \in \mathbb{C}$ *is the* **trace** *of the operator* A*.*

The next proposition states other useful properties of nuclear operators in Hilbert spaces: in particular – and precisely as in the finite-dimensional case – the trace is invariant under cyclic permutations. We remark that the operators of the statements below need not necessarily be all of trace class (an important fact in physical applications).

Proposition 4.36. *Let* H *be a Hilbert space. The trace enjoys the following properties.*

(a) *If* $A, B \in \mathfrak{B}_1(\mathsf{H})$ *and* $\alpha, \beta \in \mathbb{C}$*, then:*

$$trA^* = \overline{trA}\,, \tag{4.47}$$

$$tr(\alpha A + \beta B) = \alpha\, trA + \beta\, trB\,. \tag{4.48}$$

(b) *If* A *is of trace class and* $B \in \mathfrak{B}(\mathsf{H})$*, or* A *and* B *are both Hilbert–Schmidt operators, then*

$$trAB = trBA\,. \tag{4.49}$$

(c) *Let* $A_1, A_2, \dots A_n$ *be in* $\mathfrak{B}(\mathsf{H})$*. If one is of trace class, or two of them are Hilbert–Schmidt operators, then the trace is* **invariant under cyclic permutations***:*

$$tr\,(A_1A_2\cdots A_n) = tr\,(A_{\sigma(1)}A_{\sigma(2)}\cdots A_{\sigma(n)})\,, \tag{4.50}$$

where $(\sigma(1), \sigma(2), \dots, \sigma(n))$ *is a cyclic permutation of* $(1, 2, \dots, n)$*.*

Proof of Proposition 4.36 and of part (ii) in Theorem 4.32(b). (a) Immediate by definition of trace.
(b) Let us begin by showing the statement holds if A and B are both HS operators. By Theorem 4.24(b), equation (4.49) is equivalent to

$$(A^*|B)_2 = (B^*|A)_2 . \tag{4.51}$$

The proof of (4.51) is straightforward using the polarisation formula (valid for any inner product and the induced norm)

$$4(X|Y) = ||X+Y||^2 + ||X-Y||^2 - i||X+iY||^2 + i||X-iY||^2 ,$$

and recalling that, for HS norms, $||Z||_2 = ||Z^*||_2$ ((i) in Theorem 4.24(a)).

Now suppose A is of trace class and $B \in \mathfrak{B}(\mathsf{H})$. Then $A = CD$, with C and D HS operators by Theorem 4.32(a). In addition, DB and BC are Hilbert–Schmidt, for $\mathfrak{B}_2(\mathsf{H})$ is a two-sided ideal in $\mathfrak{B}(\mathsf{H})$. By swapping two HS operators at a time:

$$trAB = tr((CD)B) = tr(C(DB)) = tr((DB)C) = tr(D(BC)) = tr((BC)D)$$

$$= tr(B(CD)) = trBA .$$

(c) Since $\mathfrak{B}_1(\mathsf{H})$ is a two-sided ideal in $\mathfrak{B}(\mathsf{H})$, one operator of trace class among $A_1,\ldots,A_n$ is enough to render their product of trace class. In particular, using Theorem 4.32(a) and the fact that $\mathfrak{B}_2(\mathsf{H})$ is a two-sided ideal of $\mathfrak{B}(\mathsf{H})$ we see clearly that if two among $A_1,\ldots,A_n$ are HS, their product is of trace class. Then (4.50) is clearly equivalent to:

$$tr\,(A_1A_2\cdots A_n) = tr\,(A_2A_3\cdots A_nA_1)\,; \tag{4.52}$$

in fact, one permutation at a time, we obtain any cyclic permutation. Let us prove (4.52). Consider first the case of two HS operators A_i, A_j, $i < j$. If $i = 1$, the claim follows from (b) with $A = A_1$ and $B = A_2\cdots A_n$. If $i > 1$, the four operators (i) $A_1\cdots A_i$, (ii) $A_{i+1}\cdots A_n$, (iii) $A_{i+1}\cdots A_nA_1$, (iv) $A_2\cdots A_i$ are necessarily HS, for they involve either A_i or A_j as factor (never both). Hence:

$$tr(A_1\cdots A_n) = tr(A_1\cdots A_iA_{i+1}\cdots A_n) = tr(A_{i+1}\cdots A_nA_1A_2\cdots A_i)$$

$$= tr(A_2\cdots A_iA_{i+1}\cdots A_nA_1)\,,$$

which is what we wanted.

Let us prove invariance under permutations assuming A_i is of trace class. If $i = 1$ it follows from part (b) by taking $A = A_1$ and $B = A_2\cdots A_n$. So suppose $i > 1$. Then $A_1\cdots A_i$ and $A_2\cdots A_i$ are of trace class because both contain A_i, and then:

$$tr(A_1\cdots A_n) = tr(A_1\cdots A_iA_{i+1}\cdots A_n) = tr(A_{i+1}\cdots A_nA_1A_2\cdots A_i)$$

$$= tr(A_2\cdots A_iA_{i+1}\cdots A_nA_1)\,,$$

recalling part (b). Invariance under permutations allows to prove part (ii) in Theorem 4.32(b). Using (4.46) we have to prove $tr|A| = tr|A^*|$. By the corollary to the polar decomposition theorem (Theorem 3.71) we deduce $|A^*| = U|A|U^*$, where $U|A| = A$ is the polar decomposition of A. Hence

$$||A^*||_1 = tr|A^*| = tr(U|A|U^*) = tr(U^*U|A|) = tr|A| = ||A||_1 ,$$

where we used $U^*U|A| = |A|$, for U is isometric on $Ran(|A|)$ (Theorem 3.71). □

Remark 4.37. (1) A useful property of trace-class operators is the following. We will be able to justify it only after proving the spectral theorem for self-adjoint operators:

Proposition 4.38. *If* H *is a Hilbert space,* $T \in \mathfrak{B}(\mathsf{H})$ *is of trace class if and only if* $\sum_{u\in N} |(u|Tu)| < +\infty$ *for any basis* $N \subset \mathsf{H}$.

Proof. See Exercise 9.3. □

(2) If $A \in \mathfrak{B}_1(\mathsf{H})$ and $A = A^*$, computing the trace of A with an eigenvector basis of A itself (this exists by Theorem 4.18), we conclude $tr(A) = \sum_{\lambda\in\sigma_p(A)} m_\lambda \lambda$, where $\sigma_p(A)$ is, as always, the set of eigenvalues of A and m_λ the dimension of the λ-eigenspace. As for finite dimensions, for *self-adjoint* operators of trace class the trace coincides with the sum of the eigenvalues. This is still true even if A is *not* self-adjoint.

Theorem 4.39 (Lidiskii). *If* $T \in \mathfrak{B}_1(\mathsf{H})$ *with* H *a complex Hilbert space, then* $tr(T) = \sum_{\lambda\in\sigma_p(T)} m_\lambda \lambda$*, where* $\sigma_p(T)$ *is the eigenvalue set of* T*,* m_λ *the dimension of the* λ*-eigenspace, and the series converges absolutely.*

The result is far from obvious, and a proof can be found in [BiSo87]. ■

Example 4.40. For the following example familiarity with Riemannian geometry is required. An example of a trace-class operator important in physics arises [Mor99] when studying the *Laplace-Beltrami operator* (or *Laplacian*) on *Riemannian manifolds* (M,g). In local coordinates $x_1,\dots,x_n$ on the n-manifold M, the Laplacian is the differential operator:

$$\Delta = \sum_{i=1}^{n} \frac{1}{\sqrt{g}} \frac{\partial}{\partial x_i} g^{ij}(x) \sqrt{g} \frac{\partial}{\partial x_j} ,$$

where g denotes the determinant of the matrix of coefficients g_{ij}, $i,j = 1,\dots,n$, that describes the metric tensor in the given coordinates, and g^{ij} are the coefficients of the inverse matrix. If $V : M \to (K,+\infty)$, for a certain $K > 0$, is an arbitrary smooth map, we consider the operator $A = -\Delta + V$ defined on the space $\mathscr{D}(M)$ of smooth complex-valued maps on M. We may view $\mathscr{D}(M)$ as a (dense) subspace in $L^2(M,\mu_g)$, where μ_g, the natural Borel measure associated to the metric, reads $\sqrt{g}dx_1\cdots dx_n$ in local coordinates. The operator A is positive, not bounded, and admits a unique inverse (also positive): $A^{-1} : L^2(M,\mu_g) \to \mathscr{D}(M)$. Thinking A^{-1} as an $L^2(M,\mu_g)$-valued operator, it turns out that $A^{-1} \in \mathfrak{B}(L^2(M,\mu_g))$. The interesting point is that $A^{-1} \in \mathfrak{B}_\infty(L^2(M,\mu_g))$. But there is more to the story. A theorem due to Weyl proves the eigenvalues $\lambda_j \in \sigma_p(A)$ of A (where j tags the eigenvector ϕ_j and not the corresponding eigenvalue λ_j, so that $\phi_i \neq \phi_k$ if $k \neq i$ but $0 < K \leq \lambda_0 \leq \lambda_1, \leq \lambda_2, \leq \cdots$) satisfy an estimate:

$$\lim_{j\to+\infty} j^{-1}\lambda_j^{n/2} = \frac{k_n}{vol(M)} , \tag{4.53}$$

where $vol(M)$ is the manifold's volume (finite by compactness) and k_n a universal constant that depends only upon the manifold's dimension. Furthermore, the eigenvectors $\{\phi_j\}_{j\in\mathbb{N}}$ form a basis of $L^2(M,\mu_g)$, which implies the eigenvalue set $\sigma_p(A^{-k})$ of A^{-k} satisfies $\sigma_p(A^{-k}) = \{\lambda_j^{-k}\}_{j\in\mathbb{N}}$. Computing the norms $||\ ||_1$ and $||\ ||_2$ of $A^{-k} = |A|^{-k}$, and using the basis of eigenvectors $\{\phi_j\}_{j\in\mathbb{N}}$ of A, we have:

$$||A^{-k}||_1 = \sum_{j=0}^{+\infty} \lambda_j^{-k} \quad \text{and} \quad ||A^{-k}||_2^2 = \sum_{j=0}^{+\infty} \lambda_j^{-2k}.$$

The estimate (4.53) implies $A^{-k} \in \mathfrak{B}_1(L^2(M,\mu_g))$ if $k > n/2$, and $A^{-k} \in \mathfrak{B}_2(L^2(M,\mu_g))$ if $k > n/4$. ■

4.5 Introduction to the Fredholm theory of integral equations

Integral equations are a central branch of functional analysis, especially for applications in physics (for instance in the theory of quantum scattering [Pru81] and the study of inverse problems) and other sciences. In the sequel we shall deal with general results due to Fredholm in most part, and we will regularly assume an abstract viewpoint, where integral operators are seen as particular compact operators on Hilbert spaces (even though several results can be extended to general Banach spaces). We shall essentially follow the treatment of [KoFo99].

To fix ideas, let us consider a measure space (X,Σ,μ), where $\mu : \Sigma \to [0,+\infty]$ is a positive (σ-additive) measure that is σ-finite and separable, and take a map $K \in L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)$ with no further properties. $T_K \in \mathfrak{B}_2(\mathsf{H})$ is the usual integral operator (cf. Examples 4.16(3), (4) and 4.25(1), (2)) on $\mathsf{H} = L^2(\mathsf{X},\mu)$ defined by:

$$(T_K\psi)(x) := \int_{\mathsf{X}} K(x,y)\psi(y)d\mu(y)\,. \tag{4.54}$$

We wish to study, in broad terms, the integral equation:

$$T_K\varphi - \lambda\varphi = f\,, \tag{4.55}$$

where $f \in \mathsf{H}$ is given, $\varphi \in \mathsf{H}$ the unknown, $\lambda \in \mathbb{C}$ a constant.

To begin wth, consider the case $\lambda = 0$. This gives the so-called **Fredholm equation of the first kind** on the Hilbert space H.

From the abstract point of view we have to solve for $\varphi \in \mathsf{H}$ the equation:

$$A\varphi = f\,,$$

where $A : \mathsf{H} \to \mathsf{H}$ is a compact operator (in the concrete case A is T_K, a HS operator) and $f \in \mathsf{H}$ a given element.

An important general result, valid also with an infinite-dimensional Banach space B replacing H, which assumes A compact, is that the equation has *no* solution for certain $f \in \mathsf{H}$, irrespective of $A \in \mathfrak{B}_\infty(\mathsf{H})$. This follows from the next proposition.

Proposition 4.41. *If $A \in \mathfrak{B}_\infty(\mathsf{B})$ with B Banach space of infinite dimension, then $Ran(A) \neq \mathsf{B}$.*

Proof. We can write $\mathsf{B} = \cup_{n\in\mathbb{N}} B_n$, where B_n is the open ball of radius n at the origin, so:

$$Ran(A) = \cup_{n\in\mathbb{N}} A(B_n)\,.$$

If $Ran(A)$ were equal to B we could write:

$$\mathsf{B} = \cup_{n\in\mathbb{N}} A(B_n) \subset \cup_{n\in\mathbb{N}} \overline{A(B_n)} \subset \mathsf{B},$$

hence

$$\mathsf{B} = \cup_{n\in\mathbb{N}} \overline{A(B_n)}\,,$$

where any $\overline{A(B_n)}$ is compact because A is compact and B_n bounded. Therefore B would become a countable union of compact sets, which is impossible by Corollary 4.6. □

A second issue with Fredholm equations of the first kind arises from the next proposition.

Proposition 4.42. *Every left inverse to a compact injective operator $A \in \mathfrak{B}_\infty(\mathsf{X})$, where X is a normed space, cannot be bounded if X is infinite-dimensional.*

Proof. The proof is in Exercise 4.1. □

Because of this result if two equations $A\varphi = f_1$ and $A\varphi = f_2$ both admit solutions, the latter may be very different, even if f_1 and f_2 are close in norm. In other terms Fredholm equations of the first kind are *ill posed* problems *à la* Goursat. This does not entail, obviously, that Fredholm equations of the first kind are mathematically uninteresting or not useful in applied sciences. What it means is that their study is hard and requires advanced and specialised topics, that reach well beyond the present elementary treatise.

We can consider the case $\lambda \neq 0$ in (4.55). The equation in this case is called **Fredholm equation of the second kind**. For a short moment we assume the operator T_K admits a *Hermitian kernel*. In other terms we consider:

$$T_K\varphi - \lambda\varphi = f\,, \tag{4.56}$$

where T_k has the form (4.54) with $\lambda \neq 0$ fixed in $\mathbb{C}$ and $K(x,y) = \overline{K(y,x)}$, so that, by (4.3), $T_K = T_K^*$. In such a case we can state a more general theorem.

Theorem 4.43 (Fredholm equations of the second kind with Hermitian kernels). *Let $\mathsf{H} = L^2(\mathsf{X},\mu)$ be a Hilbert space with σ-finite and separable positive σ-additive measure μ. Given $f \in \mathsf{H}$ and $\lambda \in \mathbb{C}\setminus\{0\}$, consider equation (4.56) in $\varphi \in \mathsf{H}$ with*

$$(T_K\varphi)(x) = \int_{\mathsf{X}} K(x,y)\varphi(y)d\mu(y)\,,$$

where $K \in L^2(\mathsf{X}\times\mathsf{X},\mu\otimes\mu)$ and $K(x,y) = \overline{K(y,x)}$. The following facts hold.

(a) *If the number λ is not an eigenvalue of T_K, equation (4.56) has always a unique solution, whichever $f \in \mathsf{H}$.*
(b) *If λ is an eigenvalue of T_K, equation (4.56) has solutions iff f is orthogonal to the λ-eigenspace. In such a case equation (4.56) admits infinitely many solutions.*

Proof. Multiplying equation (4.56) by λ^{-1} allows to study only $\lambda = 1$ (redefining $\lambda^{-1}K$ as K and $\lambda^{-1}f$ as f). Hence we prove it in this case only. We know T_K is compact by Example 4.16(4) and self-adjoint up to a possible and inessential factor $(1/\lambda)$. So we shall refer to Theorems 4.17 and 4.18. Let $\{\psi_n\}_{n\in\mathbb{N}}$ be a basis of eigenvectors of T_K. We can decompose, uniquely, any $\varphi \in \mathsf{H}$ as

$$\varphi = \sum_{n=1}^{+\infty} a_n \psi_n + \varphi' , \tag{4.57}$$

where $\varphi' \in Ker(T_K)$ and the $a_n \in \mathbb{C}$ are uniquely determined by φ. In particular

$$f = \sum_{n=1}^{+\infty} b_n \psi_n + f' .$$

Let us seek a solution to (4.56):

$$\varphi = T_K \varphi - f$$

in the form (4.57). We must find numbers a_n and the element φ' once T_K and f are given. Substituting (4.57) and the analogous for f in (4.56), we easily find

$$\sum_n a_n \psi_n + \varphi' = \sum_n a_n \lambda_n \psi_n - \sum_n b_n \psi_n - f' ,$$

where λ_n are the *non-null* eigenvalues of T_K corresponding to eigenvectors ψ_n (in general it may happen $\lambda_n = \lambda_{n'}$, for we have numbered eigenvectors and not eigenvalues). That is to say:

$$\sum_n a_n (1 - \lambda_n + b_n) \psi_n = -f' - \varphi'_n .$$

Notice the two sides are orthogonal by construction, and so are the vectors ψ_n, pairwise. Therefore the identity is equivalent to:

$\varphi' = -f'$,
$a_n(1-\lambda_n) = -b_n$, $n = 1,2,\ldots$.

In any case φ' is always determined, for it concides with f'. The existence of solutions to (4.56) tantamounts to:

$\varphi' = -f'$,
$a_n = \frac{b_n}{\lambda_n - 1}$ for $\lambda_n \neq 1$
$b_0 = 0$ for $\lambda_m = 1$.

If $\lambda_n \neq 1$ for every n, the coefficients a_n are uniquely determined by the b_n. If $\lambda_m = 1$ for some m and $b_m \neq 0$, the last condition is false, and equation (4.56) has no solution.

Instead, if $b_m = 0$ for any m such that $\lambda_m = 1$ (i.e. if f is normal to the eigenspace of T_K with eigenvalue 1), the coefficients a_m can be chosen at will, whereas the remaining a_n are determined. In this case there are infinitely many solutions to (4.56). □

To conclude we move to the general case and drop the assumption on Hermitian kernels. To stay general we shall study the abstract *Fredholm equation of the second kind* in the Hilbert space H:

$$A\varphi - \lambda\varphi = f\,, \tag{4.58}$$

where $f \in \mathsf{H}$, $\lambda \in \mathbb{C} \setminus \{0\}$ and $A \in \mathfrak{B}_\infty(\mathsf{H})$ are given on a certain Hilbert space H and $\varphi \in \mathsf{H}$ is the problem's unknown. Nothing more is assumed about A, apart compactness. In particular, we will not assume A is a Hilbert–Schmidt operator. Let us prove the following general theorem, due to Fredholm, which can be stated also for $A \in \mathfrak{B}_\infty(\mathsf{B})$, B a Banach space.

Theorem 4.44 (Fredhol). *On the Hilbert space* H *consider the abstract* **Fredholm equation of the second kind**

$$A\varphi - \lambda\varphi = f \tag{4.59}$$

in the unknown $\varphi \in \mathsf{H}$, *with* $f \in \mathsf{H}$, $\lambda \in \mathbb{C} \setminus \{0\}$ *and* $A \in \mathfrak{B}_\infty(\mathsf{H})$ *given. Consider as well the corresponding* **homogeneous equation***, the* **adjoint equation** *and the* **homogeneous adjoint equation***:*

$$A\varphi_0 \;-\lambda\varphi_0 = 0\,, \tag{4.60}$$
$$A^*\psi \;-\overline{\lambda}\psi = g\,, \tag{4.61}$$
$$A^*\psi_0 \;-\overline{\lambda}\psi_0 = 0\,, \tag{4.62}$$

respectively, with $g \in \mathsf{H}$ *given and* $\varphi_0, \psi, \psi_0 \in \mathsf{H}$ *unknown.*

(a) *The equation* $A\varphi - \lambda\varphi = f$ *admits solutions* *iff* f *is orthogonal to each solution* ψ_0 *to the homogeneous adjoint equation* $A^*\psi_0 - \overline{\lambda}\psi_0 = 0$.
(b) *Either the equation* $A\varphi - \lambda\varphi = f$ *admits exactly one solution for any* $f \in \mathsf{H}$, *or the homogeneous equation* $A\varphi_0 - \lambda\varphi_0 = 0$ *has a non-zero solution.*
(c) *The homogeneous equations* $A\varphi_0 - \lambda\varphi_0 = 0$ *and* $A^*\psi_0 - \overline{\lambda}\psi_0 = 0$ *admit the same – finite – number of linearly independent solutions.*

Remark 4.45. (1) Statement (b) is the celebrated *Fredholm alternative.*
(2) The above theorem holds in particular if A is self-adjoint, and becomes Theorem 4.43. ■

Proof of Theorem 4.44. Here, too, dividing the initial equation by $\lambda \neq 0$ permits to study the case $\lambda = 1$ only (after redefining $\lambda^{-1}A$ as A, $\lambda^{-1}f$ as f and $\overline{\lambda}^{-1}g$ as g). Henceforth, then, $\lambda = 1$. Set $T := A - I$, and observe it is bounded but not compact, for I is not compact. The theorem relies on three lemmas. Let us first notice that $Ker(T)$ is always a closed subspace in H if T is continuous, as in the present situation, whereas the subspace $Ran(T)$ is not. The first lemma shows that $Ran(T)$ is closed as well, provided $T := A - I$, $A \in \mathfrak{B}_\infty(\mathsf{H})$.

Lemma 4.46. *Under the assumptions made on T, $Ran(T)$ is closed.*

Proof. Let $y_n \in Ran(T)$, $n \in \mathbb{N}$, and suppose $y_n \to y$ as $n \to +\infty$. We need to prove $y \in Ran(T)$. By assumption:

$$y_n = Tx_n = Ax_n - x_n \tag{4.63}$$

for some sequence $\{x_n\}_{n\in\mathbb{N}} \in \mathsf{H}$. With no loss of generality we may assume $x_n \in Ker(T)^\perp$, possibly eliminating from the sequence what projects onto $Ker(T)$. If we can prove the sequence of x_n is bounded the proof ends. In fact, A being compact, there will exist a subsequence x_{n_k} such that $Ax_{n_k} \to y' \in \mathsf{H}$ for $k \to \infty$. Substituting in (4.63) we conclude $x_{n_k} \to x$ for some $x \in \mathsf{H}$ as $k \to +\infty$. By continuity of A, $Tx = Ax - x = y$, so $y \in Ran(T)$.

We will proceed by contradiction, and assume $\{x_n\}_{n\in\mathbb{N}} \subset Ker(T)^\perp$ is bounded. If not, there would be a subsequence x_{n_m} with $0 < ||x_{n_m}|| \to +\infty$ as $m \to +\infty$. Since the y_n form a convergent, hence bounded, sequence, dividing by $||x_{n_m}||$ in (4.63) would give :

$$T\frac{x_{n_m}}{||x_{n_m}||} = A\frac{x_{n_m}}{||x_{n_m}||} - \frac{x_{n_m}}{||x_{n_m}||} = \frac{y_{n_m}}{||x_{n_m}||} \to 0\,. \tag{4.64}$$

But A is compact and the $\frac{x_{n_m}}{||x_{n_m}||}$ bounded, so we could extract a further subsequence $x_{n_{m_k}}/||x_{n_{m_k}}||$ such that:

$$\frac{x_{n_{m_k}}}{||x_{n_{m_k}}||} \to x' \in \mathsf{H} \quad \text{and} \quad T\frac{x_{n_{m_k}}}{||x_{n_{m_k}}||} \to Tx' \quad \text{as } k \to +\infty.$$

By (4.64) we would infer $Tx' = 0$ i.e. $x' \in Ker(T)$. By assumption $\frac{x_{n_{m_k}}}{||x_{n_{m_k}}||} \in Ker(T)^\perp$, and $Ker(T)^\perp$ is closed, so $x' \in Ker(T)^\perp$. Consequently $x' \in Ker(T) \cap Ker(T)^\perp = \{0\}$, in contradiction to

$$||x'|| = \lim_{k\to+\infty} \frac{||x_{n_{m_k}}||}{||x_{n_{m_k}}||} = 1\,.$$

This ends the proof. □

The second lemma claims the following.

Lemma 4.47. *Under the same assumptions on T we have orthogonal decompositions:*

$$H = Ker(T) \oplus Ran(T^*) = Ker(T^*) \oplus Ran(T)\,. \tag{4.65}$$

Proof. Since $T, T^* \in \mathfrak{B}(\mathsf{H})$, Theorem 3.13(e-d) and Proposition 3.38(d) imply $Ker(T) = (Ran(T^*)^\perp)^\perp = \overline{Ran(T^*)}$, $Ker(T^*) = (Ran(T)^\perp)^\perp = \overline{Ran(T)}$, and:

$$H = Ker(T) \oplus \overline{Ran(T^*)} = Ker(T^*) \oplus \overline{Ran(T)}\,.$$

The previous lemma holds also if we replace T with T^*, for $(A-I)^* = A^* - I$ where A^* is compact if A is. □

By Lemma 4.47 the statement in Theorem 4.44(a) follows (we wanted to prove the case $\lambda = 1$). In fact, from it we have $f \perp Ker(T^*)$ iff $f \in Ran(T)$, i.e. iff there is $\varphi \in \mathsf{H}$ such that $T\varphi = f$.

To continue with the proof of part (b) in the main theorem we define the subspaces (closed by Lemma 4.46) $\mathsf{H}^k := Ran(T^k)$, $k = 1, 2, \ldots$, so that:

$$\mathsf{H} \supset \mathsf{H}^1 \supset \mathsf{H}^2 \supset \mathsf{H}^3 \supset \cdots .$$

By construction $T(\mathsf{H}^k) = \mathsf{H}^{k+1}$. And now we have the third lemma.

Lemma 4.48. *With the given definition of* H^k*,* $k = 1, 2, \ldots$*, there exists* $j \in \mathbb{N}$ *such that:*

$$\mathsf{H}^{k+1} = \mathsf{H}^k \quad \text{if } k \geq j.$$

Proof. Assume, by contradiction, that j does not exist. Then $\mathsf{H}^k \neq \mathsf{H}^h$ if $k \neq h$, and we can manifacture a sequence of orthonormal vectors x_k sucht hat $x_k \in \mathsf{H}^k$ and $x_k \perp H^{k+1}$, $k = 1, 2, \ldots$. If $l > k$

$$Ax_l - Ax_k = -x_k + (x_l + Tx_l - Tx_k)$$

so $||Ax_l - Ax_k||^2 \geq 1$ because $x_l + Tx_l - Tx_k \in \mathsf{H}^{k+1}$. But then we cannot extract any converging subsequence from Ax_k, contradicting the fact A is compact. □

Next we prove two lemmas that, combined, yield the proof of Theorem 4.44(b) (for $\lambda = 1$).

Lemma 4.49. *Under the previous assumptions on* T*,* $Ker(T) = \{0\}$ *implies* $Ran(T) = \mathsf{H}$.

Proof. Assume $Ker(T) = \{0\}$, making T one-to-one, but by contradiction suppose $Ran(T) \neq \mathsf{H}$. Then the H^k, $k = 1, 2, 3, \ldots$ would be all distinct, violating Lemma 4.48. □

Lemma 4.50. *Under the assumptions made on* T*,* $Ran(T) = \mathsf{H}$ *implies* $Ker(T) = \{0\}$.

Proof. If $Ran(T) = \mathsf{H}$, by Lemma 4.47 we have $Ker(T^*) = \{0\}$. Then the previous lemma (with T^* instead of T) guarantees $Ran(T^*) = \mathsf{H}$. Now Lemma 4.47 again forces $Ker(T) = \{0\}$. □

It is patent that Lemmas 4.49 and 4.50 together prove statement (b) in Theorem 4.44.

We finish by proving part (c) (for $\lambda = 1$).

Suppose $\dim Ker(T) = +\infty$, in rebuttal to (c). Then there is an infinite orthonormal system $\{x_n\}_{n\in\mathbb{N}} \subset Ker(T)$. By construction $Ax_n = x_n$ and so $||Ax_k - Ax_h||^2 = 2$. But this cannot be, for it would infringe the existence of a subsequence in the bounded $\{x_n\}_{n\in\mathbb{N}}$ such that $\{Ax_{n_k}\}_{k\in\mathbb{N}}$ converges, which we know is true by compactness. Hence $\dim Ker(T) = m < +\infty$. Similarly $\dim Ker(T^*) = n < +\infty$. Assume, contraditing the statement, that $m \neq n$. In particular we may suppose $m < n$. Let then

$\{\varphi_j\}_{j=1,\dots,n}$ and $\{\psi_j\}_{j=1,\dots,m}$ be orthonormal bases for $Ker(T)$ and $Ker(T^*)$ respectively. Define $S \in \mathfrak{B}(\mathsf{H})$ by:

$$Sx := Tx + \sum_{j=1}^{m} (\varphi_j|x)\psi_j \,.$$

As $S = A' - I$, with A' compact (obtained from the compact operator A by adding a compact operator, since with finite-dimensional range), the result found above for $T = A - I$ hold for S too.

We claim $Sx = 0$ implies $x = 0$. Explicitly

$$Tx + \sum_{j=1}^{m} (\varphi_j|x)\psi_j = 0 \,. \tag{4.66}$$

By virtue of Lemma 4.47, all vectors ψ_j are orthogonal to any one of the form Tx, so (4.66) implies $Tx = 0$. Moreover, $(\varphi_j|x) = 0$ if $1 \leq g \leq m$, because x is a linear combination of the φ_j and simultaneously it must be orthogonal to them, and therefore $x = 0$. Hence we have that $Sx = 0$ implies $x = 0$. From part (b), then, there is $y \in \mathsf{H}$ such that

$$Ty + \sum_{j=1}^{m} (\varphi_j|y)\psi_j = \psi_{m+1} \,.$$

Taking the inner product with ψ_{m+1} gives a contradiction: 1 on the left equals 0 on the right, because $Ty \in Ran(T)$ and $Ran(T) \perp Ker(T^*)$. This shows we cannot assume $m < n$, so it must be $m \geq n$. Swapping the roles of T and T^* gives $n \geq m$, so altogether $m = n$, ending the proof of (c). □

Examples 4.51. An interesting instance of Fredholm equation of the second kind is provided by **Volterra's equation of the second kind**:

$$\varphi(x) = \int_a^x K(x,t)\varphi(t)dt + f(x) \,, \tag{4.67}$$

where $\varphi \in L^2([a,b],dx)$ is the unknown function, $f \in L^2([a,b],dx)$ is given and the integral kernel satisfies $|K(x,t)| < M < +\infty$ for any $x,t \in [a,b]$, $t \leq x$. (Any multiplicative factor $\rho \in \mathbb{C} \setminus \{0\}$ is absorbed in K.)

This equation befits the theory of Fredholm's theorem if we rewrite the integral as an intergal over all $[a,b]$ and assuming $K(x,t) = 0$ if $t \geq x$. For this type of equation, though, there is a better result based on contraction maps (cf. Section 2.6). It turns out, namely, that a certain high power of $T_K : L^2([a,b],dx) \to L^2([a,b],dx)$ is a contraction, where T_K is the integral operator in (4.67)

$$(T_K\varphi)(x) = \int_a^x K(x,t)\varphi(t)dt \,.$$

Consequently the homogeneous equation $T_K\varphi = 0$ has one, and one only, solution by Theorem 2.108, and the solution must be $\varphi = 0$. The proof that T_K^n is a contraction if n is big enough is similar to what we saw in Example 2.109(1), where the Banach space $(C([a,b]), || \, ||_\infty)$ is replaced by $(L^2([a,b],dx), || \, ||_2)$ (cf. Exercise 4.19). That said, parts (a) and (b) in Fredholm's theorem imply that equation (4.67) has *exactly* one solution, for any choice of the source term $f \in L^2([a,b],dx)$. ■

Exercises

4.1. Prove that if X is a normed space and $T : \mathsf{X} \to \mathsf{X}$ is compact and injective, then any linear operator $S : Ran(T) \to \mathsf{X}$ that inverts T on the left cannot be bounded if $\dim \mathsf{X} = \infty$.

Solution. If S were bounded, Proposition 2.44 would allow to extend it to a bounded operator $\tilde{S} : \mathsf{Y} \to \mathsf{X}$, where $\mathsf{Y} := \overline{Ran(T)}$, so that $\tilde{S}T = I$. Precisely as in the proof of Proposition 4.9(b), we could prove $\tilde{S}T$ is compact if $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$ is compact and $\tilde{S} \in \mathfrak{B}(\mathsf{Y},\mathsf{X})$. Then $I : \mathsf{X} \to \mathsf{X}$ would be compact, and thus the unit ball in X would have compact closure, breaching Proposition 4.5.

4.2. Using Banach's Lemma 4.12 prove that in an infinite-dimensional normed space the closed unit ball is not comapct.

Outline. Let $x_1, x_2, \ldots$ be an infinite sequence of linearly independent vectors with $||x_n|| = 1$ (so, all belonging to the closure of the unit ball). Banach's lemma constructs a sequence of vectors $y_1, y_2, \ldots$ such that $||y_n|| = 1$ and $||y_{n-1} - y_n|| > 1/2$. This sequence cannot contain converging subsequences.

4.3. Prove that if $A^* = A \in \mathfrak{B}_\infty(\mathsf{H})$, H a Hilbert space, then

$$\sigma_p(|A|) = \{|\lambda| \, | \lambda \in \sigma_p(A)\} .$$

Conclude that if $A^* = A \in \mathfrak{B}_\infty(\mathsf{H})$

$$sing(A) = \{|\lambda| \, | \lambda \in \sigma_p(A) \setminus \{0\}\} .$$

Solution. Expanding the compact self-adjoint operators A and $|A|$ according to Theorem 4.18, we have

$$A = \sum_{\lambda \in \sigma_p(A)} \lambda P_\lambda \quad \text{and} \quad |A| = \sum_{\mu \in \sigma_p(|A|)} \mu P'_\mu ,$$

with the obvious notation. Squaring A and $|A|$ and using their continuity (this allows to consider all series as finite sums), idempotency and orthogonality of projectors relative to distinct eigenvectors, and also recalling $|A|^2 = A^*A = A^2$, we have

$$\sum_{\lambda \in \sigma_p(A)} \lambda^2 P_\lambda = \sum_{\mu \in \sigma_p(|A|)} \mu^2 P'_\mu . \tag{4.68}$$

Now keep in mind $P_\lambda P_{\lambda_0} = 0$ if $\lambda \neq \lambda_0$ and $P_\lambda P_{\lambda_0} = P_{\lambda_0}$ otherwise, and the same holds for the projectors in the decomposition of $|A|$. Composing with P_{λ_0} on the right in (4.68), taking adjoints and eventually right-composing with P'_{μ_0}, we find, for any $\lambda_0 \in \sigma(A)$ and $\mu_0 \in \sigma_p(|A|)$, that $\lambda_0^2 P_{\lambda_0} P'_{\mu_0} = \mu_0^2 P_{\lambda_0} P'_{\mu_0}$, i.e.

$$(\lambda_0^2 - \mu_0^2) P_{\lambda_0} P'_{\mu_0} = 0 . \tag{4.69}$$

The fact that A admits a basis of eigenvectors (Theorem 4.18) is known to be equivalent to

$$I = s\text{-} \sum_{\lambda_0 \in \sigma(A)} P_{\lambda_0} .$$

Fix $\mu_0 \in \sigma_p(|A|)$. If $P_{\lambda_0} P'_{\mu_0} = 0$ for any $\lambda_0 \in \sigma(A)$, from the above identity we would have $P'_{\mu_0} = 0$, absurd by definition of eigenspace. Therefore, (4.69) withstanding, there must exist $\lambda_0 \in \sigma(A)$ such that $\lambda_0^2 = \mu_0^2$, i.e. $\mu_0 = |\lambda_0|$. If $\lambda_0 \in \sigma_p(A)$, swapping A and $|A|$ and using a similar argument would produce $\mu_0 \in \sigma_p(|A|)$ such that $\lambda_0^2 = \mu_0^2$, i.e. $\mu_0 = |\lambda_0|$. The first assertion is thus proved. The second one is evident by the definition of singular value.

4.4. Consider the separable Hilbert space H, the basis $\{f_n\}_{n\in\mathbb{N}} \subset \mathsf{H}$ and the sequence $\{s_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$ with $|s_{n+1}| \geq |s_n|$ and $|s_n| \to 0$, $n \to +\infty$. Using the uniform topology prove

$$T := \sum_{n=0}^{+\infty} s_n f_n (f_n|\ )$$

is well defined and $T \in \mathfrak{B}_\infty(\mathsf{H})$. Show that if every s_n is real, T is also self-adjoint and every s_n is an eigenvalue of T.

Hint. With the assumptions made, if $T_N := \sum_{n=0}^{N} s_n f_n (f_n|\ )$ and $N \geq M$:

$$||T_N x - T_M x||^2 \leq |s_M|^2 \sum_{n=M}^{N-1} |(f_n|x)|^2 \leq |s_M|^2 ||x||^2 .$$

Taking the least upper bound over the unit $x \in \mathsf{H}$ gives:

$$||T_N - T_M|| \leq |s_M|^2 \to 0 \quad \text{as } N, M \to +\infty,$$

whence the first part. The rest follows by direct inspection.

4.5. Prove that if $T \in \mathfrak{B}_\infty(\mathsf{H})$ and if $\mathsf{H} \ni x_n \to x \in \mathsf{H}$ *weakly*, i.e.

$$(g|x_n) \to (g|x) \quad \text{as } n \to +\infty, \text{ for any given } g \in \mathsf{H},$$

then $||T(x_n) - T(x)|| \to 0$ as $n \to +\infty$. Put otherwise, compact operators map weakly converging sequences to sequences converging in norm. Extend the result to the case $T \in \mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$, X and Y normed spaces.

Solution. Let $x_n \to x$ weakly. If we bear in mind Riesz's theorem, the set $\{x_n\}_{n\in\mathbb{N}}$ is immediately weakly bounded in the sense of Corollary 2.60 to Banach–Steinhaus. According to the latter, $||x_n|| \leq K$ for any $n \in \mathbb{N}$ and some $K > 0$. So let us define $y_n := Tx_n$, $y := Tx$ and note that for any $h \in \mathsf{H}$,

$$(h|y_n) - (h|y) = (T^*h|x_n) - (T^*h|x) \to 0 \quad \text{as } n \to +\infty,$$

hence also the y_n converge weakly to y. Suppose, by contradiction, $||y_n - y|| \not\to 0$ as $n \to +\infty$. Then there exist $\varepsilon > 0$ and a subsequence $\{y_{n_k}\}_{k\in\mathbb{N}}$ with $||y - y_{n_k}|| \geq \varepsilon$ for any $k \in \mathbb{N}$. Since $\{x_{n_k}\}_{k\in\mathbb{N}}$ is bounded by K, and T is compact, there must be a subsequence $\{y_{n_{k_r}}\}_{r\in\mathbb{N}}$ converging to a $y' \neq y$. This subsequence $\{y_{n_{k_r}}\}_{r\in\mathbb{N}}$ has to converge to y' also weakly. But this cannot be, for $\{y_n\}_{n\in\mathbb{N}}$ converges weakly to $y \neq y'$. Therefore $y_n \to y$ in the norm of H. The argument works in the more general case where $T \in \mathfrak{B}_\infty(\mathsf{X}, \mathsf{Y})$, X and Y normed spaces, by interpreting $\mathsf{X} \ni x_n \to x \in \mathsf{X}$ in *weak* sense:

$$g(x_n) \to g(x) \quad \text{as } n \to +\infty, \text{ for any given } g \in \mathsf{X}',$$

because Corollary 2.60 to Banach–Steinhaus still holds. In the proof note that $h \in \mathsf{Y}'$ implies $h \circ T \in \mathsf{X}'$, by composition of continuous linear mappings.

4.6. Referring to Examples 4.25, take $T_K, T_{K'} \in \mathfrak{B}_2(L^2(\mathsf{X}, \mu))$ (μ assumed separable) with integral kernels K, K'. Prove that the HS operator $aT_K + bT_{K'}$, $a, b \in \mathbb{C}$, has kernel $aK + bK'$.

4.7. Given $T_K \in \mathfrak{B}_2(L^2(\mathsf{X}, \mu))$ by the integral kernel K, prove the HS operator T_K^* has integral kernel K^*, $K^*(x,y) = \overline{K(y,x)}$.

4.8. With the same hypotheses of Exercise 4.6 show the integral kernel of $T_K T_{K'}$ is

$$K''(x,z) := \int_{\mathsf{X}} K(x,y)K'(y,z)\, d\mu(y)\,.$$

4.9. If $L^2(\mathsf{X}, \mu)$ is separable, consider the mapping $L^2(\mathsf{X} \times \mathsf{X}, \mu \otimes \mu) \ni K \mapsto T_K \in \mathfrak{B}_2(L^2(\mathsf{X}, \mu))$. Prove it is an isomorphism of Hilbert spaces. Discuss whether one can view the map as an isometry of normed spaces, assuming $\mathfrak{B}(L^2(\mathsf{X}, \mu))$ as codomain. Discuss whether the map is continuous if viewed as a homeomorphisms only.

4.10. In relationship to Exercise 4.25(3), prove that if $g \in C_0([0,1])$,

$$\left((I - \rho T)^{-1} g\right)(x) = g(x) + \rho \int_0^x e^{\rho(x-y)} g(y) dy\,.$$

Hint. Use the operator $I - \rho T$, recalling the integral expression of T and noticing $\rho e^{\rho(x-y)} = \frac{\partial}{\partial x} e^{\rho(x-y)}$.

4.11. Let $\mathfrak{B}_D(L^2(\mathsf{X}, \mu))$ be the set of *degenerate operators* on $L^2(\mathsf{X}, \mu)$ (cf. 4.25(4)), with μ separable. Prove the following are equivalent statements.

(a) $T \in \mathfrak{B}_D(L^2(\mathsf{X}, \mu))$.
(b) $Ran(T)$ has finite dimension.
(c) $T \in \mathfrak{B}_2(L^2(\mathsf{X}, \mu))$ (integral) with integral kernel $K(x,y) = \sum_{k=1}^N p_k(x) q_k(y)$, where $p_1, \ldots, p_N \in L^2(\mathsf{X}, \mu)$, $q_1, \ldots, q_N \in L^2(\mathsf{X}, \mu)$ are linearly independent.

4.12. Take the set $\mathfrak{B}_D(L^2(\mathsf{X},\mu))$ of degenerate operators (cf. 4.25(4)) on $L^2(\mathsf{X},\mu)$, with μ separable. Show $\mathfrak{B}_D(L^2(\mathsf{X},\mu))$ is a two-sided *-ideal in $\mathfrak{B}(L^2(\mathsf{X},\mu))$ and a subspace in $\mathfrak{B}_2(L^2(\mathsf{X},\mu))$. In other words, show $\mathfrak{B}_D(L^2(\mathsf{X},\mu) \subset \mathfrak{B}_2(L^2(\mathsf{X},\mu))$, that it is a closed subspace under Hermitian conjugation, and $AD, DA \in \mathfrak{B}_D(L^2(\mathsf{X},\mu)$ if $A \in \mathfrak{B}(L^2(\mathsf{X},\mu))$ and $D \in \mathfrak{B}_D(L^2(\mathsf{X},\mu))$.

4.13. Consider $\mathfrak{B}_D(L^2(\mathsf{X},\mu))$ (cf. 4.25(4)) with μ separable. Does the closure of $\mathfrak{B}_D(L^2(\mathsf{X},\mu))$ in $\mathfrak{B}_2(L^2(\mathsf{X},\mu))$ in the norm topology of $\mathfrak{B}(L^2(\mathsf{X},\mu))$ coincide with $\mathfrak{B}_2(L^2(\mathsf{X},\mu))$?

Hint. Consider the operator

$$T := \sum_{n=0}^{+\infty} \frac{1}{\sqrt{n}} T_{K_n}\,,$$

where $K_n(x,y) = \phi_n(x)\phi_n(y)$, $\{\phi_n\}_{n\in\mathbb{N}}$ a basis of $L^2(\mathsf{X},\mu)$ and the convergence is in uniform norm. Prove $T \in \mathfrak{B}(L^2(\mathsf{X},\mu))$ is well defined, but $T \notin \mathfrak{B}_2(L^2(\mathsf{X},\mu))$ since $||T\phi_n||^2 = 1/n$.

4.14. Under the assumptions of Mercer's Theorem 4.27, prove that if $T_K \in \mathfrak{B}_1(L^2(\mathsf{X},d\mu))$ then $tr(T_K) = \int_{\mathsf{X}} K(x,x)d\mu(x)$.

Hint. Expand the trace in the basis of eigenvectors given by the continuous maps in the theorem's statement. Since the series that defines K converges uniformly on the compact set X of finite measure, a smart use of dominated convergence allows to see

$$\int_{\mathsf{X}} K(x,x)d\mu(x) = \int_X \sum_{\lambda,i} \lambda u_{\lambda,i}(x)\overline{u_{\lambda,i}(x)}d\mu(x) = \sum_{\lambda,i} \lambda \int_X u_{\lambda,i}(x)\overline{u_{\lambda,i}(x)}d\mu(x)$$

$$= \sum_{\lambda,i} \lambda = tr(T_K)\,.$$

4.15. Consider an integral operator T_K on $L^2([0,2\pi],dx)$ with integral kernel:

$$K(x,y) = \frac{1}{2\pi} \sum_{n\in\mathbb{Z}\setminus\{0\}} \frac{1}{n^2} e^{in(x-y)}\,.$$

Prove T_K is a compact, HS operator of trace class.

4.16. Consider the T_K of Exercise 4.15 and the differential operator

$$A := -\frac{d^2}{dx^2}\,,$$

defined on smooth maps over $[0,2\pi]$ that satisfy periodicity conditions (including all derivatives). What is $T_K A$?

Hint. Let $\mathbf{1}$ be the constant map 1 on $[0,2\pi]$, and $P_0 : f \mapsto (\frac{1}{2\pi}\int_0^{2\pi} f(x)dx)\mathbf{1}$ the orthogonal projector onto the space of constant maps in $L^2([0,2\pi],dx)$. Then $T_K A = I - P_0$.

4.17. Consider an integral operator T_s on $L^2([0,2\pi],dx)$ with kernel:

$$K_s(x,y) = \frac{1}{2\pi} \sum_{n\in\mathbb{N}\setminus\{0\}} \frac{1}{n^{2s}} e^{in(x-y)} .$$

Prove that for $Re\, s$ sufficiently large (how large?) the following indentity makes sense:

$$tr(T_s) = \zeta_R(2s) ,$$

where $\zeta_R = \zeta_R(s)$ is *Riemann's zeta function.*

4.18. If $B \in \mathfrak{B}(\mathsf{H})$, H Hilbert space, and there is a basis N such that $\sum_{u\in N} ||Bu|| < +\infty$, prove $B \in \mathfrak{B}_1(\mathsf{H})$.

Hint. Observe $|||B|\psi|| = ||B\psi||$ and $|(\psi||B|\psi)| \le ||\psi||\,|||B|\psi||$.

4.19. Consider the integral operator $T_K : L^2([a,b],dx) \to L^2([a,b],dx)$

$$(T_K\varphi)(x) = \int_a^x K(x,t)\varphi(t)dt .$$

where K is a measurable map such that there exists $M \in \mathbb{R}$ with $|K(x,t)| \le M$ if $x,t \in [a,b]$, $t \le x$. Prove

$$||T_K^n|| \le \frac{M^n(b-a)^n}{\sqrt{(n+1)!}}$$

and hence that there exists a positive integer n rendering T_K^n a contraction.

Solution. In the ensuing computations $\varphi \in L^2([a,b],dx)$ implies $\varphi \in L^1([a,b],dx)$ by the Cauchy-Schwarz inequality, because the constant map 1 is in $L^2([a,b],dx)$. By construction, if $\theta(z) = 1$ for $z \ge 0$ and $\theta(z) = 0$ otherwise, we have

$$|(T_K^n\varphi)(x)| = \int_a^b dx_1 \int_a^b dx_2 \cdots \int_a^b dx_n \theta(x-x_1)\theta(x_1-x_2)\cdots\theta(x_{n-1}-x_n)$$
$$\times K(x,x_1)K(x_1-x_2)\cdots K(x_{n-1},x_n)\varphi(x_n) .$$

Hence

$$|(T_K^n\varphi)(x)| \le M^n \int_{[a,b]^n} dx_1\cdots dx_n |\theta(x-x_1)\cdots\theta(x_{n-1}-x_n)|\,|\varphi(x_n)| .$$

Using Cauchy-Schwarz on $L^2([a,b]^n, dx_1\cdots dx_n)$, and from $\theta(z)^2 = \theta(z) = |\theta(z)|$, we have :

$$|(T_K^n\varphi)(x)| \le M^n \sqrt{\int_{[a,b]^n} dx_1\cdots dx_n \theta(x-x_1)\cdots\theta(x_{n-1}-x_n)}\, [b-a]^{(n-1)/2}||\varphi||_2 ,$$

i.e.

$$|(T_K^n\varphi)(x)| \le M^n \frac{(x-a)^{n/2}}{\sqrt{n!}} [b-a]^{(n-1)/2}||\varphi||_2 .$$

Consequently

$$||T_K^n \varphi||_2 \leq \frac{M^n (b-a)^n}{\sqrt{(n+1)!}} ||\varphi||_2 \,,$$

and so

$$||T_K^n|| \leq \frac{M^n (b-a)^n}{\sqrt{(n+1)!}} \,.$$

But since:

$$\lim_{n \to +\infty} \frac{M^n (b-a)^n}{\sqrt{(n+1)!}} \to 0 \quad \text{as } n \to +\infty,$$

for n large enough there will be a positive $\lambda < 1$ such that

$$||T_K^n \varphi - T_K^n \varphi'||_2 \leq \lambda ||\varphi - \varphi'||_2 \,,$$

making T_K^n a contraction operator.

5

Densely-defined unbounded operators on Hilbert spaces

Von Neumann had just about ended his lecture when a student stood up and in a vaguely abashed tone said he hadn't understood the final argument. Von Neumann answered: "Young man, in mathematics you don't understand things. You just get used to them."

David Wells

This chapter will extend the theory seen so far to unbounded operators that are not necessarily defined on the entire space.

In section one we will define, in particular, the *standard domain* of an operator built by composing operators with non-maximal domains. We will introduce *closed* and *closable operators*. Then we shall present the general notion of *adjoint operator* for unbounded and densely-defined operators, thus generalising the similar notion for bounded operators defined on the whole Hilbert space.

The second section will deal with generalisations of the notion of *self-adjoint* operator to the unbounded case. For this we will introduce *Hermitian*, *symmetric*, *essentially self-adjoint* and *self-adjoint* operators, and discuss their main properties. In particular we will define the *core* of an operator and the *deficiency index*.

Section three is entirely devoted to two examples of self-adjoint operators of the foremost importance in Quantum Mechanics, namely the operators *position* and *momentum* on the Hilbert space $L^2(\mathbb{R}^n, dx)$. We will study their mathematical properties and present several equivalent definitions.

In the final section we will discuss more advanced criteria to establish whether a symmetric operator admits self-adjoint extensions. In particular we will present *von Neumann's criterion* and *Nelson's criterion*. Technical instruments for this study are the *Cayley transform* and the notion of *analytic vector*: the latter, defined by Nelson, turned out to be crucial in the applications of operator theory to QM.

5.1 Unbounded operators with non-maximal domains

Let us introduce the theory of unbounded operators with non-maximal domains. *The domains of the operators under exam will always be vector subspaces of some ambient space*, and we will often consider the case of a dense subspace. Despite the operators of concern will not be bounded, all general definitions will reduce to the ones seen in earlier chapters if restricted to bounded operators.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_5,

5.1.1 Unbounded operators with non-maximal domains in normed spaces

The first definitions are completely general and do not require any Hilbert structure. A notion of graph was already given in Definition 2.94 for operators with maximal domain. The following definition extends 2.94 slightly. The notation $\mathsf{X} \oplus \mathsf{X}$, often used from now on, was introduced in 2.93(3).

Definition 5.1. *Let* X *be a vector space. We shall call a linear mapping*

$$T : D(T) \to \mathsf{X},$$

with $D(T) \subset \mathsf{X}$ *subspace in* X*, an* **operator on** X*, and* $D(T)$ *is called the* **domain** *of* T*. The* **graph** $G(T)$ *of the operator* T *is the subspace of* $\mathsf{X} \oplus \mathsf{X}$

$$G(T) := \{(x, Tx) \in \mathsf{X} \oplus \mathsf{X} \mid x \in D(T)\}.$$

If $\alpha \in \mathbb{C}$*, and* A*,* B *are operators on* X *with domains* $D(A)$*,* $D(B)$*, we define the following operators on* H*:*

(a) AB*, such that* $ABf := A(Bf)$ *on the* **standard domain***:*

$$D(AB) := \{f \in D(B) \,|\, Bf \in D(A)\}.$$

(b) $A + B$*, such that* $(A+B)f := Af + Bf$ *on the* **standard domain***:*

$$D(A+B) := D(A) \cap D(B).$$

(c) αA*, such that* $\alpha Af := \alpha(Af)$ *on the* **standard domain***:* $D(\alpha A) = D(A)$ *if* $\alpha \neq 0$*, and* $D(0A) := \mathsf{X}$*.*

Remarks 5.2. With these standard domains the usual *associative properties* of sum and product of operators hold: is A, B, C are operators on X:

$$A + (B + C) = (A + B) + C, \quad (AB)C = A(BC).$$

Distributive properties are weaker than expected (referring to Definition 5.3 below):

$$(A + B)C = AC + BC, \quad A(B + C) \supset AB + AC,$$

for it may happen that $(B + C)x \in D(A)$ even if Bx or Cx do not belong in $D(A)$. ■

The above notion of graph evidently coincides with the familiar graph of $T \in \mathfrak{L}(\mathsf{X})$, where the latter is nothing else than an operator on X with $D(T) = \mathsf{X}$.

Extensions of *closed* operators play a central role in the sequel. The first notion is straightforward.

Definition 5.3. *If* A *is an operator on the vector space* X*, an operator* B *on* X *is called an* **extension** *of* A*, written* $A \subset B$ *or, equivalently,* $B \supset A$*, if* $G(A) \subset G(B)$*.*

5.1.2 Closed and closable operators

Let us pass to *closed operators*, by enlarging the reach of Definition 2.94 to comprehend domains smaller than the whole space, and thereby introducing new concepts. We remind that if X is normed, the *product topology* on the Cartesian product $\mathsf{X} \times \mathsf{X}$ is the one whose open sets are $\varnothing$ and the unions of products of open balls $B_\delta(x) \times B_{\delta_1}(x_1)$ centred at $x, x_1 \in \mathsf{X}$ and with any radii $\delta, \delta_1 > 0$ (cf. Definition 1.10 and the case of normed spaces in Chapter 2.3.6).

Definition 5.4. *Let A be an operator on the normed space X.*

(a) *A is called* **closed** *if its graph is closed in the product topology of $\mathsf{X} \times \mathsf{X}$. So A is closed if and only if for any sequence $\{x_n\}_{n\in\mathbb{N}} \subset D(A)$ such that:*

(i) $x_n \to x \in \mathsf{X}$ as $n \to +\infty$ and

(ii) $Tx_n \to y \in \mathsf{X}$ as $n \to +\infty$,

it follows $x \in D(A)$ and $y = Tx$.

(b) *A is* **closable** *if the closure $\overline{G(A)}$ of its graph is the graph of a (necessarily closed) operator. The latter is denoted $\overline{A}$ and is called the* **closure** *of A.*

The next proposition characterises closable operators.

Proposition 5.5. *Let A be an operator on the normed space X. The following facts are equivalent:*

(i) *A is closable;*
(ii) *$\overline{G(A)}$ does not contain elements of type $(0,z)$, $z \neq 0$;*
(iii) *A admits closed extensions.*

Proof. (i) $\Leftrightarrow$ (ii). A is not closable iff there exist sequences in $D(A)$, say $\{x_n\}_{n\in\mathbb{N}}$ and $\{x'_n\}_{n\in\mathbb{N}}$, such that $x_n \to x \leftarrow x'_n$, but $Ax_n \to y \neq y' \leftarrow Ax'_n$. By linearity this is the same as saying there is a sequence $x''_n = x_n - x'_n \to 0$ such that $Ax''_n \to y - y' = z \neq 0$. In turn, this amounts to $\overline{G(A)}$ containing $(0,z) \neq (0,0)$.
(i) $\Leftrightarrow$ (iii). If A is closable, $\overline{A}$ is a closed extension of A. Conversely, if there is a closed extension B of A, there cannot be in $\overline{G(A)}$ elements of the kind $(0,z) \neq (0,0)$, for otherwise $G(B) = \overline{G(B)} \supset \overline{G(A)} \ni (0,z)$, since $A \subset B$ and B closed, and so B would not be linear as $B(0) \neq 0$. □

A useful general property of closable operators on Banach (hence Hilbert) spaces goes as follows.

Proposition 5.6. *Let X, Y be Banach spaces, $T \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$ and $A : D(A) \to \mathsf{Y}$ an operator on Y (in general unbounded and with $D(A)$ properly contained in Y). If*

(i) *A is closable;*
(ii) *$Ran(T) \subset D(A)$;*

then $AT \in \mathfrak{B}(\mathsf{X},\mathsf{Y})$.

Proof. As the closure of A extends A, $AT = \overline{A}T$ is well defined. Now it suffices to show $\overline{A}T : \mathsf{X} \to \mathsf{Y}$ is closed and invoke the closed graph theorem (Theorem 2.95) to conclude. To prove $\overline{A}T$ closed, assume $\mathsf{X} \ni x_n \to x \in \mathsf{X}$ and $(\overline{A}T)(x_n) \to y \in \mathsf{Y}$ as $n \to +\infty$. Then $Tx_n \to z \in \mathsf{Y}$, for T is continuous. As $\overline{A}$ is closed and $\overline{A}(Tx_n) \to y$, then $z \in D(\overline{A})$ and $\overline{A}z = y$. That is to say, $(\overline{A}T)(x) = y$. Therefore $\overline{A}T$ is closed by definition. □

5.1.3 The case of Hilbert spaces: the structure of $\mathsf{H} \oplus \mathsf{H}$ *and the* τ *operator*

Let us look at the situation in which $\mathsf{X} = \mathsf{H}$ is a Hilbert space with inner product $(\,|\,)$. There is a convenient way to define a Hilbert space structure on the direct sum $\mathsf{H} \oplus \mathsf{H}$, thus obtaining the **Hilbert sum** of H with itself. Topologically speaking, $\mathsf{H} \oplus \mathsf{H}$ can be endowed with the product topology of $\mathsf{H} \times \mathsf{H}$, used to define the closure of an operator. Let us show there is a natural Hilbert space structure on $\mathsf{H} \oplus \mathsf{H}$ inducing exactly the product topology by means of the inner product norm. Define the inner product on $\mathsf{H} \oplus \mathsf{H}$ by:

$$((x,x')|(y,y'))_{\mathsf{H}\oplus\mathsf{H}} := (x|y) + (x'|y') \quad \text{if } (x,x'),(y,y') \in \mathsf{H} \oplus \mathsf{H}. \tag{5.1}$$

In this way the two summands H of $\mathsf{H} \oplus \mathsf{H}$ are mutually orthogonal, so the sum is not just direct, but orthogonal. We claim $\mathsf{H} \oplus \mathsf{H}$ with the above product is indeed a Hilbert space. Since the induced norm $||\;||_{\mathsf{H}\oplus\mathsf{H}}$ satisfies:

$$||(z,z')||^2_{\mathsf{H}\oplus\mathsf{H}} = ||z||^2 + ||z'||^2 \quad \text{for any } (z,z') \in \mathsf{H} \oplus \mathsf{H}, \tag{5.2}$$

any Cauchy sequence $\{(x_n,x'_n)\}_{n\in\mathbb{N}} \subset \mathsf{H} \oplus \mathsf{H}$ for the norm $||\;||_{\mathsf{H}\oplus\mathsf{H}}$ determines Cauchy sequences in H: $\{x_n\}_{n\in\mathbb{N}}$ and $\{x'_n\}_{n\in\mathbb{N}}$. The latter converge to x and x' in H respectively. It is thus immediate to see $(x_n,x'_n) \to (x,x')$ as $n \to +\infty$ in norm $||\;||_{\mathsf{H}\oplus\mathsf{H}}$, by (5.2). Therefore $(\mathsf{H} \oplus \mathsf{H}, ||\;||_{\mathsf{H}\oplus\mathsf{H}})$ is complete. At last, the inner product on $\mathsf{H} \oplus \mathsf{H}$ induces the product topology on $\mathsf{H} \oplus \mathsf{H}$ as predicted. To this end, in analogy to the discussion of Section 2.3.6, it suffices to recall the inclusions

$$B_{\delta/2}(x) \times B_{\delta/2}(y) \subset B^{(\mathsf{H}\oplus\mathsf{H})}_{\delta}((x,y)) \subset B_{\delta}(x) \times B_{\delta}(y)$$

where $(x,y) \in \mathsf{H} \oplus \mathsf{H}$, $B^{(\mathsf{H}\oplus\mathsf{H})}_{\delta}((x,y))$ is the open ball in $\mathsf{H} \oplus \mathsf{H}$ with centre (x,y) and radius $\delta > 0$, $B_{\varepsilon}(z)$ the similar ball in H with centre z and radius $\varepsilon > 0$.

Remarks 5.7. Owing to this result, the notions of closed operator and closure in a Hilbert space (Definition 5.4) may refer equivalently to the topology induced by the inner product on $\mathsf{H} \oplus \mathsf{H}$. ■

A useful tool to prove results quickly is the bounded operator

$$\tau : \mathsf{H} \oplus \mathsf{H} \ni (x,y) \mapsto (-y,x) \in \mathsf{H} \oplus \mathsf{H}\,. \tag{5.3}$$

If * refers to the Hilbert space $\mathsf{H} \oplus \mathsf{H}$, we have:

$$\tau^* = \tau^{-1} = -\tau\,, \tag{5.4}$$

so, in particular, τ is *unitary* on $\mathsf{H} \oplus \mathsf{H}$. If $^\perp$ refers to $\mathsf{H} \oplus \mathsf{H}$, a direct computation shows τ and $^\perp$ *commute*: for any $F \subset \mathsf{H} \oplus \mathsf{H}$:

$$\tau(F^\perp) = (\tau(F))^\perp\,. \tag{5.5}$$

5.1.4 General properties of the Hermitian adjoint operator

We now pass to define the *adjoint* of an operator T on the Hilbert space H whose domain $D(T)$ is *dense* in H, in case T is, in general, not bounded.

We cannot use Riesz's theorem and must proceed differently. First of all let us define the *domain* $D(T^*)$ of the adjoint, bearing in mind we are aiming at obtaining $(T^*x|y) = (x|Ty)$ with $x \in D(T^*)$ and $y \in D(T)$. To this end

$$D(T^*) := \{x \in \mathsf{H} \mid \text{there exists } z_{T,x} \in \mathsf{H} \text{ with } (x|Ty) = (z_{T,x}|y) \text{ for any } y \in D(T)\}\ , \tag{5.6}$$

and later we will assume, when $x \in D(T^*)$, $T^*x := z_{T,x}$.

At any rate let us show definition (5.6) is well posed, first, and that it determines: (a) a *subspace* $D(T^*) \subset \mathsf{H}$, and (b) an *operator* $T^* : D(T^*) \ni x \mapsto z_{T,x}$.
(a) $D(T^*) \neq \varnothing$ for $0 \in D(T^*)$ if $z_{T,0} := 0$. Moreover, by (anti)linearity of the inner product and of T, if $x, x' \in D(T^*)$ and $\alpha, \beta \in \mathbb{C}$ then $\alpha x + \beta x' \in D(T^*)$; that is because $(\alpha x + \beta x'|Ty) = (z_{T,\alpha x+\beta x'}|y)$ if $z_{T,\alpha x+\beta x'} := \alpha z_{T,x} + \beta z_{T,x'}$. Hence $D(T^*)$ is a subspace.
(b) The assignment $D(T^*) \ni x \mapsto z_{T,x} =: T^*x$ defines a function, linear by construction as we saw above, only if any $x \in D(T^*)$ determines a *unique* $z_{T,x}$. We claim this is the case when $D(T^*)$ is dense, as we assumed. If $(z'_{T,x}|y) = (x|Ty) = (z_{T,x}|y)$ for any $y \in D(T)$, then $0 = (x|Ty) - (x|Ty) = (z_{T,x} - z'_{T,x}|y)$. Since $\overline{D(T)} = \mathsf{H}$, there is $\{y_n\}_{n\in\mathbb{N}} \subset D(T)$ with $y_n \to z_{T,x} - z'_{T,x}$. The inner product is continuous, so $(z_{T,x} - z'_{T,x}|y) = 0$ implies $||z_{T,x} - z'_{T,x}||^2 = 0$ and then $z_{T,x} = z'_{T,x}$.

Definition 5.8. *If T is an operator on the Hilbert space H with $\overline{D(T)} = \mathsf{H}$, the* **adjoint operator** *to T, denoted T^*, is the operator on H with domain*

$$D(T^*) := \{x \in \mathsf{H} \mid \textit{there exists } z_{T,x} \in \mathsf{H} \textit{ with } (x|Ty) = (z_{T,x}|y) \textit{ for any } y \in D(T)\}$$

and such that $T^* : x \mapsto z_{T,x}$.

Remark 5.9. (1) It is clear that by construction

$$(T^*x|y) = (x|Ty)\ , \quad \text{for any pair } (x,y) \in D(T^*) \times D(T)$$

as we wanted.
(2) If $T \in \mathfrak{B}(\mathsf{H})$, Definition 5.8 for T^* implies immediately $D(T^*) = \mathsf{H}$ by Riesz's Theorem 3.16. Hence:
Definitions 5.8 and 3.37 coincide for adjoints to operators in $\mathfrak{B}(\mathsf{H})$.
(3) If T is a densely-defined operator on the Hilbert space H, $D(T^*)$ is not automatically dense in H, so in general $(T^*)^*$ will *not* exist.
(4) If A, B are densely-defined operators on the Hilbert space H:

$$A \subset B \quad \Rightarrow \quad A^* \supset B^*. \tag{5.7}$$

The proof is straightforward from Definition 5.8.
(5) If A, B are operators on the Hilbert space H with dense domains, and $D(AB)$ is dense, then

$$B^*A^* \subset (AB)^*\ .$$

Furthermore, $B^*A^* = (AB)^*$ if $A \in \mathfrak{B}(\mathsf{H})$. ■

Theorem 5.10. *Let A be an operator on the Hilbert space H, and $\overline{D(A)} = \mathsf{H}$.*

(a) *A^* is closed and*

$$G(A^*) = \tau(G(A))^\perp . \tag{5.8}$$

(b) *A is closable $\Leftrightarrow$ $D(A^*)$ is dense, in which case*

$$A \subset \overline{A} = (A^*)^* .$$

(c) *$Ker(A^*) = [Ran(A)]^\perp$ and $Ker(A) \subset [Ran(A^*)]^\perp$, with equality if $D(A^*)$ is dense in H and A is closed.*

(d) *If A is closed then $\mathsf{H} \oplus \mathsf{H}$ decomposes orthogonally:*

$$\mathsf{H} \oplus \mathsf{H} = \tau(G(A)) \oplus G(A^*) . \tag{5.9}$$

Proof. (a) Writing $\tau(G(A))^\perp$ explicitly we find:

$$\tau(G(A))^\perp = \{(x,y) \in \mathsf{H} \oplus \mathsf{H} \mid -(x|Az) + (y|z) = 0 \text{ for any } z \in D(A)\} .$$

That is to say, $\tau(G(A))^\perp$ is the graph of A^* (so long as the operator is defined!) and (5.8) holds. $\tau(G(A))^\perp$ is closed by construction, being the orthogonal complement to a set (Theorem 3.13(a)), so A^* is closed.

(b) Consider the closure of the graph of A. Then we have $\overline{G(A)} = (G(A)^\perp)^\perp$ by Theorem 3.13. Since $\tau\tau = -I$, $S^\perp = -S^\perp$ for any set S, and because (5.5), (5.8) hold, we have:

$$\overline{G(A)} = -\tau(\,\tau(G(A))^\perp\,)^\perp = -\tau(G(A^*))^\perp = \tau(G(A^*))^\perp . \tag{5.10}$$

By Proposition 5.5, $\overline{G(A)}$ is the graph of an operator (the closure of A) iff $\overline{G(A)}$ does not contain elements $(0,z)$, $z \neq 0$. I.e., $\overline{G(A)}$ is *not* the graph of an operator iff there exists $z \neq 0$ such that $(0,z) \in \tau(G(A^*))^\perp$. More explicitly

$$\text{there exists } z \neq 0 \text{ such that } \quad 0 = ((0,z)|(-A^*x,x)) , \quad \text{for any } x \in D(A^*) .$$

Put equivalently, $\overline{G(A)}$ is *not* the graph of an operator iff $D(A^*)^\perp \neq \{0\}$, iff $D(A^*)$ is not dense in H. To sum up: $\overline{G(A)}$ is a graph if and only if $D(A^*)$ is dense in H.

If $D(A^*)$ is dense in H, then $(A^*)^*$ is defined, and by (5.10), (5.8) we have

$$\overline{G(A)} = \tau(G(A^*))^\perp = G((A^*)^*) .$$

Eventually, by definition of closure, $\overline{G(A)} = G(\overline{A})$. Substituting above:

$$G(\overline{A}) = G((A^*)^*) ,$$

so $\overline{A} = (A^*)^*$.

(c) The claims descend directly from

$$(A^*x|y) = (x|Ay) , \quad \text{for any pair } (x,y) \in D(A^*) \times D(A)$$

from the density of $D(A)$, and from (b) when A is closed.

(d) A being closed, $G(A)$ is closed and so $\tau(G(A))$ is closed because $\tau : \mathsf{H} \oplus \mathsf{H} \to \mathsf{H} \oplus \mathsf{H}$ is unitary. From (5.8) and Theorem 3.13(b–d) we have immediately (5.9). This ends the proof. □

Remarks 5.11. $D(A)$ dense and $\lambda \in \mathbb{C}$ imply $(A - \lambda I)^* = A^* - \overline{\lambda} I$, so the first equation in (c) has an immediate consequence:

$$Ker(A^* - \overline{\lambda} I) = [Ran(A - \lambda I)]^\perp ,$$

while the second equation yields:

$$Ker(A - \lambda I) \subset [Ran(A^* - \overline{\lambda} I)]^\perp .$$

In the rest of the book these relations will be used repeatedly. ■

5.2 Hermitian, symmetric, self-adjoint and essentially self-adjoint operators

We are now in a position to define in full generality *self-adjoint* operators and related objects.

Definition 5.12. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* $A : D(A) \to \mathsf{H}$ *an operator on* H.

(a) *A is called* **Hermitian** *if* $(Ax|y) = (x|Ay)$ *for any* $x, y \in D(A)$.
(b) *A is* **symmetric** *if:*

(i) *A is Hermitian;*
(ii) $D(A)$ *is dense.*

Therefore A is symmetric if and only if:

(i)' $\overline{D(A)} = \mathsf{H}$*;*
(ii)' $A \subset A^*$.

(c) *A is* **self-adjoint** *if:*

(i) $D(A)$ *is dense;*
(ii) $A = A^*$.

(d) *A is* **essentially self-adjoint** *if:*

(i) $D(A)$ *is dense;*
(ii) $D(A^*)$ *is dense;*
(iii) $A^* = (A^*)^*$ *(the adjoint is self-adjoint).*

Equivalently (by Theorem 5.10(b)), A is essentially self-adjoint if:

(i)' $D(A)$ *is dense;*
(ii)' *A is closable;*
(iii)' $A^* = \overline{A}$.

(e) *A is* **normal** *if* $A^*A = AA^*$, *where either side is defined on its standard domain.*

Remark 5.13. (1) A comment on (c) in Definition 5.12: by Theorem 5.10(a), every self-adjoint operator is automatically closed.
(2) It is worth noting that:

(i) *the definitions of Hermitian, symmetric, self-adjoint and essentially self-adjoint operator coincide when the operator's domain is the whole Hilbert space;*
(ii) the following important result holds.

Theorem 5.14 (Hellinger–Toeplitz). *A Hermitian operator with the entire Hilbert space as domain is necessarily bounded (and self-adjoint in the sense of Definition 3.51).*

Proof. Boundedness follows from Proposition 3.54(d). The operator is thus self-adjoint for Definition 3.9. □

(iii) *Bounded self-adjoint operators for Definition 3.51 are precisely the self-adjoint operators of Definition 5.12 with domain the whole space.*

(3) Essential self-adjointness is by far the most important property of the four for the applications to QM, on the following grounds. As we will explain soon, an essentially self-adjoint operator admits a unique self-adjoint extension, so it retains the information of a self-adjoint operator, essentially. For reasons we shall see later in the book, paramount operators in QM are self-adjoint; at the same time it is a fact that differential operators are the easiest to handle in QM. It often turns out that QM's differential operators become essentially self-adjoint if defined on suitable domains. Thus self-adjoint differential operators are on one hand easy to employ, on the other they carry, in essence, the information of self-adjoint operators useful in QM. Because of this we will indulge on certain features related to essential self-adjointness.
(4) Given an operator $A : D(A) \to \mathsf{H}$ on the Hilbert space H, $B \in \mathfrak{B}(\mathsf{H})$ **commutes** with A when:

$$BA \subset AB\,.$$

If the domain of A is dense and so A^* exists, it is easy to check that if $B \in \mathfrak{B}(\mathsf{H})$ commutes with A then B^* commutes with A^* (prove it as an exercise). Denote by $\{A\}'$ the **commutant** of $A : D(A) \to \mathsf{H}$:

$$\{A\}' := \{B \in \mathfrak{B}(\mathsf{H}) \mid BA \subset AB\}\,.$$

If $A = A^*$ then $\{A\}'$ is a closed *-subalgebra of $\mathfrak{B}(\mathsf{H})$ in the strong topology. Thus it is a *von Neumann algebra* (cf. Example 3.44(3)). The second commutant $\{A\}'' := \{\{A\}'\}'$ is still a von Neumann algebra, called the **von Neumann algebra** generated by A. ■

The following important, yet elementary, proposition will be frequently used without explicit mention. Its easy proof is left to the reader.

Proposition 5.15. *Let H_1, H_2 be Hilbert spaces and $U : \mathsf{H}_1 \to \mathsf{H}_2$ a unitary operator. If $A : D(A) \to \mathsf{H}_1$ is an operator on H_1, consider the operator on H_2*

$$A_2 : D(A_2) \to \mathsf{H}_2 \quad \text{with } A_2 := UA_1U^{-1} \text{ and } D(A_2) := UD(A_1).$$

Then A_2 is closable, or closed, Hermitian, symmetric, essentially self-adjoint, self-adjoint, or normal iff A_1 is alike.

Notation 5.16. From now on we shall also write $A^{**\cdots*}$ instead of $(((A^*)^*)\cdots)^*$. ■

Proposition 5.17. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* A *an operator on* H.
(a) *If* $D(A)$, $D(A^*)$, $D(A^{**})$ *are dense,*

$$A^* = \overline{A}^* = \overline{A^*} = A^{***}\,. \tag{5.11}$$

(b) *A is essentially self-adjoint* $\Leftrightarrow$ *$\overline{A}$ is self-adjoint.*
(c) *If A is self-adjoint, it is maximal symmetric, i.e. it has no proper symmetric extensions.*
(d) *If A is essentially self-adjoint, A admits only one self-adjoint extension: $\overline{A}$ (coinciding with A^*).*

Proof. (a) If $D(A)$, $D(A^*)$, $D(A^{**})$ are dense, the operators A^*, A^{**} and A^{***} exist. Moreover

$$\overline{A}^* = (A^{**})^* = A^{***} = (A^*)^{**} = \overline{A^*}$$

by Theorem 5.10(b). Since A^* is closed (by Theorem 5.10(a)) we have $\overline{A^*} = A^*$.
(b) If A is essentially self-adjoint, $\overline{A} = A^*$, and in particular $D(\overline{A}) = D(A^*)$ is dense. Compute the adjoint of $\overline{A}$ and recall Theorem 5.10(b): $\overline{A}^* = (A^*)^* = \overline{A}$, i.e. $\overline{A}$ is self-adjoint.
Vice versa, if $\overline{A}$ is self-adjoint, i.e. there exists $\overline{A}^* = \overline{A}$, then $D(A)$, $D(A^*)$, $D(A^{**})$ are dense and by part (a): $A^* = \overline{A^*} = \overline{A}^*$; hence $A^* = \overline{A}$, and A is essentially self-adjoint.
(c) Let A be self-adjoint and $A \subset B$, B symmetric. Taking adjoints gives $A^* \supset B^*$. But $B^* \supset B$ by symmetry, so

$$A \subset B \subset B^* \subset A^* = A\,,$$

and then $A = B = B^*$.
(d) Let $A^* = A^{**}$, $A \subset B$ with $B = B^*$. Taking the adjoint of $A \subset B$ we see that $B = B^* \subset A^*$. Taking the adjoint twice yields $A^{**} \subset B$, but then

$$B = B^* \subset A^* = A^{**} \subset B\,,$$

hence $B = A^{**}$. The latter coincides with $\overline{A}$ by Theorem 5.10(b). □

Now we discuss two crucial features that characterise self-adjoint and essentially self-adjoint operators.

Theorem 5.18. *Let A be a symmetric operator on the Hilbert space* H. *The following are equivalent:*
(a) *A is self-adjoint.*
(b) *A is closed and $Ker(A^* \pm iI) = \{0\}$.*
(c) *$Ran(A \pm iI) = \mathsf{H}$.*

Proof. (a) $\Rightarrow$ (b). If $A = A^*$, A is closed because A^* is. If $x \in Ker(A^* + iI)$, then $Ax = -ix$, so

$$i(x|x) = (Ax|x) = (x|Ax) = (x|-ix) = -i(x|x)\,,$$

whence $(x|x) = 0$ and $x = 0$.

The proof that $Ker(A^* - iI) = \{0\}$ is analogous.
(b) $\Rightarrow$ (c). By definition of adjoint we have (see Remark 5.11) $[Ran(A - iI)]^\perp = Ker(A^* + iI)$. Hence part (b) implies $Ran(A - iI)$ is dense in H. Now recall A is closed to show that, actually, $Ran(A - iI) = \mathsf{H}$. Fix $y \in \mathsf{H}$ arbitrarily and choose $\{x_n\}_{n\in\mathbb{N}} \subset D(A)$ so that $(A - iI)x_n \to y \in \mathsf{H}$. For $z \in D(A)$,

$$||(A - iI)z||^2 = ||Az||^2 + ||z||^2 \geq ||z||^2 ,$$

whence $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence and $x = \lim_{n\to+\infty} x_n$ exists. The closure of A forces $A - iI$ to be closed, so $(A - iI)x = y$ and then $Ran(A - iI) = \overline{Ran(A - iI)} = \mathsf{H}$. The proof of $Ker(A^* - iI) = \{0\}$ is similar.
(c) $\Rightarrow$ (a). Since $A \subset A^*$ by symmetry, it is enough to show $D(A^*) \subset D(A)$. Take $y \in D(A^*)$. Given that $Ran(A - iI) = \mathsf{H}$, there is a vector $x_- \in D(A)$ such that

$$(A - iI)x_- = (A^* - iI)y .$$

On $D(A)$ the operator A^* coincides with A and therefore, by the previous identity,

$$(A^* - iI)(y - x_-) = 0 .$$

But $Ker(A^* - iI) = Ran(A + iI)^\perp = \{0\}$, so $y = x_-$ and $y \in D(A)$. The proof of $Ran(A + iI)$ is analogous. □

Theorem 5.19. *Let A be a symmetric operator on the Hilbert space* H. *The following are equivalent:*
(a) *A is essentially self-adjoint.*
(b) $Ker(A^* \pm iI) = \{0\}$.
(c) $\overline{Ran(A \pm iI)} = \mathsf{H}$.

Proof. (a) $\Rightarrow$ (b). If A is essentially self-adjoint, then $A^* = A^{**}$ and A^* is self-adjoint (and closed). Applying Theorem 5.18 gives $Ker(A^{**} \pm iI) = \{0\}$ and so (b) holds, for $A^{**} = A^*$.
(b) $\Rightarrow$ (a). $A \subset A^*$ by assumption, and because $D(A)$ is dense so is $D(A^*)$. Consequently, Theorem 5.10(b) implies A is closable and $A \subset \overline{A} = A^{**}$ (in particular $D(A^{**}) = D(\overline{A}) \supset D(A)$ is dense). Therefore $A \subset A^*$ implies $\overline{A} \subset A^*$, and Proposition 5.17(a) tells $A^* = \overline{A}^*$. Overall, $\overline{A} \subset \overline{A}^*$, i.e. $\overline{A}$ is symmetric. Then we may apply Theorem 5.18 to $\overline{A}$, for this operator satisfies (b) in the theorem. We conclude $\overline{A}$ is self-adjoint. From Proposition 5.17(b) it follows A is essentially self-adjoint.
(b) $\Leftrightarrow$ (c). Since $Ran(A \pm iI)^\perp = Ker(A^* \mp iI)$ and $\overline{Ran(A \pm iI)} \oplus Ran(A \pm iI)^\perp = \mathsf{H}$, (b) and (c) are equivalent. □

To finish we present a useful notion for the applications: the *core* of an operator.

Definition 5.20. *Let A be a closable, densely-defined operator on the Hilbert space* H. *A dense subspace* $\mathsf{S} \subset D(A)$ *is a* **core** *of A if*

$$\overline{A\restriction_{\mathsf{S}}} = \overline{A} .$$

The next proposition is obvious, yet important.

Proposition 5.21. *If A is a self-adjoint operator on the Hilbert space* H*, a subspace* $\mathsf{S} \subset D(A)$ *is a core for A iff $A \restriction_{\mathsf{S}}$ is essentially self-adjoint.*

Proof. If $A \restriction_S$ is essentially self-adjoint, it admits a unique self-adjoint extension, which coincides with its closure by Proposition 5.17(d); in our case the extension necessarily coincides with A, which is self-adjoint. Hence $A \restriction_S$ is a core.

Conversely, if $A \restriction_S$ is a core, the closure of $A \restriction_S$ is self-adjoint because it coincides with the self-adjoint A. By Proposition 5.17(b) $A \restriction_S$ is essentially self-adjoint. □

5.3 Two major applications: the position operator and the momentum operator

To exemplify the formalism so far described we introduce and study the features of two self-adjoint operators of foremost relevance in QM, called *position operator* and *momentum operator*. Their physical meaning will be clarified in the second part of the book.

In the sequel we adopt the conventions and notations of Chapter 3.6, and $x = (x_1, \ldots, x_n)$ will be a generic point in $\mathbb{R}^n$.

5.3.1 The position operator

Definition 5.22. *Consider* $\mathsf{H} := L^2(\mathbb{R}^n, dx)$*, where dx is Lebesgue's measure on $\mathbb{R}^n$. If $i \in \{1, 2, \ldots, n\}$ is given, the operator on* H*:*

$$(X_i f)(x) = x_i f(x)\,, \tag{5.12}$$

with domain:

$$D(X_i) := \left\{ f \in L^2(\mathbb{R}^n, dx) \;\middle|\; \int_{\mathbb{R}^n} |x_i f(x)|^2 \, dx < +\infty \right\}, \tag{5.13}$$

is called ith **position operator**.

Proposition 5.23. *The operator X_i of Definition 5.22 satisfies these properties.*

(a) *X_i is self-adjoint.*
(b) *$\mathscr{D}(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$ are cores for X_i: $X_i = \overline{X_i \restriction_{\mathscr{D}(\mathbb{R}^n)}} = \overline{X_i \restriction_{\mathscr{S}(\mathbb{R}^n)}}$.*

Proof. (a) The domain of X_i is certainly dense in H for it contains the space $\mathscr{D}(\mathbb{R}^n)$ of smooth maps with compact support, and also the space of Schwartz functions $\mathscr{S}(\mathbb{R}^n)$ (see Notation 3.76), both of which are dense in $L^2(\mathbb{R}^n, dx)$. Thus X_i admits an adjoint. By definition of X_i and its domain we have $(g|X_i f) = (X_i g|f)$ if $f, g \in D(X_i)$. Consequently X_i is Hermitian and symmetric. We claim it is self-adjoint too. By symmetry

$X_i \subset X_i^*$, so it suffices to show $D(X_i^*) \subset D(X_i)$. Let us define the adjoint to X_i directly: $f \in D(X_i^*)$ if and only if there is $h \in L^2(\mathbb{R}^n, dx)$ (coinciding with $X_i^* f$ by definition) such that

$$\int_{\mathbb{R}^n} \overline{f(x)} x_i g(x) dx = \int_{\mathbb{R}^n} \overline{h(x)} g(x) dx \quad \text{for any } g \in D(X_i).$$

Since $D(X_i)$ is dense and

$$\int_{\mathbb{R}^n} [\overline{x_i f(x) - h(x)}] g(x) dx = 0 \quad \text{for any } g \in D(X_i),$$

we can also say $f \in L^2(\mathbb{R}^n, dx)$ *belongs to* $D(X_i^*)$ *if and only if* $x_i f(x) = h(x)$ *almost everywhere, with* $h \in L^2(\mathbb{R}^n, dx)$.

Hence $D(X_i^*)$ consists precisely of maps $f \in L^2(\mathbb{R}^n, dx)$ for which

$$\int_{\mathbb{R}^n} |x_i f(x)|^2 \, dx < +\infty,$$

and so $D(X_i^*) = D(X_i)$ and X_i is self-adjoint.
(b) If we define X_i as above, apart from restricting the domain to $\mathscr{D}(\mathbb{R}^n)$ or $\mathscr{S}(\mathbb{R}^n)$, the operator thus obtained is no longer self-adjoint, but stays symmetric. The adjoints to $X_i \restriction_{\mathscr{D}(\mathbb{R}^n)}$ and $X_i \restriction_{\mathscr{S}(\mathbb{R}^n)}$ both coincide with the X_i^* found above, for in the construction of X_i^* we only used that X_i is the operator that multiplies by x_i on a *dense* domain: whether this is $D(X_i)$ of (5.13), or a dense subspace, does not alter the result. If we set X_i as in (5.12) and (5.13) the adjoint X_i^* must satisfy $Ker(X_i^* \pm iI) = \{0\}$ by Theorem 5.18(b). But as X_i^* is the same as we get by restricting the domain of X_i to $\mathscr{D}(\mathbb{R}^n)$ or $\mathscr{S}(\mathbb{R}^n)$ by Theorem 5.19(b), the restricted X_i is essentially self-adjoint. Part (b) is now an immediate consequence of Proposition 5.21. □

5.3.2 The momentum operator

Let us introduce the *momentum operator*. Henceforth we make use of the definitions and conventions taken from Example 2.87, and retain Notation 3.76. First, though, we need a few definitions.

We say $f : \mathbb{R}^n \to \mathbb{C}$ is a **locally integrable** function on $\mathbb{R}^n$ if $f \cdot g \in L^1(\mathbb{R}^n, dx)$ for any map $g \in \mathscr{D}(\mathbb{R}^n)$.

Definition 5.24. *Let f be locally integrable. If α is a multi-index, $h : \mathbb{R}^n \to \mathbb{C}$ is the* **αth weak derivative** *of f, written $w\text{-}\partial^\alpha f = h$, if $h : \mathbb{R}^n \to \mathbb{C}$ is locally integrable and:*

$$\int_{\mathbb{R}^n} h(x) g(x) \, dx = (-1)^{|\alpha|} \int_{\mathbb{R}^n} f(x) \partial_x^\alpha g(x) \, dx \tag{5.14}$$

for any map $g \in \mathscr{D}(\mathbb{R}^n)$.

Remarks 5.25. (1) If it exists, a weak derivative is uniquely determined up to sets of zero measure: if h and h' are locally integrable (not necessarily $L^2(\mathbb{R}^n, dx)$, in which case the following is trivial) and satisfy (5.14), then:

$$\int_{\mathbb{R}^n} (h(x) - h'(x)) g(x) \, dx = 0 \quad \text{for any } g \in \mathscr{D}(\mathbb{R}^n). \tag{5.15}$$

This implies $h(x) - h'(x) = 0$ almost everywhere by the *Du Bois-Reymond lemma* [Vla81]:

Lemma 5.26 (Du Bois-Reymond). *If ϕ is locally integrable on $\mathbb{R}^n$ then ϕ is zero almost everywhere if and only if $\int_{\mathbb{R}^n} \phi(x) f(x)\, dx = 0$ for any $f \in \mathscr{D}(\mathbb{R}^n)$.*

(2) In case $f \in C^{|\alpha|}(\mathbb{R}^n)$, the αth weak derivative of f exists and coincides with the usual derivative (up to a zero-measure set). However, there are situations in which the ordinary derivative *does not exist* but the weak one is defined.
(3) $L^2(\mathbb{R}^n, dx)$ maps are locally integrable, for $\mathscr{D}(\mathbb{R}^n) \subset L^2(\mathbb{R}^n, dx)$ and $f \cdot g \in L^1$ if $f, g \in L^2$. ■

In order to define the *momentum operator* let us construct the operator A_j on $\mathsf{H} := L^2(\mathbb{R}^n, dx)$

$$(A_j f)(x) = -i\hbar \frac{\partial}{\partial x_j} f(x) \quad \text{with } D(A_j) := \mathscr{D}(\mathbb{R}^n),$$

where $\hbar$ is a positive constant (*Planck's constant*), whose precise value is irrelevant at present. By definition we have $(g|A_j f) = (A_j g|f)$ if $f, g \in D(A_j)$. Thus A_j is symmetric because $\overline{D(A_j)} = \mathsf{H}$. Let us find the adjoint to A_j, denoted $P_j := A_j^*$, directly from the definition. With $f \in D(A_j^*) = D(P_j)$ there must be $\phi \in L^2(\mathbb{R}^n, dx)$ (coinciding with $P_j f$ by definition) such that:

$$\int_{\mathbb{R}^n} \overline{\phi(x)} g(x) dx = -i\hbar \int_{\mathbb{R}^n} \overline{f(x)} \frac{\partial}{\partial x_j} g(x) dx\,, \quad \text{for any } g \in \mathscr{D}(\mathbb{R}^n)\,. \tag{5.16}$$

Conjugating the equation we rephrase (5.16) as follows: $f \in L^2(\mathbb{R}^n, dx)$ *belongs in $D(P_j)$ iff it admits weak derivative $\phi \in L^2(\mathbb{R}^n, dx)$.*

Definition 5.27. *Let $\mathsf{H} := L^2(\mathbb{R}^n, dx)$, dx being Lebesgue's measure on $\mathbb{R}^n$. Given $j \in \{1, 2, \ldots, n\}$, the operator on H:*

$$(P_j f)(x) = -i\hbar w\text{-}\frac{\partial}{\partial x_j} f(x)\,, \tag{5.17}$$

with domain:

$$D(P_j) := \left\{ f \in L^2(\mathbb{R}^n, dx) \;\middle|\; \textit{there exists } w\text{-}\tfrac{\partial}{\partial x_j} f \in L^2(\mathbb{R}^n, dx) \right\}, \tag{5.18}$$

is called jth **momentum operator**.

Remarks 5.28. If $n = 1$, $D(P_j)$ is identified with the *Sobolev space* $\mathsf{H}^1(\mathbb{R}, dx)$. ■

Proposition 5.29. *Let P_j be the operator of Definition 5.27.*

(a) *P_j is self-adjoint.*
(b) *$\mathscr{D}(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$ are cores of P_j. Therefore:*

$$(A_j f)(x) = -i\hbar \frac{\partial}{\partial x_j} f(x) \quad \textit{with } f \in D(A_j) := \mathscr{D}(\mathbb{R}^n)\,, \tag{5.19}$$

$$(A'_j f)(x) = -i\hbar \frac{\partial}{\partial x_j} f(x) \quad \textit{with } f \in D(A'_j) := \mathscr{S}(\mathbb{R}^n)\,, \tag{5.20}$$

are essentially self-adjoint and $\overline{A_j} = \overline{A'_j} = P_j$.

Proof. To simplify notations, in the sequel we will set $\hbar = 1$ (absorbing the constant $\hbar^{-1}$ in the unit of measure of the coordinate x_j), and denote by ∂_j the jth derivative and by w-∂_j the weak derivative. We want to prove $Ker(A_j^* \pm iI) = \{0\}$. This would imply, owing to Theorem 5.19, that A_j is essentially self-adjoint, i.e. $P_j = A_j^*$ is self-adjoint. The space $Ker(A_j^* \pm iI)$ consists of maps $f \in L^2(\mathbb{R}^n, dx)$ admitting weak derivative and such that $i(\text{w-}\partial_j f \pm f) = 0$. Let us consider the equation, with $f \in L^2(\mathbb{R}^n, dx)$:

$$\text{w-}\partial_j f \pm f = 0\,. \tag{5.21}$$

Multiplying by an exponential the above gives:

$$\text{w-}\partial_j \left(e^{\pm x_j} f\right) = 0\,. \tag{5.22}$$

So we can reduce to proving the following.

Lemma 5.30. *If $h : \mathbb{R}^n \to \mathbb{C}$ is locally integrable and*

$$w\text{-}\partial_j h = 0\,, \tag{5.23}$$

h coincides almost everywhere with a constant function in x_j.

Proof of Lemma 5.30. Without loss of generality we can suppose $j = 1$. We indicate by (x,y) the coordinates of $\mathbb{R}^n$, where x is x_1 and y the remaining $n-1$ components. Take h locally integrable satisfying (5.23). Explicitly:

$$\int_{\mathbb{R}^n} h(x,y) \frac{\partial}{\partial x} g(x,y) dx \otimes dy = 0\,, \quad \text{for any } g \in \mathscr{D}(\mathbb{R}^n)\,. \tag{5.24}$$

Pick $f \in \mathscr{D}(\mathbb{R}^n)$, and choose $a > 0$ large, so to have $supp(f) \subset [-a,a] \times [-a,a]^{n-1}$. Define $\chi \in \mathscr{D}(\mathbb{R})$ with $supp(\chi) = [-a,a]$ and $\int_{\mathbb{R}} \chi(x)dx = 1$. Then there is a map $g \in \mathscr{D}(\mathbb{R}^n)$ such that

$$\frac{\partial}{\partial x} g(x,y) = f(x,y) - \chi(x) \int_{\mathbb{R}} f(u,y) du\,.$$

In fact, it is enough to consider

$$g(x,y) := \int_{-\infty}^{x} f(u,y)du - \int_{-\infty}^{x} \chi(v)dv \int_{\mathbb{R}} f(u,y)du\,. \tag{5.25}$$

This map is smooth by construction, and its x-derivative coincides with:

$$f(x,y) - \chi(x) \int_{\mathbb{R}} f(u,y)du\,.$$

Moreover the support of g is bounded: if some coordinate satisfies $|y_k| > a$, then $f(u,y) = 0$ whichever u we have, so $g(x,y) = 0$ for any x. If $x < -a$ the first integral in (5.25) vanishes, and also the second one, for χ is supported in $[-a,a]$. Conversely, if $x > a$

$$g(x,y) := \int_{-\infty}^{+\infty} f(u,y)du - 1 \int_{\mathbb{R}} f(u,y)du = 0\,,$$

where we used $supp\chi = [-a,a]$ and $\int_{\mathbb{R}} \chi(x)dx = 1$. Altogether g vanishes outside $[-a,a] \times [-a,a]^{n-1}$. Inserting g in (5.24) and using the theorem of Fubini–Tonelli gives

$$\int_{\mathbb{R}^n} h(x,y)f(x,y)\,dx\otimes dy - \int_{\mathbb{R}^n}\left(\int_{\mathbb{R}} h(x,y)\chi(x)dx\right) f(u,y)\,du\otimes dy = 0\,.$$

Relabelling variables:

$$\int_{\mathbb{R}^n}\left\{h(x,y) - \left(\int_{\mathbb{R}} h(u,y)\chi(u)du\right)\right\} f(x,y)\,dx\otimes dy = 0\,, \tag{5.26}$$

f being arbitrary in $\mathscr{D}(\mathbb{R}^n)$. Notice that

$$(x,y)\mapsto k(y) := \int_{\mathbb{R}} h(u,y)\chi(u)du$$

is locally integrable on $\mathbb{R}^n$, because

$$(x,y,u)\mapsto f(x,y)h(u,y)\chi(u)$$

is integrabile on $\mathbb{R}^{n+1}$ for any $f\in\mathscr{D}(\mathbb{R}^n)$ (it is enough to see $|f(x,y)|\leq |f_1(x)||f_2(y)|$ for suitable f_1, f_2 in $\mathscr{D}(\mathbb{R})$ and $\mathscr{D}(\mathbb{R}^{n-1})$). Equation (5.26), valid for any $f\in\mathscr{D}(\mathbb{R}^n)$, implies immediately

$$h(x,y) - \int_{\mathbb{R}} h(u,y)\chi(u)du = 0$$

almost everywhere on $\mathbb{R}^n$ by the Du Bois-Reymond Lemma 5.26. That is to say

$$h(x,y) = k(y)$$

almost everywhere on $\mathbb{R}^n$. □

In the case under scrutiny the result implies that every solution to (5.21) must have the form $f(x) = e^{\pm x_j}h(x)$, where h does *not* depend on x_j. The theorem of Fubini–Tonelli then tells $\int_{\mathbb{R}^n}|f(x)|^2dx = ||h||^2_{L^2(\mathbb{R}^{n-1})}\int_{\mathbb{R}} e^{\pm 2x_j}dx_j$. Thus h must be null almost everywhere if, as required, $f\in L^2(\mathbb{R}^n,dx)$. Therefore $Ker(A_j^*\pm iI) = \{0\}$ and so $P_j = A_j^*$ is self-adjoint (A_j is essentially self-adjoint).

Because $\mathscr{S}(\mathbb{R}^n)\supset\mathscr{D}(\mathbb{R}^n)$ it is easy to see that A'_j is symmetric, that f admits generalised derivative if $f\in D({A'}^*_j)$, and that

$${A'}^*_j f = -i\text{w-}\frac{\partial}{\partial x_j}f\,.$$

Using the same procedure, if $f\in Ker({A'}^*_j\pm I)$ then $f=0$, so A'_j is essentially self-adjoint too. Since $A_j\subset A'_j$ and A_j is essentially self-adjoint, then ${A'_j}^{**} = \overline{A'_j} = {A_j}^{**} = \overline{A_j} = P_j$ by Proposition 5.17(d). □

There is another way to introduce the operator P_j. Consider the Fourier-Plancherel transform $\hat{\mathscr{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dk)$ seen in Chapter 3.6.

We define on $L^2(\mathbb{R}^n, dk)$ the analogue to X_j, which we call K_j (conventionally, the target space $\mathbb{R}^n$ of the Fourier-Plancherel transform has coordinates $(k_1, \ldots, k_n)$). Since $\hat{\mathscr{F}}$ is unitary, the operator $\hat{\mathscr{F}}^{-1} K_j \hat{\mathscr{F}}$ is self-adjoint if defined on the domain $\hat{\mathscr{F}}^{-1} D(K_j)$.

Proposition 5.31. *Let K_j be the jth position operator on the target space of the Fourier-Plancherel transform $\hat{\mathscr{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dk)$. Then*

$$P_j = \hbar \hat{\mathscr{F}}^{-1} K_j \hat{\mathscr{F}} .$$

Proof. It suffices to show the operators coincide on a domain where they are both essentially self-adjoint. So consider $\mathscr{S}(\mathbb{R}^n)$. From Chapter 3.6 we know the Fourier-Plancherel transform is the Fourier transform on this space, and $\hat{\mathscr{F}}(\mathscr{S}(\mathbb{R}^n)) = \mathscr{S}(\mathbb{R}^n)$. Moreover, the properties of the Fourier transform imply

$$-i\hbar \frac{\partial}{\partial x_j} f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ik\cdot x}\, \hbar k_j g(k)\, dk$$

provided $g \in \mathscr{S}(\mathbb{R}^n)$ and

$$f(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{ik\cdot x}\, g(k)\, dk.$$

Therefore

$$P_j \upharpoonright_{\mathscr{S}(\mathbb{R}^n)} = \hbar \hat{\mathscr{F}}^{-1} K_j \upharpoonright_{\mathscr{S}(\mathbb{R}^n)} \hat{\mathscr{F}} .$$

Notice K_j is essentially self-adjoint on $\mathscr{S}(\mathbb{R}^n)$ by Proposition 5.23, so also $\hbar \hat{\mathscr{F}}^{-1} K_j \upharpoonright_{\mathscr{S}(\mathbb{R}^n)} \hat{\mathscr{F}}$ is, on $\mathscr{S}(\mathbb{R}^n)$, because $\hat{\mathscr{F}}$ is unitary. Since $P_j \upharpoonright_{\mathscr{S}(\mathbb{R}^n)} = A'_j$ is essentially self-adjoint as well (Proposition 5.29), and since self-adjoint extensions of essentially self-adjoint operators are unique and coincide with the closure (Proposition 5.17(d)), we conclude

$$P_j = \overline{\hbar \hat{\mathscr{F}}^{-1} K_j \upharpoonright_{\mathscr{S}(\mathbb{R}^n)} \hat{\mathscr{F}}} = \hbar \hat{\mathscr{F}}^{-1} \overline{K_j \upharpoonright_{\mathscr{S}(\mathbb{R}^n)}} \hat{\mathscr{F}} = \hbar \hat{\mathscr{F}}^{-1} K_j \hat{\mathscr{F}} . \qquad \square$$

5.4 Existence and uniqueness criteria for self-adjoint extensions

In this remaining part of the chapter we discuss a few useful criteria to determine whether an operator admits self-adjoint extensions, and how many.

5.4.1 The Cayley transform and deficiency indices

One crucial technical tool to present the criteria is the so-called *Cayley transform*, introduced below. Before that, we generalise the notion of isometry (Definition 3.6) to operators with non-maximal domain.

Definition 5.32. *An operator $U : D(U) \to \mathsf{H}$, on the Hilbert space H, is an* **isometry** *if*

$$(Ux|Uy) = (x|y) \quad \textit{for any } x, y \in D(U).$$

Remark 5.33. (1) Clearly, if $D(U) = \mathsf{H}$ the above definition pins down isometric operators in the sense of Definition 3.51.
(2) By Proposition 3.8, the above definition of isometry is the same as demanding U satisfies $||Ux|| = ||x||$, for any $x \in D(U)$. ■

The transformation $\mathbb{R} \ni t \mapsto (t-i)(t+i)^{-1} \in \mathbb{C}$ is well known to define a bijection between the real line $\mathbb{R}$ and the unit circle in $\mathbb{C}$ minus the point 1. There is a similar correspondence that maps isometric operators to symmetric operators, called *Cayley transform*, of which we will study some properties.

Theorem 5.34. *Let H be a Hilbert space.*

(a) *If A is a symmetric operator on H:*

(i) *$A + iI$ is injective;*
(ii) *the* **Cayley transform of** *A:*

$$V := (A - iI)(A + iI)^{-1} : Ran(A + iI) \to \mathsf{H}\,, \tag{5.27}$$

is well defined;
(iii) *V is an isometry with $Ran(V) = Ran(A - iI)$.*

(b) *If (5.27) holds for some operator $A : D(A) \to \mathsf{H}$ on H with $(A + iI)$ injective, then:*

(i) *$I - V$ is injective;*
(ii) *$Ran(I - V) = D(A)$ and*

$$A := i(I + V)(I - V)^{-1}\,. \tag{5.28}$$

(c) *If A is symmetric on H, A is self-adjoint $\Leftrightarrow$ its Cayley transform V is unitary on H.*
(d) *If $V : \mathsf{H} \to \mathsf{H}$ is unitary and $I - V$ injective, then V is the Cayley transform of some self-adjoint operator on H.*

Proof. (a) A direct computation using the symmetry of A and the (anti)linearity of inner products proves that

$$||(A \pm iI)f||^2 = ||Af||^2 + ||f||^2 \tag{5.29}$$

if $f \in D(A)$. Therefore if $(A + iI)f = 0$ or $(A - iI)f = 0$, $f = 0$. The operators $A \pm iI$ are thus injective on $D(A)$, making V well defined from $D(V) := Ran(A + iI)$ to H. From (5.29)

$$||(A - iI)g|| = ||(A + iI)g||$$

for any $g \in D(A)$, Set $g = (A + iI)^{-1}h$, with $h \in Ran(A + iI)$; then

$$||Vh|| = ||(A - iI)(A + iI)^{-1}h|| = ||h||\,,$$

so V is an isometry with domain $D(V) = Ran(A+iI)$ and $Ran(V) = Ran(A-iI)$.
(b) $D(V)$ consists of vectors $g = (A+iI)f$ with $f \in D(A)$. Applying V to g gives $Vg = (A-iI)f$. Adding and subtracting $g = (A+iI)f$ produces

$$(I+V)g = 2Af\,, \tag{5.30}$$

$$(I-V)g = 2if\,. \tag{5.31}$$

(5.31) tells $(I-V)$ is injective, for if $(I-V)g = 0$ then $f = 0$ and so $g = (A+iI)f = 0$. Therefore if $f \in D(A)$ we can write

$$g = 2i(I-V)^{-1}f\,. \tag{5.32}$$

Furthermore, $Ran(I-V) = D(A)$ follows immediately from (5.31). Applying $(I+V)$ to equation (5.32) and using (5.30):

$$Af = i(I+V)(I-V)^{-1}f \quad \text{for any } f \in D(A).$$

(c) Suppose $A = A^*$. By Theorem 5.18 $Ran(A+iI) = Ran(A-iI) = \mathsf{H}$. Then part (a) implies V is an isometry from $Ran(A+iI) = \mathsf{H}$ onto $\mathsf{H} = Ran(A-iI)$. Hence V is a surjective isometry, i.e. a unitary operator.

Suppose now $V : \mathsf{H} \to \mathsf{H}$ is the unitary Cayley transform of A symmetric on H. By part (a) $Ran(A+iI) = Ran(A-iI) = \mathsf{H}$. This means $A = A^*$ by Theorem 5.18.
(d) It is enough to prove V is the Cayley transform of a symmetric operator. By part (c) this symmetric operator is self-adjoint. By assumptiuon there is a bijective map $z \mapsto x$, from $D(V) = \mathsf{H}$ to $Ran(I-V)$, given by $x := z - Vz$. Define $A : Ran(I-V) \to \mathsf{H}$ as

$$Ax := i(z+Vz)\,, \quad \text{if } x = z - Vz. \tag{5.33}$$

If $x, y \in D(A) = Ran(I-V)$, $x = z - Vz$ and $y = u - Vu$ for some $z, u \in D(V)$. But V is an isometry, so

$$(Ax|y) = i(z+Vz|u-Vu) = i(Vz|u) - i(z|Vu) = (z-Vz|iu-iVu) = (x|Ay)\,,$$

and A is Hermitian. To show it is symmetric, note $D(A) = Ran(I-V)$ is dense. In fact $[Ran(I-V)]^{\perp} = Ker(I-V^*)$. If $Ker(I-V^*)$ were not $\{0\}$, a non-zero $u \in \mathsf{H}$ would exists such that $V^*u = u$, and then applying V would give $u = Vu$. But that is not possible, for $I - V$ is injective by assumption.

To finish, we prove V is the Cayley transform of A. Equation (5.33) reads:

$$2iVz = Ax - ix\,, \quad 2iz = Ax + ix\,, \quad \text{if } z \in \mathsf{H}.$$

Hence $V(Ax+ix) = Ax - ix$ for $x \in D(A)$ and $\mathsf{H} = D(V) = Ran(A+iI)$. But then V is the Cayley transform of A because $V(A+iI) = A - iI$, and so

$$V = (A-iI)(A+iI)^{-1}\,.$$

This ends the proof. □

Remarks 5.35. From the statement and proof we infer that $Ker(A \pm iI) = \{0\}$ if A is symmetric. In general, though, this is *not* true! That $Ker(A^* \pm iI) = \{0\}$ is a very restrictive condition, equivalent to the essential self-adjointness of A (if A is symmetric) by Theorem 5.19. ■

Let us pass to the consequences of Theorem 5.34 concerning the existence of self-adjoint extensions of a symmetric operator. The first result introduces the so-called *deficiency indices*.

Theorem 5.36. *If A is a symmetric operator on the Hilbert space H, we call*

$$d_\pm(A) := dim\, Ker(A^* \pm iI)\,.$$

Deficiency indices *of A. Then:*

(a) *A admits self-adjoint extensions if and only if $d_+(A) = d_-(A)$.*
(b) *If $d_+(A) = d_-(A)$, there is a bijection between self-adjoint extensions of A and surjective isometries from $Ker(A^* - iI)$ to $Ker(A^* + iI)$.*

A admits as many self-adjoint extensions as the number of above surjective isometries. In particular, A admits more than one self-adjoint extensions whenever $d_+(A) = d_-(A) > 0$.

Remarks 5.37. Deficiency indices can be defined equivalently as:

$$d_\pm(A) := dim\,[Ran(A \mp iI)]^\perp\,,$$

because $Ker(A^* \pm iI) = [Ran(A \mp iI)]^\perp$. ■

Proof of Theorem 5.36. Consider the Cayley transform V of A. Suppose A has a self-adjoint extension B and let $U : \mathsf{H} \to \mathsf{H}$ be the Cayley transform of B. It is straighforward to see U is an extension of V using (5.27), recalling $(B+iI)^{-1}$ extends $(A+iI)^{-1}$ and $B - iI$ extends $A - iI$. Hence U maps $Ran(A+iI)$ into $Ran(A-iI)$. U is unitary, so $y \perp Ran(A+iI)$ iff $Uy \perp U(Ran(A+iI))$, that is to say $U([Ran(A+iI)]^\perp) = [Ran(A-iI)]^\perp$. By Theorem 5.10(c) this means $U(Ker(A^*+iI)) = Ker(A^*-iI)$. Since U is an isometry, $dim\,Ker(A^*+iI) = dim\,Ker(A^*-iI)$, i.e. $d_+(A) = d_-(A)$.
Let us show, conversely, that if $d_+ = d_-$ then A has a self-adjoint extension, not unique in case $d_+(A) = d_-(A) > 0$. Call V the Cayley transform of A. V is bounded, so Proposition 2.44 says we can extend V, uniquely, to an isometric operator $U : \overline{Ran(A+iI)} \to \overline{Ran(A-iI)}$. The same we can do for V^{-1}, extending it to a unique isometry from $\overline{Ran(A-iI)}$ to $\overline{Ran(A+iI)}$. By continuity this operator is $U^{-1} : \overline{Ran(A-iI)} \to \overline{Ran(A+iI)}$. Now recall $\overline{Ran(A \pm iI)}^\perp = [Ran(A \pm iI)]^\perp = Ker(A^* \mp iI)$.

Having assumed $d_+(A) = d_-(A)$, we can define a unitary operator $U_0 : Ker(A + iI) \to Ker(A - iI)$. Since

$$\mathsf{H} = \overline{Ran(A+iI)} \oplus Ker(A^*-iI) = \overline{Ran(A-iI)} \oplus Ker(A^*+iI)$$

is an orthogonal decomposition into closed spaces,

$$W : (x,y) := U \oplus U_0 :\mapsto (Ux, U_0 y)\,, \quad \text{with } x \in \overline{Ran(A+iI)} \text{ and } y \in Ker(A^* - iI)\,,$$

is a unitary operator on H. Moreover $I - W$ is injective. In fact, $Ker(I-W)$ consists of pairs $(x,y) \neq (0,0)$ with $Ux = x$ and $U_0 y = y$: the first condition has only the solution $x = 0$ because U is an isometry, and the second one implies $y \in Ker(A^* + iI) \cap Ker(A^* - iI)$, giving $y = 0$. Therefore Theorem 5.34(d) applies: W is the Cayley transform of a self-adjoint operator B. As W extends U, B is a self-adjoint extension of A. There are many choices for U_0 if $d_+(A) = d_-(A) > 0$, and each one produces a different self-adjoint extension of A.

We claim that the correspondence between self-adjoint extensions of A and surjective isometries U_0 is bijective. Parts (a) and (b) of Theorem 5.34 tell that two symmetric operators are distinct iff their Cayley transforms differ. So let us consider self-adjoint extensions of A. Each self-adjoint extension B has a unitary Cayley transform W extending U (as above) to a unitary operator on H. Since

$$U : \overline{Ran(A+iI)} \to \overline{Ran(A-iI)}$$

is surjective and isometric, and

$$\mathsf{H} = \overline{Ran(A+iI)} \oplus Ker(A^* - iI) = \overline{Ran(A-iI)} \oplus Ker(A^* + iI)\,,$$

and finally W extends U, the only possibility is that W determines a surjective isometry $U_0 : Ker(A^* - iI) \to Ker(A^* + iI)$. Two distinct self-adjoint extensions B, B' give distinct operators U_0, U_0', otherwise the Cayley transforms W, W' would coincide. Altogether, then, the map sending the self-adjoint extension B of A to the relative surjective isometry U_0 is 1-1. The map is also onto by what we said above, because the choice of U_0 determines a self-adjoint extension of A, i.e. the one with Cayley transform $W := U \oplus U_0$. □

Next comes the first important corollary to Theorem 5.36.

Theorem 5.38. *A symmetric operator A on the Hilbert space H is essentially self-adjoint if and only if it admits a unique self-adjoint extension.*

Proof. If A is essentially self-adjoint it has a unique self-adjoint extension by Proposition 5.17(d). Theorem 5.36 implies a symmetric operator A has self-adjoint extensions only if $d_+ = d_-$. In particular, if the extension is unique $d_+ = d_- = 0$. But then Theorem 5.19(b) forces A to be essentially self-adjoint. □

5.4.2 Von Neumann's criterion

Another consequence to Theorem 5.34, proved by von Neumann, establishes sufficient conditions for a symmetric operator to admit self-adjoint extensions. First we need two definitions.

Definition 5.39. *Let* X *and* X' *be* $\mathbb{C}$*-vector spaces with Hermitian inner products* $(\,|\,)_{\mathsf{X}}$ *and* $(\,|\,)_{\mathsf{X}'}$ *respectively. A surjective map* $V : \mathsf{X} \to \mathsf{X}'$ *is an* **antiunitary operator** *if:*

(a) *V is antilinear:* $V(\alpha x + \beta y) = \overline{\alpha} Vx + \overline{\beta} Vy$ *for any* $x, y \in \mathsf{X}$, $\alpha, \beta \in \mathbb{C}$.
(b) *V is anti-isometric:* $(Vx|Vy)_{\mathsf{X}'} = \overline{(x|y)_{\mathsf{X}}}$ *for any* $x, y \in \mathsf{X}$.

Remarks 5.40. Despite the complex conjugation in (b), note that $||Vz||_{\mathsf{X}'} = ||z||_{\mathsf{X}}$ for any $z \in \mathsf{X}$. Moreover, V is bijective. ■

Definition 5.41. *If* $(\mathsf{H}, (\,|\,))$ *is a Hilbert space, an antiunitary operator* $C : \mathsf{H} \to \mathsf{H}$ *is a* **conjugation operator**, *or* **conjugation**, *if it is involutive, i.e.* $CC = I$.

Remarks 5.42. A conjugation is defined on a complex vector space with Hermitian inner product. In general it is *different* from an involution in the sense of Definition 3.40, which – on the contrary – is a map defined on an algebra. ■

Theorem 5.43 (von Neumann). *Let A be a symmetric operator on the Hilbert space* H. *If there is a conjugation* $C : \mathsf{H} \to \mathsf{H}$ *such that* $C(D(A)) \subset D(A)$ *and*

$$AC = CA\,,$$

then A admits self-adjoint extensions.

Proof. To begin with, let us show $C(D(A^*)) \subset D(A^*)$ and $A^*C = CA^*$. By definition of adjoint $(A^*f|Cg) = (f|ACg)$ for any $f \in D(A^*)$ and $g \in D(A)$. As C is antiunitary, $(CCg|CA^*f) = (CACg|Cf)$. As C commutes with A and $CC = I$, we have $(g|CA^*f) = (Ag|Cf))$, i.e. $(CA^*f|g) = (Cf|Ag)$ for any $f \in D(A^*)$ and $g \in D(A)$. By definition of adjoint, this means $Cf \in D(A^*)$ if $f \in D(A^*)$ and $CA^*f = A^*Cf$.

Let us pass to existence, using Theorem 5.36. According to what we have just proved, if $A^*f = if$, applying C and using that C is antilinear and commutes with A^*, we obtain $A^*Cf = -iCf$. Thus C is a map (injective because it preserves norms) from $Ker(A^* - iI)$ to $Ker(A^* + iI)$. It is also onto, for if $A^*g = -ig$, with $f := Cg$ we have $A^*f = +if$. Applying C to f again (recall $CC = I$) gives $Cf = g$. Therefore C is a bijection from $Ker(A^* - iI)$ to $Ker(A^* + iI)$. That it is also anti-isometric, i.e. it preserves orthonormal vectors, implies it must map bases to bases. In particular it preserves their cardinality, so $d_+(A) = d_-(A)$. The claim now follows from Theorem 5.36. □

5.4.3 Nelson's criterion

We present, in conclusion, *Nelson's criterion*, that provides sufficient conditions for a symmetric operator to be essentially self-adjoint. Although one part of the final proof will be fully understandable only after going into spectral theory (Chapter 8 and 9), we believe it is better to present the result at this point. The reader might want to postpone the proof until he becomes familiar with the material of those chapters. First, though, a few preliminaries are in order.

Definition 5.44. *Let A be an operator on the Hilbert space* H.

(a) *A vector $\psi \in D(A)$ such that $A^n\psi \in D(A)$ for any $n \in \mathbb{N}$ ($A^0 := I$) is called a C^∞* **vector for** *A, and the vector subspace in H of C^∞ vectors for A is denoted $C^\infty(A)$.*
(b) *$\psi \in C^\infty(A)$ is an* **analytic vector for** *A, if:*

$$\sum_{n=0}^{+\infty} \frac{||A^n\psi||}{n!} t^n < +\infty \quad \textit{for some } t > 0.$$

(c) *A vector $\psi \in C^\infty(A)$ is a* **vector of uniqueness for** *A, if $A\restriction_{D_\psi}$ is essentially self-adjoint as operator on the Hilbert space $\mathsf{H}_\psi := \overline{D_\psi}$, where D_ψ is the (finite) span of $A^n\psi$, $n = 0, 1, 2\dots$ in H.*

If ψ is an analytic vector for A, the series:

$$\sum_{n=0}^{+\infty} \frac{||A^n\psi||}{n!} t^n$$

converges for some $t > 0$. Known results on convergence of power series guarantee the complex series

$$\sum_{n=0}^{+\infty} \frac{||A^n\psi||}{n!} z^n$$

converges absolutely and uniformly for any $z \in \mathbb{C}$, $|z| < t$. Furthermore, for $|z| < t$, also the series of derivatives of any order

$$\sum_{n=0}^{+\infty} \frac{||A^{n+p}\psi||}{n!} z^n$$

will converge, for any given $p = 1, 2, 3, \dots$. The last fact has an important consequence, easily proved, that comes from using the triangle inequality and the norm's homogeneity repeatedly.

Proposition 5.45. *If ψ is an analytic vector for A, operator on the Hilbert space H, every vector in D_ψ is analytic for A. More precisely, if the series*

$$\sum_{n=0}^{+\infty} \frac{||A^n\psi||}{n!} t^n ,$$

converges for $t > 0$ and $\phi \in D_\psi$, then the series

$$\sum_{n=0}^{+\infty} \frac{||A^n\phi||}{n!} s^n ,$$

converges for any $s \in \mathbb{C}$ with $|s| < t$.

We have a proposition, called *Nussbaum lemma*.

Proposition 5.46 ("Nussbaum lemma"). *Let A be a symmetric operator on the Hilbert space H. If $D(A)$ contains a set of vectors of uniqueness whose linear span is dense in H, A is essentially self-adjoint.*

Proof. By Theorem 5.19 it is enough to prove the spaces $Ran(A \pm iI)$ are dense. With our assumptions, given $\phi \in \mathsf{H}$ and $\varepsilon > 0$, there is a finite linear combination of vectors of uniqueness ψ_i with $||\phi - \sum_{i=1}^N \alpha_i \psi_i|| < \varepsilon/2$. Since $\psi_i \in \mathsf{H}_\psi$ and $A \upharpoonright_{D_\psi}$ is essentially self-adjoint on this Hilbert space, Theorem 5.19(c) implies there exist vectors $\eta_i \in \mathsf{H}_\psi$ with $||(A \upharpoonright_{D_\psi} + iI)\eta_i - \psi_i|| \leq \varepsilon/2 \left(\sum_{j=1}^N |\alpha_j|\right)^{-1}$. Setting $\eta := \sum_{i=1}^N \alpha_i \eta_i$ and $\psi := \sum_{i=1}^N \alpha_i \psi_i$, we have $\eta \in D(A)$ and

$$||(A+iI)\eta - \phi|| \leq ||(A \upharpoonright_{D_\psi} + iI)\eta - \psi|| + ||\phi - \psi|| < \varepsilon\,.$$

But $\varepsilon > 0$ is arbitrary, so $Ran(A+iI)$ is dense. The claim about $Ran(A-iI)$ is similar. So, A is essentially self-adjoint by Theorem 5.19(c). □

The above result prepares the ground for the proof of Nelson's theorem, which – as mentioned – needs the spectral theory of unbounded self-adjoint operators (this is logically independent from Nelson' criterion albeit presented in Chapter 8, 9).

Theorem 5.47 (Nelson's criterion). *Let A be a symmetric operator on the Hilbert space H. If $D(A)$ contains a set of analytic vectors for A whose finite span is dense in H, A is essentially self-adjoint.*

Proof. By Proposition 5.46 it suffices to show any analytic vector ψ_0 for A is a vector of uniqueness for A. Note $A \upharpoonright_{D_{\psi_0}}$ is surely a symmetric operator on $\mathsf{H}_{\psi_0} := \overline{D_{\psi_0}}$, because it is Hermitian and its domain is dense in H_{ψ_0}. Suppose $A \upharpoonright_{D_{\psi_0}}$ has a self-adjoint extension B in H_{ψ_0}. (NB: we are talking about self-adjoint extensions of $A \upharpoonright_{D_{\psi_0}}$ on the Hilbert space H_{ψ_0}, and *not* on H!) Let μ_ψ be the spectral measure of $\psi \in D_{\psi_0}$ for the PVM of the spectral expansion of B (cf. Theorems 8.50(c) and 9.10), defined by $\mu_\psi(E) := (\psi|P_E^{(B)}\psi)$ for any Borel set $E \subset \sigma(B) \subset \mathbb{R}$, where $P_E^{(B)}$ is the PVM associated to the self-adjoint operator B. As ψ_0 is analytic

$$\sum_{n=0}^{+\infty} \frac{||A^n \psi_0||}{n!} t_0^n < +\infty \quad \text{for some } t_0 > 0.$$

By Remark 5.42 then,

$$\sum_{n=0}^{+\infty} \frac{||A^n \psi||}{n!} t^n < +\infty \quad \text{for any } t < t_0,\, t \geq 0.$$

If $z \in \mathbb{C}$ and $0 < |z| < t_0$,

$$\sum_{n=0}^{+\infty} \int_{\sigma(B)} \left|\frac{z^n}{n!} x^n\right| d\mu_\psi(x) = \sum_{n=0}^{+\infty} \left|\frac{z^n}{n!}\right| \int_{\sigma(B)} 1 \cdot |x^n| d\mu_\psi(x)$$
$$\leq \sum_{n=0}^{+\infty} \frac{t_0^n}{n!} \left(\int_{\sigma(B)} d\mu_\psi(x)\right)^{1/2} \left(\int_{\sigma(B)} x^{2n} d\mu_\psi(x)\right)^{1/2}$$
$$= \sum_{n=0}^{+\infty} \frac{t_0^n}{n!} ||\psi|| \, ||B^n \psi|| = ||\psi|| \sum_{n=0}^{+\infty} \frac{t_0^n}{n!} ||A^n \psi|| < +\infty ,$$

where we used Theorem 9.4(c) for the spectral measure $P^{(B)}$ of the spectral expansion of B (spectral Theorem 9.10). The theorem of Fubini–Tonelli implies, for $0 < |z| < t_0$, that we can swap series and integral

$$\sum_{n=0}^{+\infty} \int_{\sigma(B)} \frac{z^n}{n!} x^n = \int_{\sigma(B)} \sum_{n=0}^{+\infty} \frac{z^n}{n!} x^n d\mu_\psi(x) .$$

Hence if $0 \leq |z| < t_0$ and if ψ belongs to the domain of e^{zB} (cf. Definition 9.11),

$$(\psi|e^{zB}\psi) = \int_{\sigma(B)} e^{zx} d\mu_\psi(x) = \int_{\sigma(B)} \sum_{n=0}^{+\infty} \frac{z^n}{n!} x^n d\mu_\psi(x) = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \int_{\sigma(B)} x^n d\mu_\psi(x)$$
$$= \sum_{n=0}^{+\infty} \frac{z^n}{n!} (\psi|A^n\psi) .$$

In particular this happens if $z = it$ (with $|t| < t_0$) because the domain of e^{itB} is the entire Hilbert space, by Corollary 9.5:

$$(\psi|e^{itB}\psi) = \sum_{n=0}^{+\infty} \frac{(it)^n}{n!} (\psi|A^n\psi) . \tag{5.34}$$

(Note the power series on the right converges on an open disc of radius t_0, i.e. it defines an analytic extension of the left-hand-side function when it is replaced by z in the disc, even if ψ does not belong to the domain of e^{zB}.) Now consider another self-adjoint extension of $A_{D_{\psi_0}}$, say B'. Arguing as before, for $|t| < t_0$ we have

$$(\psi|e^{itB'}\psi) = \sum_{n=0}^{+\infty} \frac{(it)^n}{n!} (\psi|A^n\psi) . \tag{5.35}$$

Then (5.34) and (5.35) imply, for any $|t| < t_0$ and any $\psi \in D_{\psi_0}$,

$$(\psi|(e^{itB} - e^{itB'})\psi) = 0 .$$

But D_{ψ_0} is dense in H_{ψ_0}, so (cf. Exercise 3.18) for any $|t| < t_0$:

$$e^{itB} = e^{itB'} .$$

Compute the strong derivatives at $t = 0$ and invoke Stone's theorem (Theorem 9.29), to the effect that

$$B = B'$$

(recall $t_0 > 0$ by assumption). Therefore all possible self-adjoint extensions of $A \upharpoonright_{D_\psi}$ are the same. We claim there exists at least one. Define $C : D_{\psi_0} \to \mathsf{H}_{\psi_0}$ by

$$C : \sum_{n=0}^{N} a_n A^n \psi_0 \mapsto \sum_{n=0}^{N} \overline{a_n} A^n \psi_0 \,.$$

Easily C extends to a unique conjugation operator on H_{ψ_0}, which we still call C (see Exercise 5.15). What is more, by construction $CA \upharpoonright_{D_{\psi_0}} = A \upharpoonright_{D_{\psi_0}} C$, so $A \upharpoonright_{D_{\psi_0}}$ admits self-adjoint extensions by Theorem 5.43.

Altogether, for any analytic vector ψ_0, $A \upharpoonright_{D_{\psi_0}}$ must be essentially self-adjoint on H_{ψ_0} by Theorem 5.38, because it is symmetric and it admits precisely one self-adjoint extension. We have thus proved that any analytic vector ψ_0 is a vector of uniqueness, ending the proof. □

Examples 5.48. (1) A typical example where von Neumann's criterion applies is an operator of chief importance in QM, namely $H := -\Delta + V$. Δ is the usual Laplacian on $\mathbb{R}^n$

$$\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \,,$$

and V is a locally integrable real-valued function.

By setting the domain of H to be $\mathscr{D}(\mathbb{R}^n)$, H becomes immediately a symmetric operator on $L^2(\mathbb{R}^n, dx)$. Define C as the antiunitary operator mapping $f \in L^2(\mathbb{R}^n, dx)$ to its pointwise-conjugate function. Clearly $CH = HC$, so H admits self-adjoint extensions. By choosing a specific V it is possible to prove H is essentially self-adjoint, as we will see at the end of Chapter 10.

(2) We know the operator $A_i := -i\frac{\partial}{\partial x_i}$ on $\mathscr{D}(\mathbb{R}^n)$ (see Proposition 5.29) is essentially self-adjoint, and as such it admits self-adjoint extensions. Is there a conjugation C that commutes with A_i? (Note that it might not exist). The conjugation operator of (1) does not commute with A_i despite its invariant subspace is the domain. Another possible conjugation is $C : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dx)$ defined by $(Cf)(x) := \overline{f(-x)}$ (almost everywhere) for any $f \in L^2(\mathbb{R}^n, dx)$. It is not hard to see $C(\mathscr{D}(\mathbb{R}^n)) \subset \mathscr{D}(\mathbb{R}^n)$ and $CA_i = A_i C$.

(3) Consider the Hilbert space $\mathsf{H} := L^2([0,1], dx)$ with Lebesgue measure dx, and let $A := i\frac{d}{dx}$ with domain given by maps in $C^1([0,1])$ (i.e. maps in $C^1((0,1))$ with finite first derivatives at 0 and 1) that also vanish at 0 and 1. The operator is Hermitian, as can be seen integrating by parts and because the maps annihilate boundary terms so they vanish at the endpoints of the integral. One can also verify the domain of A is dense, making A symmetric. Let us show A is *not* essentially self-adjoint. The condition that $g \in D(A^*)$ satisfies $A^* g = ig$ (resp. $A^* g = -ig$) reads:

$$\int_0^1 \overline{g(x)} \left[f'(x) + f(x) \right] dx = 0$$

(resp. $\int_0^1 \overline{g(x)}\,[f'(x) - f(x)]\,dx = 0$) for any $f \in D(A)$. Integrating by parts shows the map $g(x) = e^x$ ($g(x) := e^{-x}$) in $L^2([0,1],dx)$ solves the above equation for any f in $C^1([0,1])$ that vanishes at 0, 1. This latter condition is crucial when integrating by parts, because the exponential does *not* vanish at 0 and 1. By virtue of Theorem 5.19 A cannot be essentially self-adjoint.

Theorem 5.43 warrants, nonetheless, the existence of self-adjoint extensions. The antilinear transformation $C : L^2([0,1],dx) \to L^2([0,1],dx)$, $(Cf)(x) := \overline{f(1-x)}$ maps the space of C^1 functions on $[0,1]$ vanishing at the endpoints to itself. In addition

$$\left(Ci\frac{d}{dx}f\right)(x) = -i\frac{d}{d(1-x)}\overline{f(1-x)} = i\frac{d}{dx}\overline{f(1-x)} = i\frac{d}{dx}(Cf)(x)\,,$$

whence $CA = AC$. There must be more than one such extension, otherwise A would be essentially self-adjoint by Theorem 5.36, a contradiction.

The argument does not change if one takes domains akin to the above, in particular the space of C^∞ maps on $[0,1]$ that vanish at 0 and 1, or smooth maps on $[0,1]$ with compact support in $(0,1)$.

(4) Take $\mathsf{H} := L^2([0,1],dx)$ with the usual Lebesgue measure dx, and consider $A := -i\frac{d}{dx}$ defined on smooth periodic maps on $[0,1]$ with periodic derivatives of any order (of period 1). Integration by parts reveals A is Hermitian. The exponential maps $e_n(x) := e^{i2\pi nx}$, $x \in [0,1]$, $n \in \mathbb{Z}$, form a basis of H, as shown in Exercise 3.32(1). They are all defined on $D(A)$, and their span is dense in H, so $D(A)$ is dense in H and A is symmetric.

Every $f \in \mathsf{H}$ corresponds bijectively to the sequence of Fourier coefficients $\{f_n\}_{n\in\mathbb{Z}} \subset \ell^2(\mathbb{Z})$ of the expansion

$$f = \sum_{n\in\mathbb{Z}} f_n e_n\,.$$

This defines a unitary operator $U : \mathsf{H} \to \ell^2(\mathbb{Z})$ such that $U : f \mapsto \{f_n\}_{n\in\mathbb{Z}}$ (see Theorem 3.28). The elementary theory of Fourier series tells that $UD(A)U^{-1} =: D(A')$ is the space of sequences $\{f_n\}$ in $\ell^2(\mathbb{Z})$ such that $n^N|f_n| \to 0$, $n \to +\infty$, for any $N \in \mathbb{N}$. Moreover, if $A' := UAU^{-1}$ and $\{f_n\}_{n\in\mathbb{Z}} \in D(A')$, then

$$A' : \{f_n\}_{n\in\mathbb{Z}} \mapsto \{2\pi n f_n\}_{n\in\mathbb{Z}}\,.$$

Replicating the argument used for X_i in the proof of Proposition 5.23 allows to arrive at

$$D(A'^*) = \left\{ \{g_n\}_{n\in\mathbb{Z}} \subset \ell^2(\mathbb{Z}) \;\middle|\; \sum_{n\in\mathbb{Z}} |2\pi n g_n|^2 < +\infty \right\}.$$

On this domain

$$A'^* : \{f_n\} \mapsto \{2\pi n f_n\}\,.$$

As in Proposition 5.23 we can verify without problems that the adjoint to this operator is the operator itself. Hence A'^* is self-adjoint and A' essentially self-adjoint. As U is unitary, also A is essentially self-adjoint and the unique self-adjoint extension $\overline{A}$ satisfies $\overline{A} = U\overline{A'}U^{-1}$. (Fill in all details as exercise.)

(5) Example (4) can be settled in a much quicker way using Nelson's criterion. The domain of A contains the functions e_n whose span is dense in $\mathsf{H} := L^2([0,1], dx)$. Moreover $Ae_n = 2\pi n e_n$. Then

$$\sum_{k=0}^{+\infty} \frac{||A^k e_n||}{k!} t^k = \sum_{k=0}^{+\infty} \frac{(2\pi|n|)^k}{k!} (t)^k = e^{2\pi|n|t} < +\infty\,,$$

for any $t > 0$. As a consequence, A is essentially self-adjoint. ■

Exercises

5.1. Let A be an operator on the Hilbert space H with dense domain $D(A)$. Take $\alpha, \beta \in \mathbb{C}$ and consider the standard domain $D(\alpha A + \beta I) := D(A)$. Prove that

(i) $\alpha A + \beta I : D(\alpha A + \beta I) \to \mathsf{H}$ admits an adjoint and

$$(\alpha A + \beta I)^* = \overline{\alpha} A^* + \overline{\beta} I\,.$$

(ii) Assuming $\alpha, \beta \in \mathbb{R}$, $\alpha A + \beta B$ is Hermitian, symmetric, self-adjoint, essentially self-adjoint $\Leftrightarrow$ A is respectively Hermitian, symmetric, self-adjoint, essentially self-adjoint.

(iii) $\alpha A + \beta I$ is closable $\Leftrightarrow$ A is closable; in that case

$$\overline{\alpha A + \beta I} = \alpha \overline{A} + \beta I\,.$$

Hint. Apply directly the definitions.

5.2. Let A and B be densely-defined operators on the Hilbert space H. If $A + B : D(A) \cap D(B) \to \mathsf{H}$ is densely defined, prove

$$A^* + B^* \subset (A + B)^*\,.$$

5.3. Let A and B be densely-defined operators on the Hilbert space H. If the standard domain $D(AB)$ is densely defined, show $AB : D(AB) \to \mathsf{H}$ admits an adjoint and

$$B^* A^* \subset (AB)^*\,.$$

5.4. Let A be a densely-defined operator on the Hilbert space H and $L : \mathsf{H} \to \mathsf{H}$ a bounded operator. Using the definition of adjoint prove that

$$(LA)^* = A^* L^*\,.$$

Then show

$$(L + A)^* = L^* + A^*\,.$$

5.5. Let $A : D(A) \to \mathsf{H}$ be a symmetric operator on the Hilbert space H. Prove that A bijective implies self-adjoint. (Bear in mind that the inverse to a self-adjoint operator, if it exists, is self-adjoint. This falls out of the spectral theorem for unbounded self-adjoint operators, that we shall see later.)

Solution. If A is symmetric so is $A^{-1} : \mathsf{H} \to D(A)$. The latter is defined on the whole Hilbert space, so it is self-adjoint. Its inverse will, in turn, be self-adjoint.

5.6. In the sequel the commutant $\{A\}'$ of an operator A on H indicates the set of operators B in $\mathfrak{B}(\mathsf{H})$ such that $BA \subset AB$. Let $A : D(A) \to \mathsf{H}$ be an operator on the Hilbert space H. If $D(A)$ is dense and A closed, prove that $\{A\}' \cap \{A^*\}'$ is a strongly closed *-subalgebra in $\mathfrak{B}(\mathsf{H})$ with unit.

5.7. Prove Proposition 5.15.

5.8. Discuss whether and where the operator $-d^2/dx^2$ is Hermitian, symmetric, and essentially self-adjoint on the Hilbert space $\mathsf{H} = L^2([0,1], dx)$. Take as domain: (i) periodic maps in $C^\infty([0,1])$, or (ii) maps in $C^\infty([0,1])$ that vanish at the endpoints.

5.9. Prove that the family of self-adjoint extensions of the operator of Example 5.48(3) can be parametrised by one real number.

Hint. Consider the dimensions of $Ker(A^* \pm iI) = \overline{Ran(A \mp iI)}$.

5.10. Prove that

$$H := -\frac{d^2}{dx^2} + x^2$$

is essentially self-adjoint on $L^2(\mathbb{R}, dx)$ if $D(H) := \mathscr{S}(\mathbb{R})$.

Hint. Seek a basis of $L^2(\mathbb{R}, dx)$ of eigenvectors of H.

5.11. Consider the Laplace operator on $\mathbb{R}^n$ seen in Example 5.48(1):

$$\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} .$$

Prove explicitly Δ is essentially self-adjoint on the Schwartz space $\mathscr{S}(\mathbb{R}^n)$ inside $L^2(\mathbb{R}^n, dx)$, and as such it admits one self-adjoint extension $\overline{\Delta}$.

Then show that if $\widehat{\mathscr{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dk)$ is the Fourier-Plancherel transform (Chapter 3.6),

$$\left(\widehat{\mathscr{F}}\overline{\Delta}\widehat{\mathscr{F}}^{-1} f\right)(k) := -k^2 f(k) ,$$

where $k^2 = k_1^2 + k_2^2 + \ldots + k_n^2$, on the standard domain:

$$\left\{ f \in L^2(\mathbb{R}^n, dk) \,\middle|\, \int_{\mathbb{R}^n} k^4 |f(k)|^2 dk < +\infty \right\} .$$

Hint. The operator Δ is symmetric on $\mathscr{S}(\mathbb{R}^3)$, so we can use Theorem 5.19, verifying condition (b). Since the Schwartz space is invariant under the action of the unitary operator $\widehat{\mathscr{F}}$ given by the Fourier-Plancherel transform, as seen in Chapter 3.6, we may consider Theorem 5.19(b) for $\widehat{\Delta} := \widehat{\mathscr{F}}\Delta\widehat{\mathscr{F}}^{-1}$. This operator is essentially self-adjoint on $\mathscr{S}(\mathbb{R}^3)$ iff Δ is defined on $\mathscr{S}(\mathbb{R}^3)$. $\widehat{\Delta}$ acts on $\mathscr{S}(\mathbb{R}^n)$ by multiplication by $-k^2 = -(k_1^2 + k_2^2 + \ldots + k_n^2)$, giving a self-adjoint operator on the aforementioned standard domain. Condition (b) can then be verified easily for $\widehat{\Delta}^*$, by using the definition of adjoint plus the fact that $\mathscr{S}(\mathbb{R}^n) \supset \mathscr{D}(\mathbb{R}^n)$. The uniqueness of self-adjoint extensions for essentially self-adjoint operators proves the last part, because $\widehat{\mathscr{F}}$ is unitary.

5.12. Recall $\mathscr{D}(\mathbb{R}^n)$ denotes the space of smooth complex functions with compact support in $\mathbb{R}^n$. Referring to the previous esercise let $\overline{\Delta}$ be the unique self-adjoint extension of $\Delta : \mathscr{S}(\mathbb{R}^n) \to L^2(\mathbb{R}^n, dx)$. Prove $\mathscr{D}(\mathbb{R}^n)$ is a core for $\overline{\Delta}$. In other words show $\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)}$ is essentially self-adjoint and $\overline{\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)}} = \overline{\Delta}$.

Hint. It suffices to show $(\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^* = \overline{\Delta}$ (because that implies, by taking adjoints, $\overline{\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)}} = ((\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^*)^* = \overline{\Delta}^* = \overline{\Delta}$). For this identity note that if $\psi \in D((\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^*)$ then $(\Delta\varphi|\psi) = (\varphi|\psi')$, with $\psi' = (\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^*\psi \in L^2(\mathbb{R}^n, dx)$, for any $\varphi \in \mathscr{D}(\mathbb{R}^n)$. Passing to the Fourier-Plancherel transform it is immediate to see that $\widehat{\mathscr{F}}\psi' = -k^2\widehat{\mathscr{F}}\psi$, since $\widehat{\mathscr{F}}(\mathscr{D}(\mathbb{R}^n))$ is dense in $L^2(\mathbb{R}^n, dk)$. Therefore we obtained $\psi \in D(\overline{\Delta})$ and $\psi' = \overline{\Delta}\psi$, and so $(\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^* \subset \overline{\Delta}$. Now suppose, conversely, $\psi \in D(\overline{\Delta})$. Using the Fourier-Plancherel transform gives $-k^2\widehat{\mathscr{F}}\psi \in L^2(\mathbb{R}^n, dk)$, and for any $\varphi \in \mathscr{D}(\mathbb{R}^n)$ we may write $(\Delta\varphi|\psi) = -\int dk k^2 \overline{(\widehat{\mathscr{F}}\varphi)}\widehat{\mathscr{F}}\psi = -\int dk \overline{(\widehat{\mathscr{F}}\varphi)} k^2 \widehat{\mathscr{F}}\psi = (\varphi|\overline{\Delta}\psi)$. By definition of adjoint we found $\psi \in D((\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^*)$ and $(\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^*\psi = \overline{\Delta}\psi$. Thus we have the other inclusion, $(\Delta\restriction_{\mathscr{D}(\mathbb{R}^n)})^* \supset \overline{\Delta}$.

5.13. Recall that if A is an operator on H, the commutant $\{A\}'$ is the set of operators B of $\mathfrak{B}(\mathsf{H})$ such that $BA \subset AB$.

Let $A : D(A) \to \mathsf{H}$ be self-adjoint and T its Cayley transform. Prove that the von Neumann algebra $(\{A\}')'$ generated by A coincides with the von Neumann algebra $(\{T\}')'$ generated by $\{T\}$ (cf. Example 3.44(3)).

5.14. Prove Proposition 5.45.

5.15. Take a symmetric operator $A : D(A) \to \mathsf{H}$ on the Hilbert space H and suppose $\psi \in C^\infty(A)$ is such that finite linear combinations of $A^n\psi$, $n \in \mathbb{N}$, are dense in H. Prove that for any chosen $N = 0, 1, 2, \ldots$ and $a_n \in \mathbb{C}$,

$$C : \sum_{n=0}^{N} a_n A^n \psi \mapsto \sum_{n=0}^{N} \overline{a_n} A^n \psi$$

determines a conjugation operator $C : \mathsf{H} \to \mathsf{H}$ (Definition 5.39).

Outline. The first thing to prove is C is well defined as a map, i.e. if $\sum_{n=0}^{N} a_n A^n \psi = \sum_{n=0}^{N_1} a'_n A^n \psi$ then $\sum_{n=0}^{N} \overline{a_n} A^n \psi = \sum_{n=0}^{N_1} \overline{a'_n} A^n \psi$. For that it is enough to observe that if $\Psi = \sum_{m=0}^{M} b_m A^m \psi$, then

$$\left(\psi \left| \sum_{n=0}^{N} a_n A^n \Psi \right.\right) = \left(\psi \left| \sum_{n=0}^{N_1} a'_n A^n \Psi \right.\right) \text{ so that } \left(\left. \sum_{n=0}^{N} \overline{a_n} A^n \psi \right| \Psi \right) = \left(\left. \sum_{n=0}^{N_1} \overline{a'_n} A^n \psi \right| \Psi \right).$$

Since the vectors Ψ are dense, $\sum_{n=0}^{N} \overline{a_n} A^n \psi = \sum_{n=0}^{N_1} \overline{a'_n} A^n \psi$, as required. By construction one verifies that if Ψ and Ψ' are as above then $(C\Psi|C\Psi') = \overline{(\Psi|\Psi')}$. Since the Ψ are dense in H and $||C\Psi|| = ||\Psi||$, it is straightforward to see C extends to H by continuity and antilinearity. This latter antilinear operator satisfies $(C\Psi|C\Psi') = \overline{(\Psi|\Psi')}$ on H and is onto, as one obtains by extending by continuity the equation $CC\Psi = I\Psi$.

6

Phenomenology of quantum systems and Wave Mechanics: an overview

Two are the possible outcomes: if the result confirms the hypotheses, you only took a measurement. But if the result contradicts the assumptions, then you made a discovery.
Enrico Fermi

In this chapter we try to arouse in the user a naïve feel about what the terms *quantum system* and *quantum phenomenology* underlie. The more mathematically-oriented reader, perhaps not so interested in the genesis of QM's notions in physics, may skip the sections following the first. From sections two, in fact, we will mention a number of experimental facts, and briefly review the theoretical "proto-quantum" methods that led to the formulation of *wave mechanics* first, and then to proper QM. Many of the physics details can be found in [Mes99, CCP82]. We shall eschew discussing important steps in this historical development, e.g. atomic spectroscopy, models of the atom (Rutherford's, Bohr's, Bohr-Sommerfeld's), the Franck-Hertz experiment, for which we recommend physics textbooks (e.g. [Mes99, CCP82]). This overview is meant to shed light on the basic theoretical model behind QM, developed in ensuing chapters.

Notation 6.1. As customary in physics texts, in this and sometimes other chapters too, we will denote vectors in three-space (identified with $\mathbb{R}^3$ once Cartesian coordinate have been fixed on a frame system), by boldface letters, e.g. $\mathbf{x}$. In the same way, Lebesgue's measure on $\mathbb{R}^3$ will be written $d^3\mathbf{x}$. ■

6.1 General principles of quantum systems

We use the term *physical system* loosely, as a manner of speaking. It is quite hard to define, from a physical point of view, what a *quantum system* actually is. We can start by saying that rather than talking of a *physical quantum state* it may be more suitable to discuss a physical system with *quantum behaviour*, thus distinguishing these systems more by their phenomenological/experimental aspects than by theoretical ones. Within the theoretical formulation of QM there is no clear border line separating classical systems from quantum systems. The divide is forced artificially; demarcation issues are very debated, today more than in the past, and the object of intense theoretical and experimental research work.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_6, © Springer-Verlag Italia 2013

Generically speaking we can talk of quantum nature for *microphysical* systems, i.e. *molecules*, *atoms*, *nuclei* and *subatomic particles* when taken singularly or in small numbers. Physical systems made of several copies of those subsystems (like *crystals*) can show quantum behaviour. Certain macroscopic systems behave in a typical quantum way only under specific circumstances that are hard to achieve (e.g. *Bose-Einstein condensates*, or *L.A.S.E.R.*). There is a way to refine slightly the rough distinction between the above micro- and macrosystems. We may say that when any physical system behaves in a quantum manner, the system's characteristic action, i.e. the number of physical dimensions of *Energy* × *Time* (equivalently *Momentum* × *Length* or *Angular Momentum*), obtained by combining suitably the characteristic physical dimensions (mass, speed, length, ...) in the processes examined, is of order smaller than **Planck's constant**:

$$h = 6.6262 \cdot 10^{-34} Js.$$

Planck's constant, and the word *quantum* labelling Quantum Mechanics, were first introduced by Planck in a 1900 work on the *black-body theory*, to deal with the issue of the theoretically-infinite total energy of a physical system consisting in the electromagnetic radiation in thermodynamic equilibrium with the walls of an enclosure at fixed temperature. Planck's theoretical prediction, later proved to be correct, was that the radiation could exchange with the walls quantities of energy proportional to the frequencies of atomic oscillators in the walls, whose universal factor is the aforementioned *Planck constant*. These packets of energy were called by the Latin name *quanta*. If we return to the criterion for distinguishing quantum from classical systems using h, let us for instance look at an electron orbiting around a hydrogen nucleus. A characteristic action of the electron is, for example, the product of its mass ($\sim 9 \cdot 10^{-31} Kg$), the estimated orbiting speed ($\sim 10^6 m/s$) and the value of Bohr's radius for the hydrogen atom ($\sim 5 \cdot 10^{-11} m$). This gives $4.5 \cdot 10^{-35} Js$, smaller than Planck's constant. Thus one would expect the hydrogen electron behave in a quantum manner, and this is indeed the case. A similar computation can be carried out for macroscopic systems like a pendulum, of mass a few grams and length one centimetre, swinging under gravity's pull. A characteristic action for this can be the maximum kinetic energy times the period of oscillation, and the value is several orders of magnitude bigger than h.

Remarks 6.2. The set of values taken by physical quantities, like energy, that charaterise a quantum systems's state is called *spectrum* in the jargon; one of the peculiarities of quantum systems is that their spectrum is usually not like the spectrum measured on comparable macroscopic systems. Sometimes the difference is astonishing, for one passes from a *continuous* spectrum of possible values in the classical case, to a *discrete* spectrum in quantum situations. It is important to point out that a discrete spectrum of values for a given physical quantity is *not* essential in QM: there are quantum quantities in QM with a continuous spectrum. This misunderstanding is the cause – or the consequence, at times – of a recurrent oversimplified interpretation, of the word *quantum* in QM. ■

6.2 Particle aspects of electromagnetic waves

Under special experimental circumstances electromagnetic waves (hence light as well) reveal a conduct that is typical of *collections of particles*. The mathematical description of these irregular, from a classical viewpoint, behaviours involves Planck's constant. In this respect we can cite two examples of classically deviant behaviour: the *photoelectric effect* and *Compton's effect*. In the infant stages of QM's development these played a fundamental role in the construction of the proto-quantum models meant to explain them.

6.2.1 The photoelectric effect

The *photoelectric effect* describes the emission of electrons (a current) by a metal irradiated with an electromagnetic wave, a phenomenon known since the first half of the XIX century. Some of its features remained with no explanation within the classical theory of interactions between matter and electromagnetic waves for a long time [Mes99, CCP82]. One conundrum, in particular, was to establish the minimum frequency of light below which no emission is measured when beaming a metal, a threshold that depends on the metal. At the time it did not seem possible to explain why the emission started instantaneously once that particular value was exceeded. According to the classical theory electronic emission should be detected independently of the frequency used, as long as enough time lapses to allow the metal's electrons to absorb sufficient energy to bond with atoms.

In 1905 A. Einstein proposed a very daring model to account for the strange properties of the photoelettric effect,[1] with an outstanding precision if compared to experimental data. Following Planck, Einstein's point was that a *monochromatic* electromagnetic wave, i.e. one with fixed frequency ν, was in reality made of particles of matter, called *light quanta*, each having energy prescribed by Planck's radiation formula:

$$E = h\nu\,. \tag{6.1}$$

The total energy of the electromagnetic wave in this model would then be the sum of the energies of the single quanta of light "associated" to the wave.

All this was, and still is, in contrast with classical electromagnetism, according to which an electromagnetic wave is a continuous system whose energy is proportional to the wave's *amplitude* rather than its *frequency*. What happened in the photoelectric effect, Einstein said, was that by irradiating the metal with a monochromatic wave each energy packet associated to the wave was absorbed by an electron in the metal, and transformed into kinetic energy. To justify the experimental evidence Einstein postulated, more precisely, that the packet could be either absorbed *completely* or *not at all*, without intermediate possibilities. If, and only if, the energy of the quantum was equal to, or bigger than, the electron's bonding energy E_0 to the metal (which depends on the metal, and can be measured irrespective of the photoelectric effect),

[1] Einstein was awarded the Nobel Prize in Physics for this work.

would the electron be *instantaneously* emitted, transforming the energetic excess of the absorbed quantum into kinetic energy. The frequency $\nu_0 := E_0/h$ would thus detect the threshold observed experimentally. This conjecture turned out to match the experimental data *perfectly*.

6.2.2 The Compton effect

The first observation and study of the *Compton effect* dates back 1923. It concerns the scattering of monochromatic electromagnetic waves of extremely high frequency – X rays ($> 10^{17} Hz$) and γ rays ($> 10^{18} Hz$) caused by matter (gases, fluids and solids). It is useful to remind that monochromatic electromagnetic waves have both a fixed frequency ν and a fixed wavelength λ, whose product is constant to the speed of light, $\nu\lambda = c$, irrespective of the kind of wave. Hence in the sequel we will talk about the wavelength of monochromatic waves. Simplifying as much as possible, the Compton effect consists in the following. Suppose we irradiate a substance (the *obstacle*) with a plane monochromatic electromagnetic wave that moves along the direction z with given wavelength λ. Then we observe a wave scattered by the obstacle and decomposed into several components (i.e. several wavelengths or frequencies). One component is scattered in every direction and has the same wavelength of the incoming wave. Every other component has a wavelength $\lambda(\theta)$, depending on the angle θ of observation, that is slightly bigger than λ. If we define θ to be the angle between the the wave's incoming direction z and the outgoing direction (after hitting the obstacle, wavelength $\lambda(\theta)$), we have the equation

$$\lambda(\theta) = \lambda + f(1 - cos\theta), \tag{6.2}$$

where the constant f has the dimension of a length and comes from the experimental data. Its value [2] is $f = 0.024(\pm 0.001)$ Å. The region around the z-axis is isotropic. Classical electromagnetic theory was, and still is, inadaguate to explain this phenomenon. However, as Compton proved, the effect could be clarified by making three assumptions, all incompatible with the classical theory but in agreement with Planck and Einstein's speculations about light quanta.

(a) The electromagnetic wave is made of particles that carry energy according to Planck's equation (6.1), exactly as Einstein predicted.
(b) Each quantum of light possesses a momentum

$$\mathbf{p} := \hbar\mathbf{k}, \tag{6.3}$$

where $\mathbf{k}$ is the *wavevector* of the wave associated to the quantum (see below).
(c) Quanta interact, via collisions (in relativistic regime, in general), with the outermost atoms and electrons of the obstacle, satisfying the conservation laws of momentum and energy.

[2] Recall 1 Å$= 10^{-10} m$.

Remarks 6.3. Concerning assumption (c), let us stress that the wavevector **k** associated to a plane monochromatic wave has, by definition, the same direction and orientation of the travelling wave, and its modulus is $2\pi/\lambda$, where λ is the wavelength. Equivalently, if ν denotes the frequency,

$$|\mathbf{k}| = 2\pi/\lambda = 2\pi\nu/c\,, \tag{6.4}$$

where we have used the well-known relationship

$$\nu\lambda = c \tag{6.5}$$

for monochromatic electromagnetic waves and $c = 2.99792 \cdot 10^8 m/s$ is the speed of light in vacuum. ■

The interested reader will find below a few more details. Under the assumptions made above, the energy and momentum conservation laws to be used in a relativistic regime, i.e. when (certain) speeds approach c, read as follow:

$$m_e c^2 + h\nu = \frac{m_e c^2}{\sqrt{1 - v^2/c^2}} + h\nu(\theta)\,, \tag{6.6}$$

$$\hbar\mathbf{k} = \frac{m_e \mathbf{v}}{\sqrt{1 - v^2/c^2}} + \hbar\mathbf{k}(\theta)\,. \tag{6.7}$$

On the left we have the quantities before the collision, on the right after the interaction. $m_e = 9.1096 \cdot 10^{-31} Kg$ is the electronic mass. The electron is thought of as at rest prior to colliding with the quatum of light. After the collision the quantum is scattered in the direction θ, while the electron travels at velocity **v**. The wavevector before the collision, **k**, is parallel to z (this direction is arbitrary, but fixed), while the wavevector of the quantum after the collision, $\mathbf{k}(\theta)$, forms an angle θ with z.

By (6.7) and by definition of wavevector,

$$\frac{m_e^2 c^2}{1 - v^2/c^2} = \frac{h^2\nu^2}{c^2} + \frac{h^2\nu(\theta)^2}{c^2} - 2\frac{h\nu}{c}\frac{h\nu(\theta)}{c}\cos\theta\,.$$

Eliminating v from this and (6.6) gives

$$\nu(\theta) = \nu - \frac{h\nu\nu(\theta)}{m_e c^2}\,(1 - cos\theta)\,. \tag{6.8}$$

Because of (6.1) and $\nu = c/\lambda$, we easily find equation (6.2) written like

$$\lambda(\theta) = \lambda + \frac{h}{m_e c}\,(1 - cos\theta), \tag{6.9}$$

so that $f = h/(m_e c)$. The actual numerical value coincides with the experimental one when one substitutes the values for h, m_e, c. Observe also that in the formal limit as $m_e \to +\infty$, equation (6.9) gives $\lambda(\theta) \to \lambda$. This explains the isotropic component of the scattered wave with identical wavelength (to the incoming one), as if this component were due to quanta of light interacting with particles of much bigger mass than the electron's (an atom of the substance, or the entire obstacle).

Remarks 6.4. The models of Einstein and Compton explain the photoelectric effect and formula (6.2) perfectly, both in quantitative and qualitative terms. Yet they must be considered *ad hoc* models, unrelated and actually conflicting with the physics knowlegde of their times: the idea that electromagnetic waves, hence also light, have a particle structure cannot account for wavelike phenomena – such as *interference* and *diffraction* – known since Newton and Huygens. In some way the wave and corpuscular nature of light (electromagnetic waves) must *co-exist* in the real world: this is forbidden in the paradigm of classical physics, but possible in the totally-relativistic formulation of QM by introducing the notion of *photon*, a massless particle which we shall not examine thoroughly. ■

6.3 An overview of Wave Mechanics

In this text we will not discuss quantum properties of light, which would require a deeper study of QM's formalism. In a reversal of viewpoint, the ideas mentioned previously about the early attempts to describe light from a quantum perspective proved very practical to formulate *wave mechanics*, which has rights to be considered the first step towards a QM formulation.

Wave mechanics is among the first rudimentary versions[3] of QM *for particles with mass*. We will spend only a little time on spelling out the foundational ideas of wave mechanics that shed light on the cornerstones of QM's proper formalism. In particular, we will forego result that are historically related to Schrödinger's stationary equation and the ensuing description of the energy spectrum of the hydrogn atom. We will return to these issues after the formalism has been set up.

6.3.1 De Broglie waves

A quantum of light, according to Einstein and Compton, is associated to a monochromatic plane electromagnetic wave with wavenumber $\mathbf{k} = \mathbf{p}/\hbar$ and *angular frequency* $\omega = E/\hbar$, which is just the 2π-multiple of the frequency ν. Each component of the plane wave (one along each of the three orthonormal vectors of the electric or magnetic field vibrating perpendicularly to $\mathbf{k}$) has the form of a scalar wave:

$$\psi(t,\mathbf{x}) = Ae^{i(\mathbf{k}\cdot\mathbf{x}-t\omega)} . \tag{6.10}$$

Actually only the *real part* of the above has any physical meaning, but it is much more convenient to use complex-valued waves for a number of reasons. For instance, they appear in Fourier's decomposition (see Section 3.6) of a general solution to the electromagnetic field equations (Maxwell's equations) or, more generally, d'Alembert's equation. In terms of *momentum* and *energy* of the quantum of light, the same wave may be written as

$$\psi(t,\mathbf{x}) = Ae^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-tE)} . \tag{6.11}$$

[3] Another version, developed in parallel by Heisenberg, consisted in the so-called *matrix mechanics*, which we will not treat.

Note how only the momentum and the energy of the quantum of light appear. In 1924 de Broglie put forward a truly revolutionary conjecture: *just like particles (photons) are associated to electromagnetic waves in certain experimental contexts, so one should be able to relate some sort of waves to matter particles.* According to de Broglie, these 'waves of matter' should be of the form (6.11), where now $\mathbf{p}$ and E are meant as momentum and (kinetic) energy of *particles*. The wavelength associated to a particle of momentum $\mathbf{p}$,

$$\lambda = h/|\mathbf{p}|\,, \tag{6.12}$$

is called *de Broglie wavelength of the particle*.

It was not at all clear what could be the nature of these alleged waves until 1927, when *experimental evidence* was gained of waves associated to electronic behaviour through two experiments carried out by Davisson and Germer, and independently G.P. Thompson. Without going into details, let us just say the following. It is a known fact that when a (sound, electromagnetic, ...) wave hits an obstacle with an inner structure of dimensions *comparable or larger* than the wavelength, the scattered wave undergoes so-called *diffraction*. The various internal parts of the obstacle interact with the wave creating constructive and destructive interference, so that the resulting wave projects, onto a screen placed behind the obstacle, a pattern made of areas of smaller and bigger intensity (*darker* and *brighter* in the case of light beams). These figures are called *diffraction patterns*. If the obstacle is a crystal, the diffraction pattern allows to recover the structure of the crystal itself. Davisson, Germer and G.P. Thompson produced patterns by beaming electrons through crytals. More precisely, they obtained diffraction patterns from the *traces, clustered together, left by electrons* emitted by a crystalline structure of mesh 1 Å. The incredible fact, endorsing de Broglie's thesis, was that in the aforementioned experiments diffraction patterns would appear only if de Broglie's wavelength was comparable or smaller than the mesh, exactly as in electromagnetic diffraction.

Remarks 6.5. It is important to underline that diffraction is a phenomenon strictly due to the wavelike nature of waves (i.e. due to something that oscillates and the superposition principle). Diffraction patterns *cannot* be generated by particles that obey the usual laws of classical mechanics, whatever the obstacle. ■

6.3.2 Schrödinger's wavefunction and Born's probabilistic interpretation

In 1926 Schrödinger took de Broglie's ideas seriously and in two famous and extraordinary papers made a more detailed hypothesis: he associated to a particle not a single plane wave like (6.11), but rather a *wave packet* made by the superposition of de Broglie plane waves (in the sense of the Fourier transform). For *free* particles, whose energy is purely kinetic, Schrödinger's wave reads:

$$\psi(t,\mathbf{x}) = \int_{\mathbb{R}^3} \frac{e^{\frac{i}{\hbar}(\mathbf{p}\cdot\mathbf{x}-tE(\mathbf{p}))}}{(2\pi\hbar^2)^{3/2}}\,\hat{\psi}(\mathbf{p})\,d^3\mathbf{p}\,, \tag{6.13}$$

where $E(\mathbf{p}) := \mathbf{p}^2/(2m)$, and m is the particle's mass. Schrödinger observed that ray optics relies on a relation, called *eikonal equation* [GPS01, CCP82], that bears a strong formal resemblance to the *Hamilton-Jacobi* equation [FaMa06, GPS01, CCP82] of classical mechanics. He was looking for a fundamental equation for matter in a *wave mechanics* of sorts, hoping it would stand to Hamilton-Jacobi's equation in a similar way *d'Alembert's equation* approximates the eikonal equation [GPS01, CCP82]. In a nushell, wave mechanics should stand to classical mechanics as wave optics stands to ray optics. The celebrated *Schrödinger equation* was born. We will recover the equation only after having contructed the formalism. For a particle subject to a force with potential U, say $\mathbf{f}(t,\mathbf{x}) = -\nabla U(t,\mathbf{x})$, the equation reads:

$$i\hbar\frac{\partial\psi(t,\mathbf{x})}{\partial t} = \left[-\frac{\hbar^2}{2m}\Delta + U(t,\mathbf{x})\right]\psi(t,\mathbf{x}) \tag{6.14}$$

where $\Delta = \sum_{j=1}^{3}\frac{\partial^2}{\partial x_j^2}$ is the Laplacian on $\mathbb{R}^3$.

The de Broglie-Schrödinger wave ψ is a *complex*-valued function and was called *wavefunction* of the particle to which it is attached. The physical interpretation of the wavefunction ψ – at least in the standard interpretation ("Copenhagen's interpretation") of the QM formalism – came from Born in 1926:

$$\rho(t,\mathbf{x}) := \frac{|\psi(t,\mathbf{x})|^2}{\int_{\mathbb{R}^3}|\psi(t,\mathbf{y})|^2 d^3\mathbf{y}}$$

is the probability density of detecting the particle at the point $\mathbf{x}$ *and at time* t *measured during an experiment for determining the particle's position.*

Born's interpretation turned out to agree with later experience, but essentially was already in line with the experimental evidence found by Davisson, Germer and G.P. Thompson, by which the traces left by particles on the screen accumulated in regions where $\rho(t,\mathbf{x}) > 0$ and were absent where $\rho(t,\mathbf{x}) = 0$, thus giving rise to the diffraction patterns.

Remark 6.6. (1) From the mathematical point of view Born's hypothesis only requires *square-integrable* wavefunctions that are *not almost everywhere zero*. Put equivalently, *non-zero* elements in $L^2(\mathbb{R}^3, d^3\mathbf{x})$ make physical sense: a *Hilbert space* makes its appearance for the very first time in the construction of QM. (It is physically irrelevant that de Broglie's plane waves have no straighforward meaning in the light of Born's interpretation, for they do not belong in $L^2(\mathbb{R}^3, d^3\mathbf{x})$. Plane monochromatic waves, used to understand experimental results à la de Broglie, can be approximated arbitrarily well by elements of $L^2(\mathbb{R}^3, d^3\mathbf{x})$ by using distributions $\hat{\psi}(\mathbf{p})$ close to a value $\mathbf{p}_0$, which in turn determines with the desired accuracy the wavelength $\lambda_0 = |\mathbf{p}_0|/h$ of de Broglie.)
(2) Assuming Born's interpretation, and in absence of experiments to determine its position, the particle with wavefunction ψ cannot evolve in time by the laws of classical mechanics: if it followed a regular trajectory, as classically prescribed, the function $|\psi|^2$ would have to vanish almost everywhere away from the trajectory. But any

regular curve in $\mathbb{R}^3$ has zero measure, so $|\psi|^2$ would be null almost everywhere in $\mathbb{R}^3$, a contradiction. In other terms when *no* experiment is made to detect a particle's position, the particle *cannot* be thought of as a classical object, for its time evolution is governed by the evolution of the wave ψ (solution to Schrödinger's equation).
(3) If we accept, as in the Copenhagen interpretation, that the wavefunction ψ describes *in full* the physical state of the particle, then the particle's position must be *physically* indefinite before an experiment is conducted to pin it down, and also attached indissolubly to the *experiment* in a probabilistic way. It is wrong to think that the probabilistic description is meant to cover for ignorance about the system's state, as in "the position is well determined, but we do not know it". In the Copenhagen interpretation the position does *not* exist until we make an experiment to determine it and until the particle's state (the maximum amount of information about its physical properties in time) is described by ψ. In wave mechanics a quantum has, thus, a dual *wave-particle* essence, but the two *never clash because they never manifest themselves simultaneously*. ■

6.4 Heisenberg's uncertainty principle

When one tries to evaluate experimentally an arbitrary quantity in a physical system, the state of the system may be altered by interacting with it. In principle, the classical description would allow to make this perturbation negligible. In 1927 Heisenberg realised that the combined hypotheses of Planck, Einstein, Compton, and de Broglie had a momentous (and epistemologically relevant) consequence. In quantitative terms Heisenberg's principle asserts that if we consider quantum systems and particular quantities to be measured, it is not always possible to disregard (as infinitesimal) the variation in the state of the system generated by a measurement: Planck's constant bounds from below the product of certain quantities. Precisely, after having considered thought experiments involving some of the hypotheses in Planck's, Einstein's, Compton's and de Broglie's models, Heisenberg concluded that:

In trying to determine the position or the momentum of a particle moving along a given axis x, we alter the momentum or the position, respectively, along the same axis, in such a way that the product of the two minimum variations Δx, Δp (of the final values of position and momentum) obeys

$$\Delta x \Delta p \gtrsim h\,. \tag{6.15}$$

If position and momentum are measured along orthogonal axes the above product can be made arbitrarily small, instead.

Equation (6.15) is *Heisenberg's uncertainty principle* for position and momentum. An analogous relationship holds for the uncertainty ΔE of a particle's energy E and the uncertainty Δt of the instant t of measurement of the energy[4]:

$$\Delta E \Delta t \gtrsim h\,. \tag{6.16}$$

To illustrate the matter let us consider the thought experiment whereby one seeks to determine the position X of an electron, with known initial momentum P, by hitting it with a monochromatic lightbeam of wavelength λ that propagates in the direction x. Let us imagine we can read the position off a screen parallel to the axis x using a lens placed between the axis and the screen. A quantum of light that has interacted with the electron will go through the lens and hit the screen, thus producing an image X'. Since the lens' aperture is finite, the outgoing direction of the quantum of light generating X' cannot be pinned down with absolute precision. Wave optics predicts in X' a diffraction pattern by which we may measure the coordinate X with a precision not smaller than

$$\Delta X \gtrsim \frac{\lambda}{\sin\alpha}\,,$$

where α is half the angle under which we see the lens from X. To the quantum of light corresponds a momentum h/λ, so the uncertainty in the component P_x of the outgoing quantum will approximatively be $h(\sin\alpha)/\lambda$. The total momentum of the system particle-quantum of light-microscope will remain constant, hence the uncertainty in the x-component of the particle's exit momentum must equal the corresponding uncertainty in the light quantum itself:

$$\Delta P_x \gtrsim \frac{h}{\lambda}\sin\alpha\,.$$

The product of the variations along the axis x is then at least

$$\Delta X \Delta P_x \gtrsim h\,.$$

Remarks 6.7. Heisenberg's principle, at this level, bears the same logical (in)consistency of the proto-quantum models of Planck, Einstein, Compton *et al.* It should be viewed more like a *working assumption* towards a novel notion of particle, for which the classical terms position and momentum make sense only within the boundaries fixed by the principle itself: *a quantum particle is allowed only physical states in which momentum and position are neither defined, nor definable, simultaneously.*
It is worth stressing, as we will see, that Heisenberg's *principle* is a *theorem* in the final formulation of QM. ■

6.5 Compatible and incompatible quantities

Quantum phenomenology, irrespective of the uncertainty principle, shows that there are pairs of quantities A and B that are *incompatible*. This means that, in principle, arbitrarily accurate and simultaneous measurements of A, B can be carried out. More explicitly: suppose we first measure A on the system, obtaining a as result, and immediately after we measure B obtaining b. Then a further reading of A, infinitesimally

[4] This second uncertainty relationship has a controversial status and its interpretion is a much thornier issue than the former's. We will not enter this territory, and refer to classical textbooks as [Mes99] in this respect.

close to B (so not to blame the time lapse), will be a value a_1 that is typically different from a, even by far. The same happens swapping the roles of A and B.

For instance, position and momentum along a fixed direction are incompatible pairs, and comply with Heisenberg's principle. Allegiance to Heisenberg and incompatibility have to do with each other, but the precise relationship can be explained properly only after completing the formalism. In general, incompatible quantities do not satisfy the uncertainty principle.

It turns out that incompatible quantum quantities never depend on one another, nor there exist devices capable of measuring them simultaneously.

There is a point to the call the attention to: quantum phenomenology shows that *compatible* quantities A' and B' do exist. This entails that if we measure first A' on the system and read a', and immediately afterwards measure B' reading b', the next measurement of A' – as close to B' as we want (so that time evolution does not interfere with measurements) – gives the same result a'. The same happens swapping A' and B'. In particular, any physical quantity A is compatible with itself and with any function depending on A (like the position of a particle along a line and the squared position).

An example of pairs of incompatible quantities on which Heisenberg's principle does not speak is provided by two distinct components of a particle's **spin**. The spin of the electron (and of all nuclear and subnuclear particles) was introduced by Goudsmit and Uhlembeck in 1925 [Mes99, CCP82] in order to make sense of some "bizarre" properties, the so-called *anomalous Zeeman effect* for atomic energy spectra (spectral lines) in alkali metals. In semi-classical sense the spin represents the intrinsic angular momentum of the electron, which may be considered, from a certain point of view, a consequence of the nonstop rotation of the electron around its centre of mass. This explanation, however, is misleading and cannot be taken verbatim. The deeper meaning of spin emerges in Wigner's framework, whereby an elementary quantum particle is defined as an elementary system invariant under the action of the Poincaré group.

Associated to the spin is an intrinsic magnetic momentum that is responsible for the observed anomalous Zeeman effect. At any rate, spin is a vector-valued quantity with characteristic quantum features that distinguish it from a classical angular momentum, thus making it a quantum angular momentum. The first difference is in the range of the modulus of spin. In the unit $\hbar$ these values are always of the type $\sqrt{s(s+1)}$, where s is an integer fixed by the kind of particle, e.g. $s = 1/2$ for the electron. Each of the three components of the spin, with respect to a positive orthonormal frame system, can take any of the $2s+1$ discrete values $-s, -s+1, \ldots, s-1, s$. The spin's three components are pairwise *incompatible* quantities: measuring in rapid sequence two of them gives distinct readings, as explained above. It is important to say that the components of a particle's orbital momentum and total angular momentum are incompatible exactly in the same way.

Compatible quantities for a quantum particle are, for instance, the x-component of the momentum and the y-component of the position vector.

This page intentionally left blank

7

The first 4 axioms of QM: propositions, quantum states and observables

Some historians claim that it is very difficult, nowadays, to find the line separating – and at the same time joining – the experimental level from the so-called theoretical one. But their view already includes several arbitrary elements, the so-called approximations.

Paul K. Feyerabend

In this chapter we will discuss the general mathematical structure of Quantum Mechanics. The procedure to achieve this is due to von Neumann, essentially, and will be presented here in its modern account via *Gleason's theorem*: that is, an extension of classical (Hamiltonian) mechanics that keeps in account the experimental evidence about the nature of quantum systems, seen in the previous chapter.

The first section recaps the results of Chapter 6, emphasising aspects that will be fundamental later.

In the following section we re-examine facets of Hamilton's formulation of mechanics from a set-theoretical and formal/logical perspective: we present the interpretation of the theory's foundations in which elementary propositions on the physical system are described by a σ-algebra, while states can be described by Borel probability measures (possibly, Dirac measures) on the σ-algebra.

Section three will show how the classical structure may be modified to comprise the description of the quantum phenomenology. Now the σ-algebra is replaced by the lattice of projectors onto a suitable Hilbert space, and a generalised σ-additive measure on the projectors' lattice takes the place of states. Similar approaches have been explored in depth by [Mac63], [Jau73], [Pir76], [Var07]. A general critical discussion appears in [BeCa81], while more recent results can be found in [EGL09].

We enter the heart of the matter in section four. With *Gleason's theorem* we explain that the aforementioned generalised measures are nothing but positive trace-class operators with trace one. In this way we introduce the convex space of quantum states, where *pure states* (or *rays*) are identified with extreme points. We also provide the formal description of the notion of compatible propositions, measurement process and the existence of *superselection rules*, together with the decomposition of the Hilbert space into *coherent sectors*.

The fifth, and last, section is devoted to the heuristic construction of the notion of *observable* as collection of elementary propositions giving a *projector-valued measure* (PVM) on the Hilbert space of the system. The construction will also yield a physical motivation for the *spectral theorem*, proved subsequently.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_7, © Springer-Verlag Italia 2013

7.1 The pillars of the standard interpretation of quantum phenomenology

We summarise below a few cardinal properties of the behaviour of quantum systems that were briefly described in the previous chapter.

QM1. (i) On a quantum system whose state has been fixed, measurements have a probabilistic outcome. It not possible to foresee the measurement's outcome, but only its probability.
(ii) However, if a quantity has been measured and gives a certain reading, repeating the measurement immediately after (so that the system does not evolve in the mean time) will give the same result.

QM2. (i) There exist **incompatible** physical quantities, in the following sense. Call A, B two such quantities. If we first measure A on the system (in a given state) and read a as outcome, and immediately after we measure B obtaining b, a subsequent measuring of A – as close as we want to the measurement of B to avoid ascribing the result to the evolution of the state – produces a reading $a_1 \neq a$, in general. The same happens swapping A and B.
It turns out that (a) incompatible quantum quantities are never dependent on one another, and (b) there are no instruments capable of simultaneous measurements.
(ii) There exist **compatible** physical quantities in the following sense. Call A', B' two such quantities. If we first measure A' on the system (in a given state) and obtain a as result, and immediately after we measure B obtaining b, a subsequent measuring of A – as close as we want to the measurement of B to avoid attributing the result to the evolution of the state – produces the same reading a. The same happens swapping A and B.
It turns out that (a) every physical quantity is compatible with itself, and (b) if two quantities are function one of the other (e.g. the energy and its square), then they are compatible.

Remark 7.1. (1) QM1 and **QM2** refer to physical quantities that do not characterise a physical system. By this we intend quantities whose range does *not* depend on the state and thus allow to distinguish a system from another. On the contrary, the remaining quantities mentioned by QM1 and QM2 take values that *depend on the state of the system.*

The physical quantities that **QM1** and **QM2** refer to, in relation to whether the outcome of successive experiments can or not be reproduced, are of course quantities that attain discrete values. As far as continuous quantities are concerned the matter is much more delicate, and we will not examine it [BGL95]. Irrespective of the type of quantities (continuous vs. discrete) what we can say, in general, is: *two quantities are compatible if and only if there exists a device capable of simultaneous measurements.*

Furthermore, **QM1** and **QM2** refer to extremely idealised measuring processes, in particular to those in which the microscopic physical system is not destroyed by the measurement. The measuring procedures employed in the experimental practice are rather diversified.

(2) It is clear we cannot be absolutely certain that quantum systems satisfy (i) in QM1. We could be tempted to think that the stochastic outcome of measurements is really due to the lack of full knowledge scientists have of the system's state, and that by knowing it *in toto* they would be able to predict measuring outcomes. In this sense quantum probability would merely have an *epistemic* nature. In the standard interpretation of QM, the so-called **Copenhagen interpretation**, the stochastic outcome of a measurement is assumed as a *primary* feature of quantum systems. There are nonetheless interesting attempts to interpret quantum phenomenology based on alternative formalisms (the so-called formulations by *hidden variables*) [Bon97]. There, the stochastic feature is explained as it were due to partial human knowledge about the system's true state, which is described by more variables (and in different fashion) than those needed in the standard formulation. None of these attempts (despite some are indeed deep, like *Bohm's theory*) is considered nowadays completely satisfactory, and does not threaten the standard interpretation and formulation of QM when one considers also relativistic quantum theories, and relativistic QFT in particular.

But we must stress that one *cannot* build a *completely classical* physical theory (that counts non-quantum relativistic theories among classical ones) that is capable of explaining the experimental phenomenology of a quantum system in its entirety. Hidden variables, in order to agree with known evidence, must at any rate satisfy a rather unusual *contextuality* property for classical theories. Furthermore, any theory that wishes to explain the quantum phenomenology, QM included, must be *nonlocal* [Bon97]. Actually, as we shall see in Chapter 13.4.2, after the theoretical investigations of Einstein, Podolsky and Rosen first, and then Bell, experiments have proved the existence of correlations among measurements made in different regions of space and at lapses so short that transmission of information between events is out of the question, whichever physical mean moves at a speed below the speed of light in vacuum.

(3) Implicit in **QM1** and **QM2** is that physical systems of interest, both in quantum theory and quantum phenomenology, divide in two large categories: *measuring instruments* and *quantum systems*. The Copenhagen formulation assumes that measuring devices are systems obeying the laws of classical physics. These hypotheses match the data coming from experiments, and although quite crude, theoretically-speaking, they lie at the heart of the interpretation's formalism. Therefore not much can be said about them within the standard formulation. At the moment, for instance, it is not clear where to draw the line between classical and quantum systems, nor how this boundary may be described inside the formalism, and neither whether the compound system 'instrument-quantum system' can be itself considered a larger quantum system, and as such treated by the formalism. In closing, the physical interaction process between instrument and quantum system, that produces the actual measurement, is not described from within the standard quantum formalism as a dynamical process. For a deeper discussion on these stimulating and involved issues we refer to [Bon97, Des99], and also to the superb section dedicated to foundational aspects of quantum theories in the *Stanford Encyclopedia of Philosophy*[1]. ■

[1] http://plato.stanford.edu/.

7.2 Classical systems: elementary propositions and states

Let us see now how (Borel) probability measures can be employed to represent the physical states of classical systems. A generalisation will be used later to describe the states of a quantum system mathematically.

7.2.1 States as probability measures

The modern formal treatment of probability theory, due to Kolmogorov, translates into the study of *probability measures*. We recap below a few definitions taken from Chapter 1.4.

Definition. *A positive σ-additive measure μ on the measure space (X, Σ) is called* **probability measure** *if $\mu(\mathsf{X}) = 1$.*

The simplest case of a probability measure on (X, Σ) is certainly the **Dirac measure δ_x concentrated at** $x \in \mathsf{X}$, defined as:

$$\delta_x(E) = 0 \text{ if } x \notin E \quad , \quad \delta_x(E) = 1 \text{ if } x \in E, \quad \text{for any } E \in \Sigma.$$

In the sequel we shall work with *Borel measures*, so we recall the following notions from Chapter 1.4, which we have already used. They will be useful in the rest of the text.

Definition. *Let X be a topological space.*

(a) *The* **Borel σ-algebra** *of X, $\mathscr{B}(\mathsf{X})$, is the smallest (under intersections) σ-algebra containing the open sets of X.*
(b) *The elements of $\mathscr{B}(\mathsf{X})$ are the* **Borel sets** *of X.*
(c) *$f : \mathsf{X} \to \mathbb{C}$ is (***Borel***)* **measurable** *if it is measurable with respect to $\mathscr{B}(\mathsf{X})$ and $\mathscr{B}(\mathbb{C})$, i.e. $f^{-1}(E) \in \mathscr{B}(\mathsf{X})$ for any $E \in \mathscr{B}(\mathbb{C})$.*

Obviously, in (c), $\mathscr{B}(\mathbb{C})$ refers to the standard topology of $\mathbb{C}$, and the definition can be stated similarly to comprise $\mathbb{R}$-valued maps and $\mathscr{B}(\mathbb{R})$.

Definition. *If X is a locally compact Hausdorff space, a* **Borel measure** *on X is a positive, σ-additive measure on $\mathscr{B}(\mathsf{X})$.*

Consider a classical physical system with n spatial freedom degrees, so $2n$ degrees overall, including kinetical degrees ("velocities"). The *Hamiltonian formulation* [GPS01, FaMa06] of the system's dynamics, very briefly, goes as follows.

(i) The ambient space is the **phase spacetime** $\mathscr{H}_{n+1}$. This is a smooth manifold of real dimension $2n+1$ formed by the disjoint union[2] of $2n$-dimensional submanifolds

[2] $\mathscr{H}_{n+1}$ is the total space of a *fibre bundle* with base $\mathbb{R}$ (the *time axis*) and fibres $\mathscr{F}_t$ given by $2n$-dimensional *symplectic manifolds*. There is an atlas on $\mathscr{H}_{n+1}$ whose local charts have

$\mathscr{F}_t$, all diffeomorphic and smoothly depending on $t \in \mathbb{R}$:

$$\mathscr{H}_{n+1} = \bigsqcup_{t\in\mathbb{R}} \mathscr{F}_t \,.$$

(ii) The coordinate $t \in \mathbb{R}$ is **time**, every $\mathscr{F}_t$ is the **phase space** at time t and any point in $\mathscr{F}_t$ represents a **state** of the system **at time** t.
(iii) $\mathscr{H}_{n+1}$ admits an atlas with local coordinates: $t, q^1, \ldots, q^n, p_1, \ldots, p_n$ (where $q^1, \ldots, q^n, p_1, \ldots, p_n$ are **symplectic coordinates** on $\mathscr{F}_t$) in which the system's evolution equations take **Hamilton's** form:

$$\frac{dq^k}{dt} = \frac{\partial H(t,q(t),p(t))}{\partial p_k} \quad k = 1,2,\ldots,n\,, \tag{7.1}$$

$$\frac{dp_k}{dt} = -\frac{\partial H(t,q(t),p(t))}{\partial q^k} \quad k = 1,2,\ldots,n\,, \tag{7.2}$$

where H, the **Hamiltonian (function)** of the system in local coordinates, is known once the system is known.

With this representation the system's evolution in time is described by the integral curves of Hamilton's differential equations. Each integral curve determines, at any given time $t \in \mathbb{R}$, a point $(t, s(t)) \in \mathscr{H}_{n+1}$, $s(t) \in \mathscr{F}_t$, where the curve meets $\mathscr{F}_t$. $s(t)$ is the state of the system at time t.

We remark that (in absence of constraints) the choice of a frame system $\mathscr{I}$ allows to decompose locally $\mathscr{H}_{n+1}$ as a the Cartesian product $\mathbb{R} \times \mathscr{F}$, where $\mathbb{R}$ is the time axis (once the origin has been fixed) and $\mathscr{F}$ is identified with phase space at time $t = 0$. Other choices of the framing give different identifications. Similarly, the Hamiltonian H, identified with the total *mechanical energy* of the physical system in certain circumstances, depends on the reference systems; however, Hamilton's equations of motion are independent of any frame: solutions do not depend on choices, but are the same on $\mathscr{H}_{n+1}$ irrespective of the framing.

In certain, fundamental, situations, like *statistical mechanics* or *thermodynamical statistics*, the system's state is not known with abolute precision, so neither is the evolution of the system. In these cases one uses statistical ensembles [Hua87, FaMa06]: rather than considering one copy of the system, one takes a statistical ensemble of identical and independent copies of the system, with states distributed on every $\mathscr{F}_t$ with a certain probability density that is locally representable by a C^1 map $\rho = \rho(t,q,p)$. The density evolves in time in accordance to **Liouville's equation**:

$$\frac{\partial \rho}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial \rho}{\partial q^i}\frac{\partial H}{\partial p_i} - \frac{\partial H}{\partial q^i}\frac{\partial \rho}{\partial p_i} \right) = 0\,. \tag{7.3}$$

---

coordinates $t, q^1, \ldots, q^n, p_1, \ldots, p_n$, where t is the natural parameter on the base $\mathbb{R}$ while the remaining $2n$ coordinates define a local symplectic frame on each $\mathscr{F}_t$.

The function $\rho(t,s)$, with $s \in \mathscr{F}_t$, represents the probability density that the system is in the state s at time t. The interpretation of ρ requires, for any t:

$$\rho(t,s) \geq 0 \quad \text{and} \quad \int_{\mathscr{F}_t} \rho \, d\mu_t = 1 \, . \tag{7.4}$$

The measure μ_t on $\mathscr{B}(\mathscr{F}_t)$ is the Lebesgue measure $dq^1 \cdots dq^n dp_1 \cdots dp_n$ (extended to $\mathscr{F}_t$ using a partition of unity) on every local symplectic chart of $\mathscr{F}_t$. The known **Liouville theorem** states that with this choice μ_t on every phase space, the integral in (7.4) does not depend on $t \in \mathbb{R}$ provided ρ solves (7.3) [Hua87, FaMa06, CCP82].

In case one works with statistical ensembles, the density ρ_t is still thought of as the *system's state at time t*, even if this notion of state generalises the previous one. We shall abide by this convention, and distinguish **sharp states** given by points $r(t) \in \mathscr{F}_t$, from **probabilistic states** determined by a Liouville density ρ_t on $\mathscr{F}_t$. In either case the state at time t can be viewed as a *Borel probability measure* $\{\nu_t\}_{t\in\mathbb{R}}$ defined on the phase space $\mathscr{F}_t$. More precisely:

(i) for a probabilistic state[3] $\nu_t(E) := \int_E \rho(t,s) d\mu_t$ if $E \in \mathscr{B}(\mathscr{F}_t)$;
(ii) for a sharp state $\nu_t := \delta_{r(t)}$.

Remarks 7.2. In order to represent the system's states at time t in a completely general way, thereby foregoing the evolution problem and forgetting a standard Hamiltonian formulation, one could use topological manifolds $\mathscr{F}_t$ rather than smooth ones. States (at time t) could be represented in terms of probability measures for the Borel σ-algebra. The existence of a topology on $\mathscr{F}_t$ is intrinsically related to the existence of "neighbourhoods" of its points coming from experimental errors, as infinitesimal as we want but not negligible. Better said, the possibility of distinguishing points in $\mathscr{F}_t$, despite measuring errors, is expressed mathematically by requesting a Hausdorff topology on $\mathscr{F}_t$ (as happens when defining a smooth manifold). ■

7.2.2 *Propositions as sets, states as measures on them*

If we assume that the Hamiltonian description of our system retains all physical properties, then it must be possible to describe, in phase space $\mathscr{F}_t$ at time t, all statements about the system that at time t are true, false, or true with a certain probability, in some way or another. Moreover it should be possible to recover the truth value of those propositions, i.e. the probabilty they are true, from the state ν_t of the system. Here is a natural way to do this.

Observe first that every proposition P determines a subset in $\mathscr{F}_t$ that contains the points (thought of as sharp states) that render P true (at time t). We indicate this set by the same symbol $P \subset \mathscr{F}_t$. Next, suppose we work with a sharp state, so that ν_t

[3] $\mathscr{F}_t$ is a smooth manifold hence a locally compact Hausdorff space (since locally homeomorphic to $\mathbb{R}^n$). As μ_t is defined on $\mathscr{B}(\mathscr{F}_t)$ and ρ_t is continuous, ν_t is well defined on $\mathscr{B}(\mathscr{F}_t)$.

is a Dirac measure. Then proposition P is true at time t if and only if the point $r(t)$ describing the system at time t belongs to the set P. Now assign the conventional value 0 to a false proposition at time t, and 1 to a true one at t, in relation to state $\nu_t = \delta_{r(t)}$. The crucial observation is that the truth value of P is $\nu_t(P)$, when the state is ν_t, though of as *measure of* $P \subset \mathscr{F}_t$ *with respect to the (Dirac) measure* ν_t.

This fact clarifies the concrete meaning of the Dirac measure ν_t viewed as system's state at time t. Furthermore, the same interpretation can be employed when the state is probabilistic: $\nu_t(P)$ *represents the probability that proposition* $P \subset \mathscr{F}_t$ *is true at time* t *when the state* ν_t *is probabilistic.*

Remark 7.3. (1) Everything we said makes sense if the set P belongs to the σ-algebra on which the measures ν_t are defined. This is the Borel σ-algebra, and hence it is reasonably large.
(2) One proposition may be formulated in different yet equivalent ways. When we identify propositions with sets in $\mathscr{F}_t$ we are explicitly assuming that *if two propositions determine the same subset in* $\mathscr{F}_t$*, they must be considered identical.* ■

7.2.3 Set-theoretical interpretation of the logical connectives

Given two propositions P, Q, we can compose them using *logical connectives* to obtain other propositions. In particular, we can form the propositions $P \mathscr{O} Q$ and $P \mathscr{E} Q$ using the binary connectives called *disjountion* (*inclusive or*), and *conjunction* (*and*). Negifying one proposition produces its *negation* $\dashv P$.

We can interpret these propositions in terms of sets in the Borel σ-algebra on $\mathscr{F}_t$:

(i) $P \mathscr{O} Q$ corresponds to $P \cup Q$;
(ii) $P \mathscr{E} Q$ corresponds to $P \cap Q$;
(iii) $\dashv P$ corresponds to $\mathscr{F}_t \setminus P$.

There is a *partial order relation* on subsets of $\mathscr{F}_t$ given by the inclusion: $P \leq Q$ if and only if $P \subset Q$.

At the level of propositions, the most natural interpretation of $P \subset Q$ *is to say that* P *implies* Q*, i.e.* $P \Rightarrow Q$. Equivalently: each time the system is in a sharp state satisfying P, the state satisfies Q as well. For non-sharp states: for any state, the probability that Q is true is not smaller than the probability that P is true.

Remark 7.4. (1) The truth probability of composite propositions can be computed from the starting propositions using the measure ν_t, because a σ-algebra is closed under the set-theoretical operations corresponding to $\mathscr{O}$, $\mathscr{E}$, $\dashv$.
(2) It is easy to see that if ν_t is a Dirac measure, the truth probability (in this case either 0 or 1) assigned to each expression (i), (ii), (iii), coincides with the value found on the truth tables of the connective used. For instance, $P \mathscr{O} Q$ is true ($\nu_t(P \cup Q) = 1$) if and only if at least one of its constituent propositions is true ($\nu_t(P) = 1$ or $\nu_t(Q) = 1$); in fact the point x at which the Dirac measure $\delta_x = \nu_t$ concentrates lies in $P \cup Q$ iff x lies in either set P, Q. ■

7.2.4 "Infinite" propositions and physical quantities

Propositional calculus normally disregards propositions made by infinitely many propositions and connectives like $P_1 \mathscr{O} P_2 \mathscr{O} \ldots$. Interpreting propositions and connectives in terms of points and operations on a σ-algebra, though, allows to "handle" infinitely-long propositions.

We can relate some (at least) of these propositions to measurable *physical quantities* on the system. Generally speaking, we may consider the *physical quantities* defined on our Hamiltonian system as a collection of functions, somehow regular, defined on phase spacetime and real-valued: $f : \mathscr{H}_{n+1} \to \mathbb{R}$. A fairly broad choice of *regularity* is to take the class of maps that restrict to *Borel measurable* maps on each fibre $\mathscr{F}_t$. Less radical options are continuous maps, C^1 maps, or even C^∞ maps. From the point of view of physics it may seem natural to require physical quantities be described by functions that are at least continuous, because measurements are always affected by experimental errors when finding the point in $\mathscr{F}_t$ representing the state at time t: if maps were not continuous, small errors would cause enormous variations in a quantity's values. Nevertheless we must also remember there might be quantities with discrete range, for which the above issue is meaningless (discrete values can be distinguished using instruments with sufficient, finite, precision). As we are interested in the passage to the quantum case rather than in analysing the classical case, we shall not go deep into this kind of problem. We limit ourselves to working at a given instant t for which the physical quantities of concern will be measurable functions $f : \mathscr{F}_t \to \mathbb{R}$. If $f : \mathscr{F}_t \to \mathbb{R}$ is a physical quantity that can be measured on the system (at time t), using it we can construct statements of this kind:

$$P_E^{(f)} =$$

The value that f assumes on the system's state belongs to the Borel subset $E \subset \mathbb{R}$.

By considering Borel sets E, and not just open intervals or singlets for example, allows to treat quantities with both continuous and discrete ranges in the same way, and also keep track of the fact that the measurement made by an instrument is a set, not just a point, owing to the finite precision of the instrument itself. As a matter of fact $\mathscr{B}(\mathbb{R})$ contains closed sets, finite sets, countable sets and so on. In set-theoretical terms the proposition will be associated to a Borel set in $\mathscr{F}_t$ that we continue denoting by the same symbol

$$P_E^{(f)} = \text{“}f^{-1}(E) \subset \mathscr{F}_t\text{”}.$$

(As explained above, by this convention the probability that $P_E^{(f)}$ is true for the system at time t is $\nu_t(P_E^{(f)})$, once the state ν_t is known.)

Consider an interval $[a,b)$, $b \leq +\infty$. Decompose it in the disjoint union of infinitely many subintervals: $[a,b) = \cup_{i=1}^{\infty}[a_i, a_{i+1})$, where $a_1 := a$, $a_i < a_{i+1}$ and $a_i \to b$ as $i \to \infty$. Then the proposition

$$P_{[a,b)}^{(f)} =$$

The value of f on the state of the system falls in the Borel set $[a,b)$

can be clearly written as an *infinite disjunction*

$$P^{(f)}_{[a,b)} = \mathscr{O}_{i=1}^{+\infty} P^{(f)}_{[a_i,a_{i+1})}$$

of statements of the form:

$$P^{(f)}_{[a_i,a_{i+1})} =$$

The value of f on the state of the system falls in the Borel set $[a_i, a_{i+1})$.

This corresponds to writing the set $P^{(f)}_{[a,b)}$ as the disjoint union:

$$P^{(f)}_{[a,b)} = \cup_{i=1}^{+\infty} P^{(f)}_{[a_i,a_{i+1})} \, .$$

Therefore it makes physical sense to assume the existence of (certain) statements built by infinitely many connectives and propositions.

Since negifying $\mathscr{O}$ dually gives $\mathscr{E}$, if we assume the set of admissible propositions is closed under $\dashv$, then we must also accept propositions involving infinitely many $\mathscr{E}$s as physically meaningful.

The possibility of representing propositions as sets in a σ-algebra, and thus compute the probability they are true on a state using the corresponding measure, suggests to allow for propositions with countably many connectives $\mathscr{O}$ or $\mathscr{E}$, because the corresponding sets still belong in the σ-algebra, which is closed under countable unions and intersections.

To obtain a structure "isomorphic" to the σ-algebra built from atomic formulas and $\mathscr{O}$, $\mathscr{E}$, $\dashv$, we need to add two more propositions, playing the role of the sets $\varnothing$ and $\mathscr{F}_t$. These are the contradiction (whose truth probability is 0, whichever the state), denoted **0**, and the tautology (truth probability equal 1 whichever the state), written **1**.

Once propositions and sets are identified, the σ-algebra structure enables us to declare that *the set of elementary propositions P relative to the physical system of concern, equipped with the logical connectives* $\mathscr{O}$, $\mathscr{E}$, $\dashv$, *is "isomorphic" to a* σ*-algebra.*

The truth value of a proposition P, for sharp states, or its truth probability, for probabilistic states, on a given state at time t equals $\nu_t(P)$, *where now* $P \subset \mathscr{F}_t$ *is the set corresponding to the proposition.*

Remark 7.5. (1) We may ask whether the σ-algebra of all propositions on the system corresponds to the full Borel σ-algebra of $\mathscr{F}_t$, or if it is smaller. If we assume every bounded measurable real map on $\mathscr{F}_t$ is a physical quantity, then the answer is clearly yes, because among those maps are the characteristic functions of measurable subsets of $\mathscr{F}_t$.
(2) As earlier remarked, once we fix a frame system $\mathscr{I}$ (in absence of contraints, as in the cases at hand) the phase spacetime $\mathscr{H}_{n+1}$ of the physical system is mapped diffeomorphically to the Cartesian product $\mathbb{R} \times \mathscr{F}$, where $\mathscr{F}$ is the phase space at time 0 and $\mathbb{R}$ the time line (with given origin). Thus we may regard propositions at any given instant t as Borel subsets of $\mathscr{F}$, and any state at time t as a probability measure on $\mathscr{F}$. Henceforth, especially when generalising to the quantum case, we will harness this simplification of the formalism that results from a choice of frame. ■

7.2.5 Intermezzo: basics on the theory of lattices

In physical systems we can identify propositions and sets, and think states as measures on sets. In order to pass to quantum systems, where there is no notion of phase space, it is important to raise the following question. Do there exist mathematical structures that are not σ-algebras of sets but sort of isomorphic to one? The answer is yes and comes from the theory of lattices.

In the sequel we will use some basic notions from the theory of posets. We assume they are known to the reader; if not they can be found in the appendix.

Definition 7.6. *A partially ordered set* $(\mathsf{X}, \geq)$ *is a* **lattice** *when, for any* $a, b \in \mathsf{X}$,

(a) $\sup\{a,b\}$ *exists, denoted* $a \vee b$ *(sometimes called 'join').*
(b) $\inf\{a,b\}$ *exists, written* $a \wedge b$ *(sometimes 'meet').*

(The poset is not required to be totally ordered.)

Remark 7.7. (1) If $(\mathsf{X}, \geq)$ is partially ordered, as usual $a \leq b$ means $b \geq a$, for any $a, b \in \mathsf{X}$. The following three facts are easily equivalent: $a \wedge b = a$, $a \vee b = b$, $a \leq b$.
(2) On any lattice X, by definition of inf, sup we have the following properties, for any $a, b, c \in \mathsf{X}$:

Associativity: $(a \wedge b) \wedge c = a \wedge (b \wedge c)$ and $(a \vee b) \vee c = a \vee (b \vee c)$.
Commutativity: $a \vee b = b \vee a$ and $a \wedge b = b \wedge a$.
Absorption: $a \vee (a \wedge b) = a$ and $a \wedge (a \vee b) = a$.
Idempotency: $a \vee a = a$ and $a \wedge a = a$.

By the associative property we can write $a \vee b \vee c \vee d$ and $a \wedge b \wedge c \wedge d$ without ambiguity.
(3) The above properties characterise lattices: a set X equipped with binary operations $\wedge : \mathsf{X} \times \mathsf{X} \to \mathsf{X}$, $\vee : \mathsf{X} \times \mathsf{X} \to \mathsf{X}$ that satisfy the properties of (2) is partially ordered by the relation $a \geq b \Leftrightarrow b = b \wedge a$. In that case $\sup\{a,b\} = a \vee b$ and $\inf\{a,b\} = a \wedge b$. ■

Various types of lattices exist, and the next definition describes some of them.

Definition 7.8. *A lattice* $(\mathsf{X}, \geq)$ *is said:*

(a) Distributive *if* $\vee$ *and* $\wedge$ *distribute over one another: for any* $a, b, c \in \mathsf{X}$,

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee b) , \quad a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge b) .$$

(b) Bounded *if it admits a minimum* **0** *and a maximum* **1** *(sometimes called 'bottom' and 'top').*
(c) Orthocomplemented *if bounded and equipped with a mapping* $\mathsf{X} \ni a \mapsto \neg a$, *where* $\neg a$ *is the* **orthogonal complement** *of* a, *such that:*

(i) $a \vee \neg a = \mathbf{1}$ *for any* $a \in \mathsf{X}$;
(ii) $a \wedge \neg a = \mathbf{0}$ *for any* $a \in \mathsf{X}$;
(iii) $\neg(\neg a) = a$ *for any* $a \in \mathsf{X}$;
(iv) $a \geq b$ *implies* $\neg b \geq \neg a$ *for any* $a, b \in \mathsf{X}$.

(d) Modular, *if* $q \geq p$ *implies* $(q \vee p) \wedge r = q \vee (p \wedge r)$, $\forall p,q,r \in \mathsf{X}$.
(e) Orthomodular, *if orthocomplemented and* $q \geq p$ *implies* $q = p \vee ((\neg p) \wedge q)$, $\forall p,q \in \mathsf{X}$.
(f) σ-complete, *if every countable set* $\{a_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ *admits least upper bound* $\vee_{n\in\mathbb{N}} a_n$.
In an orthocomplemented lattice two elements a, b are:

orthogonal, *written* $a \perp b$, *if* $\neg a \geq b$ *(or equivalently* $\neg b \geq a$*).*
commuting, *if* $a = c_1 \vee c_3$ *and* $b = c_2 \vee c_3$ *with* $c_i \perp c_j$ *if* $i \neq j$.

A lattice with properties (a), (b) and (c) (hence (d) and (e)) is called a **Boolean algebra**. *A Boolean algebra satisfying (f) is a* **Boolean σ-algebra**.

*A (***distributive, bounded, orthocomplemented, σ-complete***)* **sublattice** *is a subset in* X *admitting a lattice structure (distributive, bounded, orthocomplemented, σ-complete) for the restrictions of* $\geq$ *and* $\neg$.

Remark 7.9. (1) It is immediate to prove that arbitrary intersections of orthocomplemented sublattices are orthocomplemented sublattices (with the same minimum, maximum and orthogonal complement of X).
(2) If X is an orthocomplemented lattice and $p,q \in \mathsf{X}$ belong to a Boolean subalgebra of X, then p and q commute, as noticed before. Strengthening the assumptions on X the converse holds too, [BeCa81].

Proposition 7.10. *Let* X *be an orthocomplemented lattice. Then* $p,q \in \mathsf{X}$ *commute if and only if the orthocomplemented sublattice spanned by* $\{p,q\}$ *(the intersection of all bounded orthocomplemented sublattices containing* $\{p,q\}$*) is a Boolean subalgebra of* X.

(3) It is easy to see that on any orthocomplemented lattice **De Morgan's laws** hold: for any $a,b \in \mathsf{X}$,

$$\neg(a \vee b) = \neg a \wedge \neg b\,, \quad \neg(a \wedge b) = \neg a \vee \neg b\,. \tag{7.5}$$

(4) In a general orthocomplemented lattice:

$$\text{if } a \perp b \text{ then } a \wedge b = \mathbf{0}.$$

A Boolean algebra X is modular, orthomodular and every pair $a,b \in \mathsf{X}$ commutes: using the distributive law, in particular, $a = (a \wedge \neg b) \vee (a \wedge b)$ and $b = (b \wedge \neg a) \vee (a \wedge b)$.
(5) By the definition of inf, sup and De Morgan's laws, a bounded orthocomplemented lattice is σ-complete iff every countable subset $\{a_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ admits a greatest lower bound $\wedge_{n\in\mathbb{N}} a_n$. In this case $\wedge_{n\in\mathbb{N}} a_n = \neg(\vee_{n\in\mathbb{N}} \neg a_n)$, so also $\vee_{n\in\mathbb{N}} a_n = \neg(\wedge_{n\in\mathbb{N}} \neg a_n)$. ∎

Definition 7.11. *If* X, Y *are lattices, a map* $h : \mathsf{X} \to \mathsf{Y}$ *is a* **(lattice) homomorphism** *when*

$$h(a \vee_{\mathsf{X}} b) = h(a) \vee_{\mathsf{Y}} h(b)\,,\ \ h(a \wedge_{\mathsf{X}} b) = h(a) \wedge_{\mathsf{Y}} h(b)\,,\ \ a,b \in \mathsf{X}$$

(with the obvious notations.) If X *and* Y *are bounded, a homomorphism h is further required to satisfy*

$$h(\mathbf{0}_{\mathsf{X}}) = \mathbf{0}_{\mathsf{Y}}\,, \quad h(\mathbf{1}_{\mathsf{X}}) = \mathbf{1}_{\mathsf{Y}}\,.$$

If X *and* Y *are orthocomplemented, a homomorphism h also satisfies*

$$h(\neg_{\mathsf{X}} a) = \neg_{\mathsf{Y}} h(x)\,.$$

If X, Y *are* σ*-complete, h further fulfills*

$$h(\vee_{n\in\mathbb{N}} a_n) = \vee_{n\in\mathbb{N}} h(a_n)\,, \textit{if } \{a_n\}_{n\in\mathbb{N}} \subset \mathsf{X}\,.$$

In all cases (bounded, orthocomplemented, σ*-complete lattices, Boolean (*σ*-) algebras) if h si bijective it is called* **isomorphism** *of the relative structures.*

Remark 7.12. **(1)** Since $b \geq a$ iff $b \wedge a = a$, the following facts hold. If $h : \mathsf{X} \to \mathsf{Y}$ is a homomorphism then for any $a, b \in \mathsf{X}$, $a \geq_{\mathsf{X}} b$ implies $h(a) \geq_{\mathsf{Y}} h(b)$ with the obvious notation.
(2) It is immediate to see that the inverse $h^{-1} : \mathsf{Y} \to \mathsf{X}$ of an isomorphism $h : \mathsf{X} \to \mathsf{Y}$ (of lattices or Boolean (σ-)algebras) is an isomorphism.
(3) Given an abstract Boolean σ-algebra X, does there exist a concrete σ-algebra of sets that is isomorphic to the previous one? In this respect the following general result holds, known as **Loomis–Sikorski theorem**. [4] This guarantees that every Boolean σ-algebra is isomorphic to a quotient Boolean σ-algebra $\Sigma/\mathscr{N}$, where Σ is a concrete σ-algebra of sets over a measureable space and $\mathscr{N} \subset \Sigma$ is closed under countable unions; moreover, $\varnothing \in \mathscr{N}$ and for any $A \in \Sigma$ with $A \subset N \in \mathscr{N}$, then $A \in \mathscr{N}$. The equivalence relation is $A \sim B$ iff $A \cup B \setminus (A \cap B) \in \mathscr{N}$, for any $A, B \in \Sigma$. It is easy to see the coset space $\Sigma/\mathscr{N}$ inherits the structure of Boolean σ-algebra from Σ with respect to the (well-defined) partial order relation $[A] \geq [B]$ if $A \supset B$, $A, B \in \Sigma$.

This is the sharpest result in the general case. Consider, for instance, the σ-algebra $\mathscr{B}([0,1])$ of Borel sets in $[0,1]$. Take the quotient $\mathscr{B}^*([0,1]) := \mathscr{B}([0,1])/\mathscr{N}$, where $\mathscr{N}$ consists of subsets in $[0,1]$ of zero (Lebesgue) measure. It can be proved that $\mathscr{B}^*([0,1])$ is isomorphic to no σ-algebra of subsets on any measurable space.

But if one restricts to Boolean algebras only, the known **Stone representation theorem**[5] asserts that an abstract Boolean algebra is always isomorphic to some concrete algebra of sets, without the need of quotienting. ■

7.2.6 The distributive lattice of elementary propositions for classical systems

We can revert to σ-algebras of sets, and with the definitions given above the following assertions are trivial, so their proof is left as exercise.

Proposition 7.13. *Every* σ*-algebra on* X *is a Boolean* σ*-algebra where:*

(i) *the partial order is the inclusion (hence* $\vee$ *corresponds to* $\cup$ *and* $\wedge$ *to* $\cap$*);*
(ii) *the maximum and minimum in the Boolean algebra are* X *and* $\varnothing$*;*
(iii) *orthocomplements correspond to set-complements with respect to* X.

[4] Sikorski S.: On the representation of Boolean algebras as field of sets. Fund. Math. **35**, 247–256 (1948).

[5] Stone M.H.: The Theory of Representations of Boolean Algebras. Trans. AMS **40**, 37–111 (1936).

Proposition 7.14. *Let Σ, Σ' be σ-algebras on X and X' respectively, and $f : \mathsf{X} \to \mathsf{X}'$ a measurable function.*

(a) *The sets $P_E^{(f)} := f^{-1}(E)$, $E \in \Sigma'$, define a Boolean σ-subalgebra of the Boolean σ-algebra of Proposition 7.13.*
(b) *The mapping $\Sigma' \ni E \mapsto P_E^{(f)}$ is a homomorphism of Boolean σ-algebras.*

The same assertions hold for the set of propositions relative to a physical system.

Proposition 7.15. *Propositions relative to a classical physical system form a distributive, bounded, orthocomplemented and σ-complete lattice, i.e. a Boolean σ-algebra, where:*

(i) *the order relation is the logical implication, the conjunction is the intersection and the disjunction is the union;*
(ii) *the maximum and minimum are the tautology* **1** *and the contradiction* **0***;*
(iii) *orthocomplementation corresponds to negation.*

If a measurable function $f : \mathscr{F} \to \mathbb{R}$ represents a physical quantity, then:

(a) *As E varies in the Borel σ-algebra of $\mathbb{R}$, the propositions*

$$P_E^{(f)} =$$

The value f takes on the state of the system belongs in the Borel set $E \subset \mathbb{R}$,

define a Boolean σ-algebra.
(b) *The map that sends a Borel set $E \subset \mathbb{R}$ to the proposition $P_E^{(f)}$ is a homomorphism of Boolean σ-algebras.*

7.3 Propositions on quantum systems as orthogonal projectors

Eventually we can move on to quantum system. In trying to follow an appoach that is as close as possible to the classical case, we first aim at finding a mathematical model for the class of elementary propositions relative to a quantum system. Then we will evaluate at time t by conducting experiments with the aid of suitable instruments, whose results is merely 0 (= *the proposition is false*) or 1 (= *the proposition is true*). We still do not know how to describe the system, but we know **QM1** and **QM2** have to hold in relationship to the quantum quantities that are measurable on the system. For the moment we concentrate on **QM2**. We know there exist incompatible *quantities*. Then it is immediate to conclude that there must be incompatible *propositions*: if A and B are incompatible, then

$$P_J^{(A)} =$$

The value of A on the state of the system belongs to the Borel set $J \subset \mathbb{R}$,

$$P_K^{(B)} =$$

The value of B on the state of the system belongs to the Borel set $K \subset \mathbb{R}$, are, in general, incompatible propositions: their truth values interfere with each other when we measure them within lapses as short as we like (so that the system's state is not responsible for the time evolution). We know no instrument exists that is capable of evaluating simultaneously two incompatible quantities. Hence it is physically meaningless, in this context, to say that the above propositions, associated to incompatible physical quantities, can assume on the system a certain truth value *simultaneously*. The propositions $P_J^{(A)}$ and $P_K^{(B)}$, in this sense, are called **incompatible**.

Important remark. It has to be clear that the propositions we are considering must be understood as statements about physical systems to which we assign a truth value, 0 or 1, *as a consequence of a corresponding experimental measuring process*. In this light the *incompatibility* of two propositions does not imply they cannot be both true, so that, for example, their conjuction is always false. The meaning is much deeper: incompatible refers to the fact that *it makes no (physical) sense to give them, simultaneously, any truth value*. Nor is it possible to make sense of propositions like $P_J^{(A)} \mathscr{O} P_K^{(B)}$ or $P_J^{(A)} \mathscr{E} P_K^{(B)}$, *because there is no experiment that can evaluate the truth of such propositions.* ■

By this remark we cannot assume, as model for the set of elementary propositions to be tested on our quantum system, a σ-algebra of sets *where* $\cap$ *and* $\cup$ *are interpreted as* $\mathscr{E}$ *and* $\mathscr{O}$ *respectively*. If we were to do so, we would then have to impose constraints on the model, e.g. veto certain symbolic combinations built connecting incompatible propositions. An alternative idea of von Neumann turned out to be successful: model elementary propositions via the *orthogonal projectors* of a complex Hilbert space. As we will see, the set of projectors is a lattice; although the structure is not a Boolean σ-algebra, it will allow us to distinguish among compatible and incompatible propositions, and to interpret $\mathscr{E}$ and $\mathscr{O}$ as the standard $\wedge$ and $\vee$ provided the former are used with compatible propositions.

7.3.1 The non-distributive lattice of orthogonal projectors on a Hilbert space

The set of orthogonal projectors on a Hilbert space enjoys certain properties, close to those of Boolean lattices, but with important differences that let us model incompatible propositions of a quantum system. First of all we deal with a number of technical features of commuting projectors.

Proposition 7.16. *Let* $(\mathsf{H}, (\,|\,))$ *be a Hilbert space and* $\mathfrak{P}(\mathsf{H})$ *the set of orthogonal projectors on* H.
The following properties hold for any $P, Q \in \mathfrak{P}(\mathsf{H})$.

(a) *The following facts are equivalent:*

(i) $P \leq Q$;
(ii) $P(\mathsf{H})$ *is a subspace of* $Q(\mathsf{H})$;
(iii) $PQ = P$;
(iv) $QP = P$.

(b) *The following facts are equivalent:*

(i) $PQ = 0$*;*
(ii) $QP = 0$*;*
(iii) $P(\mathsf{H})$ *and* $Q(\mathsf{H})$ *are orthogonal;*
(iv) $Q \leq I - P$*;*
(v) $P \leq I - Q$*.*

If (i)-(v) hold, $P + Q$ *is an orthogonal projector onto* $P(\mathsf{H}) \oplus Q(\mathsf{H})$*.*
(c) *If* $PQ = QP$ *then* PQ *is an orthogonal projector onto* $P(\mathsf{H}) \cap Q(\mathsf{H})$*.*
(d) *If* $PQ = QP$ *then* $P + Q - PQ$ *is an orthogonal projector onto the closed space* $< P(\mathsf{H}), Q(\mathsf{H}) >$*.*
(e) $PQ = QP$ *iff there exist* $R_1, R_2, R_3 \in \mathfrak{P}(\mathsf{H})$ *such that:*

$$P = R_1 + R_3\,, \quad Q = R_1 + R_2 \quad \textit{with } R_i(\mathsf{H}) \perp R_j(\mathsf{H}) \textit{ if } i \neq j\,.$$

Proof. (a) First, notice that if P is a projector onto M, then $Pu = 0$ if and only if $u \in \mathsf{M}^\perp$, by the orthogonal decomposition $\mathsf{H} = \mathsf{M} \oplus \mathsf{M}^\perp$ (Theorem 3.13(d)) and because the component of u on M is precisely Pu.
(i) $\Rightarrow$ (ii). If $P \leq Q$ then $(u|Qu) \geq (u|Pu)$. Since projectors are idempotent and self-adjoint, the latter is equivalent to $(Qu|Qu) \geq (Pu|Pu)$, i.e. $||Qu|| \geq ||Pu||$. In particular $Qu = 0$ implies $Pu = 0$, so $Q(\mathsf{H})^\perp \subset P(\mathsf{H})^\perp$. Using Theorem 3.13(e) and noting $Q(\mathsf{H})$ and $P(\mathsf{H})$ are closed, we find $P(\mathsf{H}) \subset Q(\mathsf{H})$.
(ii) $\Rightarrow$ (iii). If S is a basis for $P(\mathsf{H})$, complete it to a basis of $Q(\mathsf{H})$ by adding the orthogonal set S' to S. By Proposition 3.58(d), $P = s\text{-}\sum_{u \in S} u(u|\;)$ and $Q = s\text{-}\sum_{u \in S \cup S'} u(u|\;)$. Since S and S' are orthogonal, orthonormal systems, and because the inner product is continuous, the claim follows.
(iii) $\Leftrightarrow$ (iv). The statements follow from each other by taking adjoints.
(iii) + (iv) $\Rightarrow$ (i). If $u \in \mathsf{H}$, $(u|Qu) = ((P + P^\perp)u|Q(P + P^\perp)u)$ where $P^\perp = I - P$. Notice P and $P^\perp$ commute with Q by (iii) and (iv), and moreover $PP^\perp = P^\perp P = 0$. Expanding the right side of $(u|Qu) = (u|(P + P^\perp)Q(P + P^\perp)u)$, and neglecting terms that are null by the above considerations, gives

$$(u|Qu) = (u|PQPu) + (u|P^\perp QP^\perp u)\,.$$

On the other hand by (iii) and (iv): $(u|PQPu) = (u|PPu) = (u|Pu)$. Therefore

$$(u|Qu) = (u|Pu) + (u|P^\perp QP^\perp u)\,,$$

so $(u|Qu) \geq (u|Pu)$.
(b) Assuming $PQ = 0$ and taking adjoints gives $QP = 0$, hence $P(\mathsf{H})$ and $Q(\mathsf{H})$ are orthogonal, for $PQ = QP = 0$. If $P(\mathsf{H})$ and $Q(\mathsf{H})$ are ortogonal, fix on each a basis, N and N' respectively, by writing P and Q as prescribed by Proposition 3.58(d): $P = \sum_{u \in N}(u|\;)u$, $Q = \sum_{u \in N'}(u|\;)u$. Immediately, $PQ = QP = 0$. At last, $Q \leq I - P$ ($P \leq I - Q$) iff Q (resp. P) projects onto a subspace in the orthogonal to $P(\mathsf{H})$ (resp. $Q(\mathsf{H})$) by part (a), i.e. $P(\mathsf{H}) \perp Q(\mathsf{H})$. Using the above expressions for P, Q, recalling $N \cup N'$ is a basis of $P(\mathsf{H}) \oplus Q(\mathsf{H})$ and using again Proposition 3.58(d), implies $P + Q$ is the orthogonal projector onto $P(\mathsf{H}) \oplus Q(\mathsf{H})$.

(c) That PQ is an orthogonal projector (self-adjoint and idempotent) if $PQ = QP$, with P, Q orthogonal projectors, is straighforward. If $u \in \mathsf{H}$, then $PQu \in P(\mathsf{H})$ but also $PQu = QPu \in Q(\mathsf{H})$, so $PQu \in P(\mathsf{H}) \cap Q(\mathsf{H})$. We have shown $PQ(\mathsf{H}) \subset P(\mathsf{H}) \cap Q(\mathsf{H})$, so to conclude it suffices to see $P(\mathsf{H}) \cap Q(\mathsf{H}) \subset PQ(\mathsf{H})$. If $u \in P(\mathsf{H}) \cap Q(\mathsf{H})$ then $Pu = u$, $Qu = u$, so also $Pu = PQu = u$, i.e. $u \in PQ(\mathsf{H})$. This means $P(\mathsf{H}) \cap Q(\mathsf{H}) \subset PQ(\mathsf{H})$.
(d) That $R := P + Q - PQ$ is an orthogonal projector is straighforward. Consider the space $\overline{< P(\mathsf{H}), Q(\mathsf{H}) >}$. We can build a basis as follows. Begin with a basis N for the closed subspace $P(\mathsf{H}) \cap Q(\mathsf{H})$. Then add a basis for the space that "remains in $P(\mathsf{H})$ once $P(\mathsf{H}) \cap Q(\mathsf{H})$ has been taken out", i.e. a basis N' for the closed orthogonal complement to $P(\mathsf{H}) \cap Q(\mathsf{H})$ in $P(\mathsf{H})$: this is $P(\mathsf{H}) \cap (P(\mathsf{H}) \cap Q(\mathsf{H}))^\perp$. At last build with the same criterion a third basis N'' for $Q(\mathsf{H}) \cap (P(\mathsf{H}) \cap Q(\mathsf{H}))^\perp$. The three bases thus obtained are pairwise orthogonal and together give a basis of $\overline{< P(\mathsf{H}), Q(\mathsf{H}) >}$. All this shows that

$$\overline{< P(\mathsf{H}), Q(\mathsf{H}) >} =$$

$$(P(\mathsf{H}) \cap Q(\mathsf{H})) \oplus (P(\mathsf{H}) \cap (P(\mathsf{H}) \cap Q(\mathsf{H}))^\perp) \oplus (Q(\mathsf{H}) \cap (P(\mathsf{H}) \cap Q(\mathsf{H}))^\perp)$$

is an othogonal sum. With our assumptions the projector on the first summand is PQ by (c). Thus the projector on $(P(\mathsf{H}) \cap Q(\mathsf{H}))^\perp$ is $I - PQ$. Again by (c) the orthogonal projector on the second summand is $P(I - PQ) = P - PQ$, and similarly the third projector is

$$Q(I - PQ) = Q - PQ\,.$$

By part (b) the projector onto the whole sum $\overline{< P(\mathsf{H}), Q(\mathsf{H}) >}$ is

$$PQ + (P - PQ) + (Q - PQ) = P + Q - PQ\,.$$

Statement (e) is another way to phrase Proposition 3.60. □

Based on what we have proved, consider two orthogonal projectors $P, Q \in \mathfrak{P}(\mathsf{H})$ that commute, and suppose they are associated to statements about the physical system (i.e. propositions, denoted by the same letters). Under the correspondence

$$\begin{aligned} P \,\mathscr{E}\, Q &\longleftrightarrow PQ\,, \\ P \,\mathscr{O}\, Q &\longleftrightarrow P + Q - PQ\,, \\ \dashv P \quad &\longleftrightarrow I - P\,, \end{aligned}$$

the right-hand sides are orthogonal projectors. The latter, moreover, satisfy properties that are formally identical to those of propositional calculus. For example, $\dashv (P \,\mathscr{E}\, Q) = \dashv P \,\mathscr{O} \dashv Q$. In fact,

$$\dashv P \,\mathscr{O} \dashv Q \longleftrightarrow (I - P) + (I - Q) - (I - P)(I - Q) = 2I - P - Q - I + PQ + P + Q$$

$$= I - PQ \longleftrightarrow \dashv (P \,\mathscr{E}\, Q)$$

and in the same way one may check *every* relation written previously, provided the projectors commute. Note, further, that if P, Q commute and $P \leq Q$ then $PQ = QP = P$ and $P + Q - PQ = Q$. If we interpret the latter by their truth value we have a similar situation, by the above correspondence, to $P \mathscr{E} Q$ and $P \mathscr{O} Q$ when P logically implies Q. Thus we may say that $P \leq Q$ corresponds to Q being logical consequence of P.

The real difference between orthogonal projectors and the propositions of a classical system is the following. If the projectors P, Q *do not commute*, PQ and $P+Q-PQ$ are not even projectors in general, so *the above correspondence breaks down*.

All this seems very interesting in order to find a model for the propositions of quantum system, under axiom **QM2**. The idea is that:

the propositions of quantum systems are in 1–1 *correspondence with the orthogonal projectors of a Hilbert space. This is such that:*

(i) *the logical implication between propositions P and Q ($P \Rightarrow Q$) corresponds to the relation $P \leq Q$ of the corresponding projectors;*
(ii) *two propositions are compatible if and only if the respective projectors commute.*

Remarks 7.17. Before going any further let us shed some light on the nature of commuting orthogonal projectors. One would be led to suspect that the only cases where P and Q commute are if: (a) projection spaces are one contained in the other, or (b) projection spaces are orthogonal. With the following explicit example we show that there are other possibilities. Consider the space $L^2(\mathbb{R}^2, dx \otimes dy)$, dx, dy being Lebesgue measures on the real line, and the sets in the plane $A = \{(x,y) \in \mathbb{R}^2 \mid a \leq x \leq b\}$, $a < b$ given, and $B = \{(x,y) \in \mathbb{R}^2 \mid c \leq y \leq d\}$, $c < d$. If $G \subset \mathbb{R}^2$ is measurable, define the linear operator

$$P_G : L^2(\mathbb{R}^2, dx \otimes dy) \to L^2(\mathbb{R}^2, dx \otimes dy)$$

by $P_G f = \chi_G \cdot f$ for any $f \in L^2(\mathbb{R}^2, dx \otimes dy)$, where χ_G is, as always, the characteristic function of G and $\cdot$ is the pointwise product of two maps. The operator P_G is an orthogonal projector, and moreover

$$P_G(L^2(\mathbb{R}^2, dx \otimes dy)) = \{f \in L^2(\mathbb{R}^2, dx \otimes dy) \mid \mathit{ess\,supp}\, f \subset G\}\,.$$

Then it is immediate to prove $P_A P_B = P_B P_A = P_{A \cap B}$, whilst:

(a) none of the projection spaces $P_A(L^2(\mathbb{R}^2, dx \otimes dy))$, $P_B(L^2(\mathbb{R}^2, dx \otimes dy))$ is included in the other, and
(b) the two projection spaces are not orthogonal. ■

If the speculative correspondence between propositions about quantum systems and orthogonal projectors on a suitable Hilbert space is to be meaningful, the structural analogies of orthogonal projectors and the σ-complete Boolean algebra of propositions must reach farther than the case of two propositions. We expect, in particular, to determine the structure of a Boolean (σ-)algebra on some set of projectors representing pairwise-compatible properties. The following fact asserts that the space of *all* orthogonal projectors is a non-distributive lattice, and establishes some of its peculiarities. Referring to part (c) let us remark that if $A \subset \mathfrak{P}(\mathsf{H})$ is a set of commuting orthogonal projectors, by Zorn's lemma there exists a maximal commutative

$\mathfrak{P}_0(\mathsf{H}) \subset \mathfrak{P}(\mathsf{H})$ with $A \subset \mathfrak{P}_0(\mathsf{H})$: every projector in $\mathfrak{P}(\mathsf{H})$ commuting with any element in $\mathfrak{P}_0(\mathsf{H})$ belongs to $\mathfrak{P}_0(\mathsf{H})$.

Theorem 7.18. *Let* H *be a (complex) Hilbert space.*

(a) *The collection* $\mathfrak{P}(\mathsf{H})$ *of orthogonal projectors on* H *is a bounded, orthocomplemented,* σ*-complete lattice, typically non-distributive. More precisely:*

(i) $\geq$ *is the order relation between projectors;*
(ii) *the maximum and minimum elements in* $\mathfrak{P}(\mathsf{H})$ *are: I (identity operator) and* 0 *(null operator) respectively;*
(iii) *the orthocomplement to the projector P corresponds to*

$$\neg P = I - P\,; \tag{7.6}$$

(iv) *the projection spaces of* $P, Q \in \mathfrak{P}(\mathsf{H})$ *are orthogonal iff they are orthogonal as elements in the orthocomplemented lattice* $\mathfrak{P}(\mathsf{H})$*;*
(v) *two projectors* $P, Q \in \mathfrak{P}(\mathsf{H})$ *commute iff they commute as elements in the orthocomplemented lattice* $\mathfrak{P}(\mathsf{H})$*;*
(vi) $\mathfrak{P}(\mathsf{H})$ *is not distributive if dim* $\mathsf{H} \geq 2$.

(b) *On* $\mathfrak{P}(\mathsf{H})$ *the following hold:*

(i) *if* $P, Q \in \mathfrak{P}(\mathsf{H})$ *commute:*

$$P \wedge Q = PQ\,, \tag{7.7}$$
$$P \vee Q = P + Q - PQ\,; \tag{7.8}$$

(ii) *if* $\{Q_n\}_{n\in\mathbb{N}} \subset \mathfrak{P}(\mathsf{H})$ *consists of commuting elements:*

$$\vee_{n\in\mathbb{N}} Q_n = s\text{-}\lim_{n\to+\infty} Q_0 \vee \cdots \vee Q_n\,, \tag{7.9}$$
$$\wedge_{n\in\mathbb{N}} Q_n = s\text{-}\lim_{n\to+\infty} Q_0 \wedge \cdots \wedge Q_n\,, \tag{7.10}$$

independently of the labelling of the Q_n.

(c) *If* $\mathfrak{P}_0(\mathsf{H}) \subset \mathfrak{P}(\mathsf{H})$ *is a maximal commutative set of orthogonal projectors, then* $\mathfrak{P}_0(\mathsf{H})$ *is a Boolean* σ*-subalgebra. In particular* $\mathfrak{P}_0(\mathsf{H}) \ni 0, I$, $\mathfrak{P}_0(\mathsf{H})$ *is closed under orthocomplementation, the inf and sup of a countable subset in* $\mathfrak{P}_0(\mathsf{H})$ *exist in* $\mathfrak{P}_0(\mathsf{H})$ *and coincide with the inf and sup on* $\mathfrak{P}(\mathsf{H})$.

Proof. (a) Recall $\geq$ is a partial order on $\mathfrak{P}_0(\mathsf{H})$ by Proposition 3.54(f). By Proposition 7.16(a):

$$P \leq Q \;\text{ iff }\; P(\mathsf{H}) \subset Q(\mathsf{H})\,. \tag{7.11}$$

This partial order of orthogonal projectors corresponds one-to-one to the partial order of projection spaces. The class of closed subspaces in H is a lattice: we claim that if M, N are closed, their least upper bound and greatest lower bound are $\mathsf{M} \vee \mathsf{N} = \overline{< \mathsf{M}, \mathsf{N} >}$ and $\mathsf{M} \wedge \mathsf{N} = \mathsf{M} \cap \mathsf{N}$ respectively. Now, $\overline{< \mathsf{M}, \mathsf{N} >}$ is closed and contains M, N; moreover, any closed space L containing M, N must contain $\overline{< \mathsf{M}, \mathsf{N} >}$ as well, so $\mathsf{M} \vee \mathsf{N} = \overline{< \mathsf{M}, \mathsf{N} >}$. $\mathsf{M} \cap \mathsf{N}$ closed in M and N by construction, and if L is another such space, it must be contained in $\mathsf{M} \cap \mathsf{N}$, whence $\mathsf{M} \wedge \mathsf{N} = \mathsf{M} \cap \mathsf{N}$. Passing to projectors and using (7.11), we have that for $P, Q \in \mathfrak{P}(\mathsf{H})$, $P \vee Q$ is the orthogonal projector onto $\overline{< P(\mathsf{H}), Q(\mathsf{H}) >}$, while $P \wedge Q$ the projector onto $P(\mathsf{H}) \cap Q(\mathsf{H})$. The same argument applies to a family of orthogonal projectors $\{P_i\}_{i\in I}$ of arbitrary cardinality. In that case $\vee_{i\in I} P_i$ is the projector onto $\overline{< \{P_i(\mathsf{H})\}_{i\in I} >}$ and $\wedge_{i\in I} P_i$ the projector on $\cap_{i\in I} P_i(\mathsf{H})$, so the lattices of orthogonal projectors and closed subspaces are both σ-complete. In the lattice of closed subspaces the min and max are clearly $\{0\}$ and H. Passing to orthogonal projectors via (7.11), the minimum and maximum are the orthogonal projectors onto $\{0\}$ and H, i.e. the null operator and the identity. Orthocomplementation of projectors, $\neg P := I - P$, corresponds to complementation of closed subspaces $\neg P(\mathsf{M}) := P(\mathsf{M})^{\perp}$, by Proposition 3.58(b). Checking the properties of orthocomplementation for subspaces (hence projectors) is then immediate by (b), (d), (e) in Theorem 3.13. Part (iv) in (a) follows directly from Proposition 7.16(b), whilst (v) in (a) descends from Proposition 7.16(e). To prove (vi), we exhibit a counterexample to distributivity.

Consider a two-dimensional subspace S in a (complex) Hilbert space H of dimension ≥ 2. Idenitify S with $\mathbb{C}^2$ by fixing an orthonormal basis $\{e_1, e_2\}$. Now consider the subspaces: $\mathsf{H}_1 :=< e_1 >$, $\mathsf{H}_2 :=< e_2 >$ and $\mathsf{H}_3 :=< e_1 + e_2 >$.

From $\mathsf{H}_1 \wedge (\mathsf{H}_2 \vee \mathsf{H}_3) = \mathsf{H}_1 \wedge S = \mathsf{H}_1$ and $(\mathsf{H}_1 \wedge \mathsf{H}_2) \vee (\mathsf{H}_1 \wedge \mathsf{H}_3) = \{0\} \vee \{0\} = \{0\}$ follows

$$\mathsf{H}_1 \wedge (\mathsf{H}_2 \vee \mathsf{H}_3) \neq (\mathsf{H}_1 \wedge \mathsf{H}_2) \vee (\mathsf{H}_1 \wedge \mathsf{H}_3)\,.$$

Let us prove (b) and (c) together. If the projectors P and Q commute, or if $\{Q_n\}_{n\in\mathbb{N}}$ pairwise commute, by Zorn's lemma there is a maximal commuting $\mathfrak{P}_0(\mathsf{H})$ containing P, Q, or $\{Q_n\}_{n\in\mathbb{N}}$ respectively. Now choose such a maximal element $\mathfrak{P}_0(\mathsf{H})$ as in the proof of part (b).

Clearly 0 and I belong to $\mathfrak{P}_0(\mathsf{H})$ because they commute with everything in $\mathfrak{P}_0(\mathsf{H})$. The same happens for $\neg P = I - P$ if $P \in \mathfrak{P}_0(\mathsf{H})$. We have to prove, for any $P, Q \in \mathfrak{P}_0(\mathsf{H})$, the existence of the sup and the inf of $\{P, Q\}$ inside $\mathfrak{P}_0(\mathsf{H})$, that they are computed as prescribed in part (b), and that these projectors actually coincide with the sup and inf of $\{P, Q\}$ inside $\mathfrak{P}(\mathsf{H})$. The distributive laws of $\vee$ and $\wedge$ follow easily from (7.8) and (7.7), from the projectors' commutation and from the idempotency of any projector, $PP = P$.

By Proposition 7.16(c), the projector onto $M \cap N$, corresponding to $P \wedge Q$ in $\mathfrak{P}(\mathsf{H})$, is exactly PQ, and this belongs to $\mathfrak{P}_0(\mathsf{H})$ because by construction it commutes with any element of the maximal $\mathfrak{P}_0(\mathsf{H})$. Therefore

$$P \wedge Q := \inf_{\mathfrak{P}_0(\mathsf{H})} \{P, Q\} = \inf_{\mathfrak{P}(\mathsf{H})} \{P, Q\} = PQ\,.$$

As P, Q commute, the projector onto $\overline{< \mathsf{M}, \mathsf{N} >}$, corresponding to $P \vee Q$ in $\mathfrak{P}(\mathsf{H})$, is $P + Q - PQ$ by Proposition 7.14(d); the latter lives in $\mathfrak{P}_0(\mathsf{H})$ for it commutes with $\mathfrak{P}_0(\mathsf{H})$. As before,

$$P \vee Q := \sup_{\mathfrak{P}_0(\mathsf{H})} \{P, Q\} = \sup_{\mathfrak{P}(\mathsf{H})} \{P, Q\} = P + Q - PQ\,.$$

This makes $\mathfrak{P}_0(\mathsf{H})$ a Boolean algebra.

To conclude we show $\mathfrak{P}_0(\mathsf{H})$ is σ-complete. Consider a countable family of projectors $\{Q_n\}_{n\in\mathbb{N}}$ and associate to each the projector P_n defined recursively by: $P_0 := Q_0$, and for $n = 1,2,\ldots$:

$$P_n := Q_n(I - P_1 - \ldots - P_{n-1})\,.$$

By induction we can prove with ease:

(i) $P_nP_m = 0$ if $n \neq m$;
(ii) $Q_1 \vee \cdots \vee Q_n = P_1 \vee \cdots \vee P_n = P_1 + \cdots + P_n$, $n = 0,1,\ldots$.

If we introduce bounded operators

$$A_n := P_1 + \cdots + P_n,$$

then:

(iii) $A_n = A_n^*$ and $A_nA_n = A_n$ for any $n = 0,1,\ldots$, i.e. the A_n are orthogonal projectors, so $A_n \leq I$, for any $n = 0,1,\ldots$ by Proposition 3.58(e);
(iv) $A_n \leq A_{n+1}$ for any $n = 0,1,\ldots$

By virtue of Proposition 3.65 there exists a bounded self-adjoint operator A defined by the strong limit:

$$A = \text{s-}\lim_{n\to+\infty} P_n = \text{s-}\lim_{n\to+\infty} Q_0 \vee \cdots \vee Q_n\,.$$

Immediately, then, $AA = A$, making A an orthogonal projector in $\mathfrak{P}_0(\mathsf{H})$ because (strong) limit of operators commuting with $\mathfrak{P}_0(\mathsf{H})$. Still by Proposition 3.65, $A_n \leq A$ and in particular $Q_n \leq Q_1 \vee \cdots \vee Q_n \leq A$ for any $n \in \mathbb{N}$. We claim A is the least upper bound of the Q_n, in $\mathfrak{P}(\mathsf{H})$ and in $\mathfrak{P}_0(\mathsf{H})$. Suppose the orthogonal projector $K \in \mathfrak{P}(\mathsf{H})$ satisfies $K \geq Q_n$ for any $n \in \mathbb{N}$. Then $KQ_n = Q_n$ by Proposition 7.16(a). By definition of the P_n, $KP_n = P_n$ and hence $KA_n = A_n$, so also $K \geq A_n$ for any natural number n, by Proposition 7.16(a). Now Proposition 3.65 warrants $K \geq A$. In other words $A \in \mathfrak{P}_0(\mathsf{H})$ bounds the Q_n from above, and any other upper bound $K \in \mathfrak{P}(\mathsf{H})$ is larger than A. By definition of sup, $A = \sup_{\mathfrak{P}(\mathsf{H})}\{Q_n\}_{n\in\mathbb{N}} =: \vee_{n\in\mathbb{N}} Q_n$. As $A \in \mathfrak{P}_0(\mathsf{H})$, A is also the sup in $\mathfrak{P}_0(\mathsf{H})$. In the above identity

$$\vee_{n\in\mathbb{N}} Q_n = \text{s-}\lim_{n\to+\infty} Q_0 \vee \cdots \vee Q_n$$

the indexing order of the Q_n is not relevant, given that the left-hand side, i.e. the supremum of $\{Q_n\}_{n\in\mathbb{N}}$, does not depend on any ordering. Formula (7.10) is easy using $\neg$ and (7.9). □

This result explains that it makes sense to describe propositions about quantum system in terms of the (non-Boolean) lattice of orthogonal projectors on a Hilbert space, and incompatible propositions in terms of non-commuting projectors. Moreover, it makes sense to assign the usual meaning to $\wedge$, $\vee$ in terms of the connectives $\mathscr{E}$, $\mathscr{O}$, provided the former are employed with projectors describing compatible propositions.

Foundational studies on the role of the lattice of projectors, in relationship to the logical formulation of QM, are found in [Mac63, Jau73, Pir76, BeCa81, DCGi02, Var07, EGL09] besides [Bon97]. The reader can find a different approach in [Emc72]: based on *Jordan algebras*, it prepares for the *algebraic formulation* following ideas of Segal.

7.3.2 Recovering the Hilbert space from the lattice

A reasonable question to ask is whether there are better reasons for choosing to describe quantum systems via a lattice of orthogonal projectors, other than the kill-off argument "it works". To tackle the problem we start by listing special properties of the projectors's lattice, whose proof is elementary.

Theorem 7.19. *The bounded, orthocomplemented, σ-complete lattice $\mathfrak{P}(\mathsf{H})$ of Theorem 7.18 satisfies these additional properties:*

(i) **separability** *(for H separable): if $\{P_a\}_{a\in A} \subset \mathfrak{P}(\mathsf{H})$ satisfies $P_i \perp P_j$, $i \neq j$, then A is at most countable;*
(ii) **atomicity***: there exist elements in $A \in \mathfrak{P}(\mathsf{H}) \setminus \{0\}$, called* **atoms***, for which $0 \leq P \leq A$ implies $P = 0$ or $P = A$; for any $P \in \mathfrak{P}(\mathsf{H}) \setminus \{0\}$ there exists an atom A with $A \leq P$ ($\mathfrak{P}(\mathsf{H})$ is then called* **atomic***);*
(iii) **orthomodularity***: $P \leq Q$ implies $Q = P \vee ((\neg P) \wedge Q)$;*
(iv) **covering property***: if $A, P \in \mathfrak{P}(\mathsf{H})$, A atom, satisfy $A \wedge P = 0$, then (1) $P \leq A \vee P$ with $P \neq A \vee P$, and (2) $P \leq Q \leq A \vee P$ implies $Q = P$ or $Q = A \vee P$;*
(v) **irreducibility***: only 0 and I commute with every element of $\mathfrak{P}(\mathsf{H})$.*

The only orthogonal projectors onto one-dimensional spaces are the atoms of $\mathfrak{P}(\mathsf{H})$.

Based on the experimental evidence of quantum systems we could try to prove, in absence of any Hilbert space, that elementary propositions with experimental outcome in $\{0,1\}$ form a poset. More precisely, we could attempt to find a bounded, orthocomplemented σ-complete lattice that verifies conditions (i)–(v) above, and then prove this lattice is described by the orthogonal projectors of a Hilbert space.

The partial order relation of elementary propositions can be variously defined. But it will always correspond to the logical implication, in some way or another. Starting from [Mac63] a number of approaches (either of essentially physical nature, or of formal character) have been developed to this end: in particular, those making use of the notion of quantum state, which we will see in a short while for the concrete case of propositions represented by orthogonal projectors. We refer to the aforementioned literature for more information. More difficult is to justify that the poset thus obtained is a lattice, i.e. that it admits greatest lower bound $P \vee Q$ and least upper bound $P \wedge Q$ for any P, Q. There are several proposals, very differing in nature, to introduce this lattice structure [Jau73, Mac63] (see [BeCa81] and [EGL09] for a general treatise) and make the physical meaning explicit in terms of measurement outcome. Althought we will not discuss them, let us just mention a suggestive idea (albeit ridden with issues) of Jauch, based on a result of von Neumann (Theorem 13.7 in [Neu50]).

Theorem 7.20 (von Neumann's theorem on iterated projectors). *Let* H *be a complex Hilbert space and* $P, Q : \mathsf{H} \to \mathsf{H}$ *orthogonal projectors, in general not commuting. Calling, as usual,* $P \wedge Q$ *the orthogonal projector onto* $P(\mathsf{H}) \cap Q(\mathsf{H})$*, we have:*

$$(P \wedge Q)x = \lim_{n \to +\infty} (PQ)^n x \quad \textit{for any } x \in \mathsf{H}. \tag{7.12}$$

Proof. See Exercise 7.5. □

There is an extremely interesting physical point of view that interprets the right-hand side of (7.12) as the *consecutive and alternated* measurement of an *infinite* sequence of propositions P, Q. Proposition $P \wedge Q$ is true for a state of a quantum system only if all propositions in the sequence are true.

If we accept the lattice structure on elementary propositions of a quantum system, then we may define the operation of orthocomplementation by the familiar logical/physical negation. Compatible propositions can then be defined in terms of commuting propositions as of Definition 7.8 (by (v) in Theorem 7.18(a) this notion of compatibility is the usual one when propositions are interpreted via projectors). Now fully-fledged with an orthocomplemented lattice and the notion of compatible propositions, we can attach a physical meaning (an interpretation backed by experimental evidence) to the requests that the lattice be bounded, complete, atomic, irreducible and that it have the covering property [BeCa81]. Under these hypotheses and assuming there exist at least 4 pairwise-orthogonal atoms, Piron ([Pir64, JaPi69], [BeCa81, Chapter 21], Aerts in [EGL09]) used projective geometry techniques to show that the lattice of quantum propositions can be canonically identified with the closed (in a generalised sense) subsets of a generalised Hilbert space of sorts. In the latter: (a) the field is replaced by a division ring (usually not commutative) equipped with an involution, and (b) there exists a certain definite Hermitian form associated with the involution. It has been conjectured by many people (see [BeCa81]) that if the lattice is also orthomodular and separable, the division ring can only be picked among $\mathbb{R}, \mathbb{C}$ or $\mathbb{H}$ (quaternion algebra). More recently Solèr[6], Holland[7] and Aerts–van Steirteghem[8] have found sufficient hypotheses, in terms of the existence of orthogonal systems, for this to happen. If the ring is $\mathbb{R}$ or $\mathbb{C}$ and we further assume the involution is continuous in the topology induced by the Hermitian inner product, we obtain precisely the lattice of orthogonal projectors of the separable Hilbert space [BeCa81]. In case of $\mathbb{H}$ one gets a similar structure. In all these arguments the assumption of irreducibility is not really crucial: if property (v) fails, the lattice can be split into irreducible sub-lattices [Jau73, BeCa81]. Physically-speaking this situation is natural in presence of *superselection rules*, of which more soon.

Our study takes place in complex Hilbert spaces, and $R := P \wedge Q$ denotes simply the projector onto the intersection of the targets of P, Q. R may or not have a meaning

[6] Solèr M.P.: Characterization of Hilbert spaces by orthomodular spaces. Communications in Algebra **23**, 219–243 (1995).

[7] Holland S.S.: Orthomodularity in infinite dimensions; a theorem of M. Solèr. Bulletin of the American Mathematical Society **32**, 205–234 (1995).

[8] Aerts D., van Steirteghem B.: Quantum Axiomatics and a theorem of M.P. Solèr. International Journal of Theoretical Physics **39**, 497–502 (2000).

as statement about the system, but as we have noted earlier it *does not correspond to the proposition P & Q when P, Q relate to incompatible propositions*. Conversely, the approach of Birkhoff and von Neumann, that befits the so-called *standard quantum logic*, uses $\vee$ and $\wedge$ as proper connectives (yielding an algebra different from the usual one), even if they operate between projectors of incompatible propositions (i.e. for which no instrument can evaluate the truth of P, Q simultaneously. This is the reason why the viewpoint of *Quantum Logic* has been criticised by physicists (cf. [Bon97, Chapter 5] for a thorough discussion). In the past years, alongside the modern development of Birkhoff's and von Neumann's approach [EGL09], many authors have introduced new formal strategies that differ from *Quantum Logic à la* Birkhoff–von Neumann, in particular by means of *topos theory* [DI08, HLS09].

7.3.3 Von Neumann algebras and the classification of factors

An important point for developing von Neumann's theory is that the lattice of elementary propositions on a quantum system should satisfy the *modularity* condition (Definition 7.8(d)). We will not go into explaining the manifold reasons for this (see Rédei in [EGL09]). It will be enough to remark, as von Neumann himself proved, that $\mathfrak{P}(\mathsf{H})$ is not modular if H does not have finite dimension, due to the existence of unbounded self-adjoint operators whose domains intersect trivially (in the null vector). The way out proposed by von Neumann and Murray is to reduce the number of elementary observables on the quantum system, so to guarantee modularity. The point is to start from a *von Neumann algebra* $\mathfrak{R}$ (Definition 3.47) on a separable Hilbert space H, as opposed to the lattice of projectors $\mathfrak{P}(\mathsf{H})$. As the double commutant theorem (Theorem 3.46) shows, $\mathfrak{R} = \mathfrak{P}_{\mathfrak{R}}(\mathsf{H})''$, where $\mathfrak{P}_{\mathfrak{R}}(\mathsf{H})$ indicates the subset in $\mathfrak{P}(\mathsf{H})$ of orthogonal projectors belonging to $\mathfrak{R}$. It can be proved $\mathfrak{P}_{\mathfrak{R}}(\mathsf{H})$ inherits the structure of bounded, orthomodular, σ-complete lattice from $\mathfrak{P}(\mathsf{H})$. Hence $\mathfrak{P}_{\mathfrak{R}}(\mathsf{H})$ represents elementary propositions associated to the physical system; since the spectral theorem (see Chapter 9) uses the strong topology, for which $\mathfrak{R} = \mathfrak{P}_{\mathfrak{R}}(\mathsf{H})''$ is closed, self-adjoint elements of $\mathfrak{R}$ determine the bounded observables on the system. Ultimately it is possible to choose $\mathfrak{R}$ so that $\mathfrak{P}_{\mathfrak{R}}(\mathsf{H})$ is modular. This happens for special von Neumann algebras, belonging to the class of so-called *factors*: von Neumann algebras $\mathfrak{R}$ with trivial centre, $\mathfrak{R}' \cap \mathfrak{R} = \{cI\}_{c\in\mathbb{C}}$. The classification of factors of Murray and von Neumann proves, among other things, that the lattice $\mathfrak{P}_{\mathfrak{R}}(\mathsf{H})$ inside the so-called *factors of type* II_1 is modular. Although modularity is nowadays no longer deemed fundamental, and some of von Neumann motivations have proved indefensible (see Rédei in [EGL09]), *von Neumann algebras*, *factors*, and the classification and study of factors have been decisive for the development of the mathematical formulation of quantum theories, especially quantum field theories [Haa96].

7.4 Propositions and states on quantum systems

In this section we discuss the first two axioms of the general formulation of QM, and explain the mathematical description of propositions and states of quantum systems by using a suitable Hilbert space. We characterise those states in an important theorem due to Gleason. We also show that quantum states form a convex set, and can be obtained as linear combinations of extreme states. The latter, called *pure states*, are in one-to-one correspondence with elements (*rays*) of the projective space associated to the physical system's Hilbert space.

7.4.1 Axioms A1 *and* A2*: propositions, states of a quantum system and Gleason's theorem*

Given what we saw in the previous section, we shall assume the following QM axiom. Propositions and projectors are denoted by the same symbol, as we have already done.

A1. *Let S be a quantum system described in a frame system $\mathscr{I}$. Then testable propositions on S at any given time correspond bijectively to (a subset of) the lattice* $\mathfrak{P}(\mathsf{H}_S)$ *(for the inclusion) of orthogonal projectors on a separable (complex) Hilbert space* H_S*, called* **Hilbert space associated to** S*. Moreover:*

(1) *compatible propositions correspond to commuting orthogonal projectors.*
(2) *the logical implication of compatible propositions $P \Rightarrow Q$ corresponds to the relation $P \leq Q$ of the associated projectors.*
(3) *I (identity operator) and* 0 *(null operator) correspond to the tautology and the contradiction.*
(4) *the negation $\neg P$ of a proposition P corresponds to the projector $\neg P = I - P$.*
(5) *only when the propositions P, Q are compatible, the propositions $P \,\mathscr{O}\, Q$, $P \,\mathscr{E}\, Q$ make physical sense and correspond to the projectors $P \vee Q$, $P \wedge Q$.*
(6) *if $\{Q_n\}_{n\in\mathbb{N}}$ is a countable set of pairwise-compatible propositions, the propositions corresponding to $\vee_{n\in\mathbb{N}} Q_n$, $\wedge_{n\in\mathbb{N}} Q_n$* make sense.

Remark 7.21. (1) Proper motivation for why H_S should be *separable* will be seen later, when we consider concrete quantum system and give an explicit representation of H_S. The assumption is also necessary to many theoretical results in this book.
(2) From now on we shall assume that the subset of $\mathfrak{P}(\mathsf{H}_S)$ describing the system's propositions is the entire $\mathfrak{P}(\mathsf{H}_S)$, leaving out for the moment *superselection rules*. As we saw in Chapter 7.3.3 the matter is quite subtle. A weaker assumption would be to have elementary propositions described by the sublattice of orthogonal projectors of a *von Neumann algebra* $\mathfrak{R}_S \subset \mathfrak{B}(\mathsf{H}_S)$, actually a *factor* if we neglect superselection rules. Self-adjoint elements $\mathfrak{R}_S$ identify with bounded observables on S, as we will have time to explain.
(3) The set of elementary propositions about a quantum system contains the subset of so-called **atomic propositions**, corresponding to atoms in the lattice of projectors (Theorem 7.19), i.e. orthogonal projectors onto *one-dimensional* subspaces. Atomic propositions are characterised as follows. P is atomic if there is no $P' \in \mathfrak{P}(\mathsf{H}_S)$ such that $P' \Rightarrow P$ apart from $P' = 0$, $P' = P$. Morevoer, if P, Q are distinct atomic propositions, they are compatible ($PQ = QP$) if and only if mutually exclusive: $PQ = 0$.

Supposing H_S separable, any $Q \in \mathfrak{P}(\mathsf{H}_S)$ can be written as disjuntion, at most countable, of atomic propositions, because of Proposition 3.58(d).

The existence of a propositional subset with the properties of atoms is physically remarkable and far from obvious.

That they exist in classical systems, by the way, is also not obvious at all (see [Jau73]). It implies that two propositions P, Q are compatible if and only if they can be written, separately, as disjuntion (at most countable) of sets N_P, N_Q of atomic propositions, so that the atomic propositions of the union $N_P \cup N_Q$ are pairwise compatible. The proof, by the above argument, follows immediately Proposition 3.60.

(4) Apart from S, the Hilbert space H_S depends on the frame system of choice. Picking a different (inertial) system boils down to having a new Hilbert space, but isomorphic to the former as we will see in Chapter 12, 13.

Another formulation, alternative to ours, to build the quantum formalism is the following. Given a quantum physical system S, one assigns to any instant $t \in \mathbb{R}$ a Hilbert space $\mathsf{H}_S(t)$ that *does not* depend on any frame system. This reminds of *absolute space at time t* in classical physics, a notion that is independent of frames. The way the various $\mathsf{H}_S(t)$ are related depends on the frame sistem and the time evolution, the latter described by isomorphisms between two $\mathsf{H}_S(t)$ with different t; it does not depend upon the chosen frame, in contrast to what we will obtain in Chapter 13 (albeit the formalism will be equivalent).

If we chose to use frame-independent (but time-depending) Hilbert spaces $\{\mathsf{H}_S(t)\}_{t\in\mathbb{R}}$, we would not be able to describe the evolution by a one-parameter group of unitary operators on the same Hilbert space. That is precisely what happens after having fixed a frame and the Hilbert space once and for all, as we will see in Chapter 13. [9] ■

Let us pass to the second axiom about quantum states. The point, relying on **QM1** and **QM2** (and ensuing remarks) is that a quantum state at time t gives the "probability" that every proposition of the system is true. So the idea is to generalise the notion of σ-additive probability measure. Instead of defining it on a σ-algebra, we must think of it as living on the set of associated projectors. We know every maximal set of compatible propositions defines a σ-finite Boolean algebra, itself an extension of a σ-algebra where measures are defined. So this is the natural principle.

A2 (preliminary form). *A state ρ at time t on a quantum system S is a mapping ρ : $\mathfrak{P}(\mathsf{H}_S) \to [0,1]$ such that:*

(1) $\rho(I) = 1$;

(2) *if $\{P_i\}_{i\in\mathbb{N}} \subset \mathfrak{P}(\mathsf{H}_S)$ satisfy $P_iP_j = 0$, $i \neq j$, then*

$$\rho\left(s\text{-}\sum_{i=0}^{+\infty} P_i\right) = \sum_{i=0}^{+\infty} \rho(P_i)\,.$$

[9] Although we will not do so, one could also use two-parameter groupoids of unitary transformations between different Hilbert spaces.

Remark 7.22. (1) Demand (1) just says the tautology is true on every state.
(2) Demand (2) clearly holds for finitely many propositions: it is enough that $P_i = 0$ definitely.
(3) (2) can be rephrased as:

$$\rho\left(\vee_{i\in\mathbb{N}}P_i\right) = \sum_{i=1}^{+\infty}\rho\left(P_i\right),$$

for any collection $\{P_i\}_{i\in\mathbb{N}} \subset \mathfrak{P}(\mathsf{H}_S)$ of compatible, mutually-exclusive propositions, so that $\sum_{i=0}^{+\infty} P_i = \vee_{i\in\mathbb{N}}P_i$ exists by Theorem 7.18.

The proof of the existence of $\sum_{i=0}^{+\infty} P_i$ is spelt out next, at any rate. Under the assumptions, partial sums give self-adjoint idempotent operators, hence orthogonal projectors. Therefore $\sum_{i=0}^{N} P_i \leq I$ by Proposition 3.58(e). Moreover $\sum_{i=0}^{N+1} P_i \geq \sum_{i=0}^{N} P_i$, as is easy to see. So by Proposition 3.65 the sequence admits strong limit. Immediately, this limit is idempotent and self-adjoint, hence a projector.
(4) Every state ρ clearly determines the equivalent of a positive σ-additive probability measure on any maximal commutative set of projectors $\mathfrak{P}_0(\mathsf{H}_s)$, which, as seen before, generalises a σ-algebra. Thus we have extended the notion of probability measure.
(5) The reader should beware identifying the "probability measure" ρ with an honest probability measure on a σ-algebra: *the fact we now consider quantum incompatible propositions alters drastically the rules of conditional probability*. The probability that "P is true when Q holds" abides by a different set of rules from the classical theory in case P and Q are incompatible in quantum sense.
(6) If H_S is separable a "probability measure" ρ on $\mathfrak{P}(\mathsf{H}_S)$ in the sense of **A2** is completely determined by its range over *atomic propositions* (see Remark 7.21(3)), i.e. over orthogonal projectors onto subspaces of dimension 1 in H. The proof follows directly property (2) in **A2**, to which ρ is subjected. ■

Important remark. When we assign a state there will be propositions with probability 1 of being true if the system undergoes a measurement, and propositions with probability less than 1 if the system is tested on them. We may view the first class as properties the system really possesses in the state considered.

Under the standard interpretation of QM, where probability *has no epistemic meaning*, we are forced to conclude that the properties relative to the second class of propositions *are not defined* for the state examined.

An important example for physics is this. Take propositions P_E, corresponding to properties of a system formed by a quantum particle on the real line, of the form: "the particle's position is in the Borel set $E \in \mathscr{B}(\mathbb{R})$". If the state ρ assigns to each P_E, E bounded, probability less than 1 (it is not hard to come up with such a state, as we shall see when dealing with Heisenberg's uncertainty principle/theorem) then we must conclude that *the particle's position, in state ρ, is not defined.*

From this point of view, the spatial description of particles as points in a manifold – here $\mathbb{R}$, representing the "physical space at rest" of a frame system – does not play a central role anymore, unlike for classical physics. In some sense all the properties of a system (which may vary with the state) are on a par, and the "space" in which system and states are described is a Hilbert space. ■

From the mathematical perspective the first question to raise is whether maps ρ as in **A2** exist at all.

Given a Hilbert space H we show they do exist. The proof clearly works, trivially, also when H is finite dimensional. Recall $\mathfrak{B}_1(\mathsf{H})$ denoted *trace-class operators* on H (Chapter 4).

Proposition 7.23. *Let* H *be a separable Hilbert space and* $T \in \mathfrak{B}_1(\mathsf{H})$ *a positive (hence self-adjoint) operator with trace* 1. *Define* $\rho_T : \mathfrak{P}(\mathsf{H}) \to \mathbb{R}$ *by* $\rho_T(P) := tr(TP)$ *for any* $P \in \mathfrak{P}(\mathsf{H})$. *Then:*

(a) $\rho_T(P) \in [0,1]$ *for any* $P \in \mathfrak{P}(\mathsf{H})$.
(b) $\rho_T(I) = 1$.
(c) *if* $\{P_i\}_{i\in\mathbb{N}} \subset \mathfrak{P}(\mathsf{H})$ *satisfies* $P_iP_j = 0$, $i \neq j$, *then*

$$\rho_T\left(s\text{-}\sum_{i=0}^{+\infty} P_i\right) = \sum_{i=1}^{+\infty} \rho_T(P_i)\,.$$

Proof. TP is of trace class for any $P \in \mathfrak{P}(\mathsf{H})$ by Theorem 4.32(b), for P is bounded, hence we can compute $tr(TP)$. T's positivity ensures the eigenvalues of T are non-negative ((c) in Proposition 3.54). We claim they all belong to $[0,1]$. T is compact and self-adjoint (as positive). Using the decomposition of Theorem 4.21, since $|A| = A$ ($A \geq 0$) and so in $A = U|A|$ we have $U = I$,

$$T = \sum_{\lambda\in\sigma_p(A)} \sum_{i=1}^{m_\lambda} \lambda\,(u_{\lambda,i}|\ )\,u_{\lambda,i}\,.$$

The set $\sigma_p(A)$ consists of the eigenvalues of A, and if $\lambda > 0$, $\{u_{\lambda,i}\}_{i=1,\ldots,m_\lambda}$ is a basis for the eigenspace relative to $\lambda \in \sigma_p(A)$. The convergence is in the uniform topology. Let us write the above expansion as

$$T = \sum_j \lambda_j (u_j|\ )u_j\,. \tag{7.13}$$

Above we labelled over $\mathbb{N}$ (or a finite subset thereof, if $dim(\mathsf{H}) < +\infty$) the set of eigenvectors $u_j = u_{\lambda,i}$, $\lambda > 0$, where λ_j is the eigenvalue of u_j; moreover, the set of eigenvectors was completed to a basis of H by adding a basis for the kernel of T (the overall basis is at most countable because H is separable).

Computing the trace of T with respcet to the u_j gives

$$1 = tr(T) = \sum_j \lambda_j\,,$$

so $\lambda_j \in [0,1]$. Note that the above equation proves part (b) as well, for $TI = I$. Take now $P \in \mathfrak{P}(\mathsf{H})$ and compute the trace of TP in said basis:

$$tr(TP) = \sum_j \lambda_j (u_j|Pu_j)\,.$$

As $(u_j|Pu_j) = (Pu_j|Pu_j)$, we have $0 \le (u_j|Pu_j) \le ||P||^2||u_j||^2 \le 1$, where we used $||u_j|| = 1$ and $||P|| \le 1$ (Proposition 3.58(e)). Therefore

$$0 \le \sum_j \lambda_j (u_j|Pu_j) \le \sum_j \lambda_j = 1$$

and (a) holds.

Let us prove (c). Choose a basis $\{u_{i,j}\}_{j\in I_i}$ in each $P_i(\mathsf{H})$. We leave the reader to prove that $B := \{u_{i,j}\}_{j\in I_i, i\in\mathbb{N}}$ is a basis for the closed projection space of $P = s\text{-}\sum_{i=0}^{+\infty} P_i$. Now we complete B to a basis of H by adding B' basis of $P(\mathsf{H})^\perp$. Since T is continuous, by Proposition 3.58(d):

$$\rho_T(P) = tr\left[T\left(s\text{-}\sum_{u\in B}(u|\ )u\right)\right] = tr\left(s\text{-}\sum_{u\in B}(u|\ )Tu\right).$$

Now we compute the trace using $B\cup B'$:

$$\rho_T(P) = \sum_{v\in B\cup B'}\left(v\left|\sum_{u\in B}(u|v)Tu\right.\right) = \sum_{u\in B}(u|Tu)\,, \tag{7.14}$$

where we used that $(v|u) = \delta_{uv}$ for $u, v \in B\cup B'$, and B is orthogonal to B'. Analogously we can prove

$$\sum_{i=1}^{+\infty}\rho_T(P_i) = \sum_{i=1}^{+\infty}\sum_{j=1}^{I_i}(u_{i,j}|Tu_{i,j})\,.$$

We can always enlarge any set I_i to $\mathbb{N}$ by setting $u_{i,j} := 0$ if $j > \sup I_i$. Remembering Proposition 3.21(c), noting $(u_{i,j}|Tu_{i,j}) \ge 0$ for any i, j, hence also

$$\int_{\mathbb{N}}(u_{i,j}|Tu_{i,j})d\mu(j) \ge 0 \quad \text{for any } i\in\mathbb{N},$$

where μ is the counting measure on $\mathbb{N}$, we eventually get

$$\sum_{i=1}^{+\infty}\rho_T(P_i) = \sum_{i=1}^{+\infty}\sum_{j=1}^{\infty}(u_{i,j}|Tu_{i,j}) = \int_{\mathbb{N}}d\mu(i)\int_{\mathbb{N}}(u_{i,j}|Tu_{i,j})d\mu(j)\,.$$

With the same interpretation the last sum in (7.14) reads

$$\rho_T(P) = \sum_{u\in B}(u|Tu) = \int_{\mathbb{N}\times\mathbb{N}}\left(u_{i,j}\left|Tu_{i,j}\right.\right)d\mu(i)\otimes d\mu(j)\,.$$

As $(u_{i,j}|Tu_{i,j})$ are non-negative numbers and the integral converges, the theorem of Fubini–Tonelli guarantees

$$\sum_{i=1}^{+\infty}\rho_T(P_i) = \int_{\mathbb{N}}d\mu(i)\int_{\mathbb{N}}(u_{i,j}|Tu_{i,j})d\mu(j) = \int_{\mathbb{N}\times\mathbb{N}}\left(u_{i,j}\left|Tu_{i,j}\right.\right)d\mu(i)\otimes d\mu(j) = \rho_T(P)\,,$$

i.e. statement (c). □

The next result, due to Gleason [Gle57], is paramount, in that it provides a complete characterisation of the functions that satisfy axiom **A2**.

Theorem 7.24 (Gleason). *Let* H *be a Hilbert space of finite dimension* ≥ 3*, or infinite-dimensional and separable. For any map* $\mu : \mathfrak{P}(\mathsf{H}) \to [0,+\infty)$ *with* $\mu(I) < +\infty$ *satisfying statement (2) in* **A2***, there exists a positive operator* $T \in \mathfrak{B}_1(\mathsf{H})$ *such that*

$$\mu(P) = tr(TP) \quad \textit{for any } P \in \mathfrak{P}(\mathsf{H}).$$

Sketch of proof. Take a Hilbert space H, either separable and infinite-dimensional, or just finite-dimensional. Define a non-negative *frame function* on H to be a function $f : \mathbb{S}_{\mathsf{H}} \to [0,+\infty)$, $\mathbb{S}_{\mathsf{H}} := \{x \in \mathsf{H} \mid ||x|| = 1\}$, for which there exists $W \in [0,+\infty)$ such that

$$\sum_{i\in K} f(x_i) = W$$

for any basis $\{x_i\}_{i\in K} \subset \mathsf{H}$. A lengthy argument relying on results of von Neumann (cf. Gleason, *op. cit.*) proves the following lemma.

Lemma 7.25. *On any Hilbert space, either separable or of finite dimension* ≥ 3*, for any non-negative frame function* f *there exists a bounded, self-adjoint operator* T *such that* $f(x) = (x|Tx)$*, for every* $x \in \mathbb{S}_{\mathsf{H}}$.

Consider the projectors $P_x := (x|\ )x$, $x \in \mathbb{S}_{\mathsf{H}}$. With the assumption made on μ it is straightforward that $f(x) := \mu(P_x)$ is a non-negative frame function, since $\mu \geq 0$ and

$$\sum_{i\in K} f(x_i) = \sum_{i\in K} \mu(P_{x_i}) = \mu\left(\sum_{i\in K} P_{x_i}\right) = \mu\,(I) < +\infty\,.$$

By the lemma there is a self-adjoint operator T such that $\mu(P_x) = (x|Tx)$ for any $x \in \mathbb{S}_{\mathsf{H}}$. Since $(x|Tx) \geq 0$ for any x, T is positive, so $T = |T|$ (in fact: $|T|^2 = T^*T$ by polar decomposition, but now $T^*T = T^2$ because positive roots are unique (Theorem 3.66)). If $\{x_i\}_{i\in K} \subset \mathsf{H}$ is a basis,

$$+\infty > \mu(I) = \sum_{i\in K} f(x_i) = \sum_{i\in K} (x_i|Tx_i) = \sum_{i\in K} (x_i|\,|T|x_i)\,.$$

By Definition 4.30 $T = |T|$ is then of trace class. Take now $P \in \mathfrak{P}(\mathsf{H})$ and pick a basis $\{x_i\}_{i\in J}$ of $P(\mathsf{H})$, complete it to H by adding a basis $\{x_i\}_{i\in J'}$ of $P(\mathsf{H})^\perp$. Then J is countable (or finite) by Theorem 3.30, plus:

$$P = s\text{-}\sum_{i\in J} P_{x_i}$$

by Proposition 3.58(d). Eventually,

$$P_{x_i} P_{x_j} = 0$$

if $i \neq j$ are in J. Since $Px_i = x_i$ if $i \in J$, and $Px_i = 0$ if $i \in J'$, we have

$$\mu(P) = \sum_{i\in J} \mu(P_{x_i}) = \sum_{i\in J} (x_i|Tx_i) = \sum_{i\in J\cup J'} (x_i|TPx_i) = tr(TP)\,.$$

The proof's sketch ends here. □

Remark 7.26. **(1)** Gleason's proof works for real Hilbert spaces too.
(2) The operator T has trace 1 if $\mu(I) = 1$, as in the case of **A2**.
(3) If the Hilbert space is complex, as in **A2** and always in this text, the operator T associated to μ is unique: any other T' of trace class such that $\mu(P) = tr(T'P)$ for any $P \in \mathfrak{P}(\mathsf{H})$ must also satisfy $(x|(T - T')x) = 0$ for any $x \in \mathsf{H}$. If $x = 0$ this is clear, while if $x \neq 0$ we may complete the vector $x/||x||$ to a basis, in which $tr((T - T')P_x) = 0$ reads $||x||^{-2}(x|(T - T')x) = 0$, where P_x is the projector onto $< x >$. By Exercise 3.18 we obtain $T - T' = 0$.
(4) Imposing $\dim \mathsf{H} > 2$ is mandatory, as the next example shows. On $\mathbb{C}^2$ the orthogonal projectors are $0, I$ and any matrix of the form

$$P_n := \frac{1}{2}\left(I + \sum_{i=2}^{3} n_i \sigma_i\right) , \quad \text{with } n = (n_1, n_2, n_3) \in \mathbb{R}^3 \text{ such that } |n| = 1,$$

where $\sigma_1, \sigma_2, \sigma_3$ are the *Pauli matrices*:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} , \ \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} , \ \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} . \tag{7.15}$$

There is a one-to-one correspondence between projectors P_n and points $n \in \mathbb{S}^2$ on the unit two-sphere. The functions μ of Gleason's theorem can be thought of as maps on $\mathbb{S}^2 \cup \{0, I\}$. Gleason's assumptions boil down to $\mu(0) = 0$, $\mu(I) = 1$ and $\mu(n) = 1 - \mu(-n)$. Positive trace-class operators with unit trace are precisely those of the form:

$$\rho_u = \frac{1}{2}\left(I + \sum_{i=2}^{3} u_i \sigma_i\right) \quad \text{with } u \in \mathbb{R}^3 \text{ such that } |u| \leq 1 . \tag{7.16}$$

If $\cdot$ is the standard dot product on $\mathbb{R}^3$, a direct computation using Pauli's matrices gives

$$tr(\rho_u P_n) = \frac{1}{2}(1 + u \cdot n) .$$

The function μ defined by $\mu(0) = 0$, $\mu(I) = 1$ and

$$\mu(P_n) = \frac{1}{2}\left(1 + (v \cdot n)^3\right) ,$$

for any $n \in \mathbb{S}^2$ and a fixed $v \in \mathbb{S}^2$, satisfies Gleason's theorem. It is easy to prove, however, there are no operators ρ_u like (7.16) such that $\mu(P_n) := tr(\rho_u P_n)$ for any projector P_u; that is to say, there are no $u \in \mathbb{R}^3$, $|u| \leq 1$, such that

$$(1 + u \cdot n) = \left(1 + (v \cdot n)^3\right) \quad \text{for any } n \in \mathbb{S}^2. \qquad \blacksquare$$

Gleason's theorem and the previous considerations, plus the fact that every quantum system known has a Hilbert space satisfying Gleason's assumptions[10], lead to a reformulation of axiom **A2**.

[10] Particles with spin $1/2$ admit a Hilbert space – in which the observable spin is defined – of dimension 2. The same occurs to the Hilbert space in which the polarisation of light is described (cf. helicity of photons). When these systems are described in full, however, for instance including freedom degrees relative to position or momentum, they are representable on a separable Hilbert space of infinite dimension.

A2. *A* **state** *ρ at time t, on a quantum system S with associated Hilbert space H_S, is a positive trace-class operator with unit trace on H_S. The probability that the proposition $P \in \mathfrak{P}(\mathsf{H}_S)$ is true on state ρ equals $tr(\rho P)$.*

In conclusion, and more generally, we can say the following.

Definition 7.27. *Let H be a Hilbert space (not necessarily separable nor finite-dimensional). A positive trace-class operator with trace 1 is called a* **state** *on H. The set of states on H is denoted by $\mathfrak{S}(\mathsf{H})$.*

7.4.2 The Kochen–Specker theorem

Gleason's theorem has a momentous consequence in physics, which distinguishes the states of classical systems from quantum ones. Classical systems admit *completely deterministic* states, described by what we have called sharp states: Dirac measures with support at a point in phase space at the time considered. Each such measure maps sets either to 0 or to 1. These are states on which every statement is either true or false, and there is no intermediate option. *States of this kind do not occur in quantum system because of the following important fact*[11].

Theorem 7.28 (Kochen–Specker). *If H is a Hilbert space, separable or of finite dimension ≥ 3, there is no function $\rho : \mathfrak{P}(\mathsf{H}) \to [0,1]$ fulfilling (1) and (2) in axiom* **A2** *(preliminary form) and taking values in $\{0,1\}$.*

Proof. If x belongs to $\mathbb{S}_\mathsf{H}$ (unit length) and P_x is the orthogonal projector $(x|\cdot)x$, any such ρ gives by Gleason's theorem (the dimension is ≥ 3) a map $\mathbb{S}_\mathsf{H} \ni x \mapsto \rho(P_x) = (x|Tx)$, where $T \in \mathfrak{B}_1(\mathsf{H})$ with $T \geq 0, trT = 1$ is determined uniquely by ρ. This map is patently continuous for the topology of $\mathbb{S}_\mathsf{H}$ induced by the ambient H. We claim $\mathbb{S}_\mathsf{H}$ is path-connected, i.e., for any $x, y \in \mathbb{S}_\mathsf{H}$ there is a continuous path $\gamma : [a,b] \to \mathbb{S}_\mathsf{H}$ starting at $\gamma(a) = x$ and ending at $\gamma(b) = y$. If so, since $\mathbb{S}_\mathsf{H} \ni x \mapsto \rho(P_x) = (x|Tx)$ is continuous, its image is clearly path-connected (as composite of paths in $\mathbb{S}_\mathsf{H}$ with ρ itself). As this image belongs in $\{0,1\}$, the possibilities are that it is $\{0,1\}$, or $\{0\}$, or $\{1\}$. But there is no path joining 0 and 1 *contained in* $\{0,1\}$, so necessarily $\rho(P_x) = 0$ for any $x \in \mathbb{S}_\mathsf{H}$, or $\rho(P_x) = 1$ for any $x \in \mathbb{S}_\mathsf{H}$. In the former case $(x|Tx) = 0$ for any x, hence $tr(T) = 0$, violating $tr(T) = 1$. In the latter case $(x|Tx) = 1$ for any x, again contradi cting $tr(T) = 1$ by dimensional reasons.

To conclude we must show $\mathbb{S}_\mathsf{H}$ is indeed path connected. Taking $x, y \in \mathbb{S}_\mathsf{H}$ we have two options. The first is that $x = e^{i\alpha_0}y$ for some $\alpha_0 > 0$, so x is joined to y by the curve $[0, \alpha_0] \ni \alpha \mapsto e^{i\alpha}x$. Note the curve is continuous for the Hilbert topology and totally contained in S. The second option is that x is a linear combination of y and some $y' \in \mathbb{S}_\mathsf{H}$ orthogonal to y, obtained from completing y to an orthonormal basis for the span of y, x. Since $||x|| = ||y|| = ||y'|| = 1$ and $y \perp y'$, then $x = e^{i\alpha}(\cos\beta)y + e^{i\delta}(\sin\beta)y'$ for three reals α, β, δ. But then x is joined to y by the

[11] Kochen S., Specker E.P.: The problem of Hidden Variables in Quantum Mechanics. J. Math. Mech. **17**(1), 59–87 (1967).

continuous curve, all contained in $\mathbb{S}_{\mathsf{H}}$, defined by varying each of the three parameters on suitable adjacent intervals. □

This no-go result is relevant when one tries to construct classical models of QM by introducing "hidden variables" of classical type, essentially, because these severely restrict the models. For a general discussion on the use of hidden variables and the obstruction due to the lack of dispersion-free states, i.e. sharp states (also in more general contexts than the formulation of QM in Hilbert spaces), we recommend [Jau73, ch. 7] and [BeCa81, Chapter 25].

7.4.3 Pure states, mixed states, transition amplitudes

Let us now study the set $\mathfrak{S}(\mathsf{H}_S)$ of states if H_S is the Hilbert space associated to the quantum system S. A few reminders are useful.

Given a vector space X, a finite linear combination $\sum_{i\in F}\alpha_i x_i$ is called **convex** if $\alpha_i \in [0,1]$, $i\in F$, and $\sum_{i\in F}\alpha_i = 1$.

Moreover (Definition 2.61) $C\subset \mathsf{X}$ is called **convex** if for any pair $x,y\in C$, $\lambda x+(1-\lambda)y\in C$ for all $\lambda\in[0,1]$ (and thus every convex combination of elements in C belongs to C).

If C is convex, $e\in C$ is called **extreme** if it cannot be written as $e=\lambda x+(1-\lambda)y$, with $\lambda\in(0,1)$, $x,y\in C\setminus\{e\}$.

Definition 7.29. *Let X be a vector space over $\mathbb{K}=\mathbb{C}$ or $\mathbb{R}$ and consider the equivalence relation:*

$$u\sim v \;\Leftrightarrow\; v=\alpha u \text{ for some } \alpha\in\mathbb{K}\setminus\{0\}.$$

The quotient space $\mathsf{X}/\!\sim$ is the **projective space** *over X. We call elements of $\mathsf{X}/\!\sim$ other than $[0]$ (equivalence class of the null vector)* **rays** *of X.*

Proposition 7.30. *Let $(\mathsf{H},(\;|\;))$ be a separable Hilbert space.*

(a) *$\mathfrak{S}(\mathsf{H})$ is a convex subset in $\mathfrak{B}_1(\mathsf{H})$.*
(b) *Extreme points in $\mathfrak{S}(\mathsf{H})$ are those of the form:*

$$\rho_\psi := (\psi|\;)\psi\,, \quad \text{for every vector } \psi\in\mathsf{H} \text{ with } ||\psi||=1.$$

This sets up a bijection between extreme states and rays of H, which maps the extreme state $(\psi|\;)\psi$ to the ray $[\psi]$.
(c) *Any state $\rho\in\mathfrak{S}(\mathsf{H})$ satisfes*

$$\rho\geq\rho\rho\,,$$

and is extreme if and only if

$$\rho\rho=\rho\,.$$

(d) *Any state $\rho\in\mathfrak{S}(\mathsf{H})$ is a linear combination of extreme states, including infinite combinations in the uniform topology. In particular there is always a decomposition*

$$\rho=\sum_{\phi\in N} p_\phi(\phi|\;)\phi\,,$$

where N is an eigenvector basis for ρ, $p_\phi \in [0,1]$ for any $\phi \in N$, and

$$\sum_{\phi \in N} p_\phi = 1\,.$$

Proof. (a) Take two states ρ, ρ'. It is clear $\lambda\rho + (1-\lambda)\rho'$is of trace class because trace-class operators form a subspace in $\mathfrak{B}(\mathsf{H})$ (Theorem 4.32). By the trace's linearity (Proposition 4.34):

$$tr[\lambda\rho + (1-\lambda)\rho'] = \lambda tr\rho + (1-\lambda)tr\rho' = \lambda 1 + (1-\lambda)1 = 1\,.$$

At last, if $f \in \mathsf{H}$ and $\lambda \in [0,1]$, since ρ and ρ' are positive:

$$(f|(\lambda\rho + (1-\lambda)\rho')f) = \lambda(f|\rho f) + (1-\lambda)(f|\rho' f) \geq 0\,.$$

Hence $\lambda\rho + (1-\lambda)\rho'$ is a state if ρ, ρ' are states and $\lambda \in [0,1]$.
(b) and (d) Consider $\rho \in \mathfrak{S}(\mathsf{H})$. ρ is a compact and self-adjoint operator (as positive). Using the decomposition of Theorem 4.21, and since $|\rho| = \rho$ ($\rho \geq 0$), so $U = I$ in the polar decomposition of $\rho = U|\rho|$, we find:

$$\rho = \sum_{\lambda \in \sigma_p(\rho)} \sum_{i=1}^{m_\lambda} \lambda\,(u_{\lambda,i}|\ )\,u_{\lambda,i}\,. \tag{7.17}$$

Above, $\sigma_p(\rho)$ is the set of eigenvectors of ρ, and if $\lambda > 0$, $\{u_{\lambda,i}\}_{i=1,\ldots,m_\lambda}$ is a basis of the eigenspace of $\lambda \in \sigma_p(\rho)$. At last, convergence is understood in uniform topology. This expansion alone proves (d).

Completing $\cup_{\lambda>0}\{u_{\lambda,i}\}_{i=1,\ldots,m_\lambda}$ by adding a basis for $Ker\rho$, by Proposition 4.34 we obtain:

$$1 = tr(\rho) = \sum_{\lambda \in \sigma_p(\rho)} m_\lambda \lambda\,. \tag{7.18}$$

Suppose now $\rho_\psi := (\psi|\ )\psi$, $||\psi|| = 1$. Immediately, $\rho_\psi \in \mathfrak{S}(\mathsf{H})$. We want to prove ρ_ψ is extreme in $\mathfrak{S}(\mathsf{H})$. So assume there are $\rho, \rho' \in \mathfrak{S}(\mathsf{H})$ and $\lambda \in (0,1)$ such that

$$\rho_\psi = \lambda\rho + (1-\lambda)\rho'\,.$$

We will show $\rho = \rho' = \rho_\psi$.

Consider the orthogonal projector $P_\psi = (\psi|\ )\psi$. It is clear (completing ψ to a basis) that $tr(\rho_\psi P_\psi) = 1$, so

$$1 = \lambda tr(\rho P_\psi) + (1-\lambda)tr(\rho' P_\psi)\,.$$

As $\lambda \in (0,1)$ and $0 \leq tr(\rho P_\psi) \leq 1$, we have $0 \leq tr(\rho' P_\psi) \leq 1$, possible only if $tr(\rho P_\psi) = tr(\rho' P_\psi) = 1$. So let us prove that $tr(\rho P_\psi) = 1$ and $tr(\rho' P_\psi) = 1$ imply $\rho = \rho' = \rho_\psi$.

Decomposing ρ as in (7.17), $tr(\rho P_\psi) = 1$ becomes

$$\sum_j \lambda_j |(u_j|\psi)|^2 = 1\,, \tag{7.19}$$

where we labelled with $\mathbb{N}$ (or a finite subset in case $dim(\mathsf{H}) < +\infty$) the eigenvectors $u_j = u_{\lambda,i}$, $\lambda > 0$, we called λ_j the eigenvalue of u_j and we completed to a basis of H the previous eigenvectors by adding a basis for the null space of ρ (the total basis is at most countable because H is separable). By assumption

$$\sum_j \lambda_j = 1\,, \tag{7.20}$$

$$\sum_j |(u_j|\psi)|^2 = 1\,. \tag{7.21}$$

Since $\lambda_j \in [0,1]$ and $|(u_j|\psi)|^2 \in [0,1]$ for any $j \in \mathbb{N}$, we obtain

$$\sum_j \lambda_j^2 \leq 1\,, \tag{7.22}$$

$$\sum_j |(u_j|\psi)|^4 \leq 1\,. \tag{7.23}$$

Thus the sequences of the λ_j and the $|(u_j|\psi)|^2$ belong to $\ell^2(\mathbb{N})$. Identity (7.19), plus (7.22), (7.23) and the Cauchy-Schwarz inequality in $\ell^2(\mathbb{N})$, give

$$\sum_j \lambda_j^2 = 1\,, \tag{7.24}$$

$$\sum_j |(u_j|\psi)|^4 = 1\,. \tag{7.25}$$

Since $\lambda_j \in [0,1]$ for any $j \in \mathbb{N}$, (7.20) and (7.24) are consistent only if all λ_i vanish except one, say λ_p, which equals 1. Likewise, since $|(u_j|\psi)|^2 \in [0,1]$ for any $j \in \mathbb{N}$, (7.21) and (7.25) are consistent only if all $|(u_j|\psi)|$ are zero except for $|(u_k|\psi)| = 1$. As the u_i are a basis and $||\psi|| = 1$, necessarily $\psi = \alpha u_k$, with $|\alpha| = 1$. Clearly, then, $k = p$, for otherwise $tr(\rho P_\psi) = 0$. But

$$\rho = \sum_j \lambda_j (u_j|\ )u_j\,,$$

so eventually

$$\rho = \lambda_k(u_k|\ )u_k = 1 \cdot (u_k|\ )u_k = \overline{\alpha^{-1}}\alpha^{-1}(\psi|\ )\psi = |\alpha|^{-1}(\psi|\ )\psi = (\psi|\ )\psi = \rho_\psi\,.$$

In the same way we can prove $\rho' = \rho_\psi$.

If a state ρ is not of the type $(\psi|\ )\psi$, we can still decompose it as

$$\rho = \sum_j \lambda_j (u_j|\ )u_j\,,$$

where at least two $p \neq q$ (with $\lambda_p \neq \lambda_q$) do not vanish. Hence in particular $\lambda_p, 1-\lambda_p \in (0,1)$. Then we can write ρ as

$$\rho = \lambda_p(u_p|\ )u_p + (1-\lambda_p)\sum_{j\neq p} \frac{\lambda_j}{(1-\lambda_p)}(u_j|\ )u_j\,.$$

$(u_p|\ )u_p$ is a state as already said, and easily, we also have

$$\rho' := \sum_{j\neq p} \frac{\lambda_j}{(1-\lambda_p)}(u_j|\ )u_j$$

is a state of $\mathfrak{S}(\mathsf{H})$ (obviously $\rho' \neq (u_p|\ )u_p$ by construction, as $u_q \nsim u_p$). So we have proved ρ is not extreme.

The function f mapping the extreme state $(\psi|\ )\psi$ to the ray $[\psi]$ is well defined. In fact, let us first notice that $||\psi|| = 1$ by definition of extreme state, so $\psi \neq 0$ and $[\psi]$ is a ray. One extreme state may be expressed differently: namely (as is immediate to see from $||\phi|| = 1$) $(\psi|\ )\psi = (\phi|\ )\phi$ iff $\psi = e^{i\alpha}\phi$ for some $\alpha \in \mathbb{R}$. But then by definition of ray $[\psi] = [\phi]$. We claim the function f is one-to-one: if ϕ, ψ are unit and $[\psi] = [\phi]$, then $\psi = e^{i\alpha}\phi$ for some $\alpha \in \mathbb{R}$, so $(\psi|\ )\psi = (\phi|\ )\phi$. The function is also onto, because if $[\phi]$ is a ray, $||\phi|| \neq 0$ so there is a $\psi \in [\phi]$ with $||\psi|| = 1$. Then $f((\psi|\ )\psi) = [\phi]$ since $\psi = \alpha\phi$ for some non-zero $\alpha \in \mathbb{C}$.
(c) Begin with the second claim. If ρ is extreme, $\rho\rho = \rho$ using the form in part (b) for extreme points. Decomposing a state ρ as (see the meaning above):

$$\rho = \sum_j \lambda_j (u_j|\ )u_j\,,$$

gives

$$\rho\rho = \sum_j \lambda_j^2 (u_j|\ )u_j\,.$$

If $\rho\rho = \rho$, passing to traces gives

$$\sum_j \lambda_j^2 = \sum_j \lambda_j = 1$$

with $\lambda_j \in [0,1]$. This is possible only if all λ_j are zero but one, $\lambda_k = 1$. Then

$$\rho = \sum_j \lambda_j (u_j|\ )u_j = 1 \cdot (u_k|\ )u_k\,,$$

which is an extreme state by (b).

Now to the first claim. Let $x = \sum_j \alpha_j u_j$ be arbitrary in H (the u_j are a basis of H). Since $\lambda_j \in [0,1]$,

$$(x|\rho\rho x) = \sum_j \lambda_j^2 (x|u_j)(u_j|x) = \sum_j \lambda_j^2 |\alpha_j|^2 \leq \sum_j \lambda_j |\alpha_j|^2$$
$$= \sum_j \lambda_j (u_j|x)(u_j|x) = (x|\rho x)\,.$$

Therefore $\rho\rho \leq \rho$. □

Now we have a definition.

Definition 7.31. *Let* $(\mathsf{H}, (\,|\,))$ *be a separable Hilbert space.*

(a) *Extreme elements in* $\mathfrak{S}(\mathsf{H})$ *are called* **pure states***, and their set is denoted* $\mathfrak{S}_p(\mathsf{H})$*; non-extreme states are* **mixed states***,* **mixtures** *or* **nonpure states***.*
(b) *If:*

$$\psi = \sum_{i \in I} \alpha_i \phi_i \,,$$

with I finite or countable (and the series converges in the topology of H *in the second case), where the vectors* $\phi_i \in \mathsf{H}$ *are all non-null and* $0 \neq \alpha_i \in \mathbb{C}$*, one says the state* $(\psi|\;)\psi$ *is a* **coherent superposition** *of the states* $(\phi_i|\;)\phi_i/||\phi_i||^2$*.*
(c) *If* $\rho \in \mathfrak{S}(\mathsf{H})$ *satisfies:*

$$\rho = \sum_{i \in I} p_i \rho_i$$

with I finite, $\rho_i \in \mathfrak{S}(\mathsf{H})$*,* $0 \neq p_i \in [0,1]$ *for any* $i \in I$*, and* $\sum_i p_i = 1$*, the state* ρ *is said* **incoherent superposition** *of states* ρ_i *(possibly pure).*
(d) *If* $\psi, \phi \in \mathsf{H}$ *satisfy* $||\psi|| = ||\phi|| = 1$*:*

- (i) *the complex number* $(\psi|\phi)$ *is the* **transition amplitude** *or* **probability amplitude** *of the state* $(\phi|\;)\phi$ *on the state* $(\psi|\;)\psi$*;*
- (ii) *the non-negative real number* $|(\psi|\phi)|^2$ *is the* **transition probability** *of the state* $(\phi|\;)\phi$ *on the state* $(\psi|\;)\psi$*.*

Remark 7.32. (1) The vectors of the Hilbert space of a quantum system associated to pure states are often said, in physics' literature, **wavefunctions**. The reason for such a name is due to the first formulation of Quantum Mechanics in terms of Wave Mechanics (see Chapter 6).
(2) The possibility of creating pure states by non-trivial combinations of vectors associated to other pure states is called, in the jargon of QM, **superposition principle of (pure) states**.
(3) In (c), in case $\rho_i = \psi_i(\psi_i|\;)$, we do *not* require $(\psi_i|\psi_j) = 0$ if $i \neq j$. However it is immediate to see that if I is finite, if ρ_i is a mixed or pure state and if $p_i \in [0,1]$ for any $i \in I$, $\sum_i p_i = 1$, then:

$$\rho = \sum_{i \in I} p_i \rho_i$$

is of trace class (obvious: trace-class operators are a vector space and every ρ_i is of trace class), positive (as positive linear combination of positive operators), and it has unit trace: this because by the trace's linearity (Proposition 4.34), we have

$$tr\rho = tr\left(\sum_{i \in I} p_i \rho_i\right) = \sum_{i \in I} p_i tr \rho_i = \sum_{i \in I} p_i \cdot 1 = 1 \,.$$

The decomposition of ρ over an eigenvector basis can be considered a limiting case of the above: when I is countable, in fact, $\rho_i = \psi_i(\psi_i|\;)$ and $(\psi_i|\psi_j) = \delta_{ij}$.

It is important to remark that *in general, a given mixed state admits several incoherent decompositions by pure and nonpure states.*

(4) Consider the pure state $\rho_\psi \in \mathfrak{S}_p(\mathsf{H})$, written $\rho_\psi = (\psi|\ )\psi$ for some $\psi \in \mathsf{H}$ with $||\psi|| = 1$. What we want to emphasise is that this pure state is also an orthogonal projector $P_\psi := (\psi|\ )\psi$, so it must correspond to a proposition about the system.

The naïve and natural interpretation[12] *of the statement is this: "the system's state is the pure state given by the vector* ψ*".*

This interpretation is due, if $\rho \in \mathfrak{S}(\mathsf{H})$, to the fact that $tr(\rho P_\psi) = 1$ if and only if $\rho = (\psi|\ )\psi$. In fact, if $\rho = (\psi|\ )\psi$, by completing ψ to a basis and taking the trace, we have $tr(\rho P_\psi) = 1$. Conversely, suppose $tr(\rho P_\psi) = 1$ for the state ρ. Then $\rho = (\psi|\ )\psi$ from the proof of Proposition 7.30.

(5) Part (4) allows to interpret the square modulus of the transition amplitude $(\phi|\psi)$. If $||\phi|| = ||\psi|| = 1$, as the definition of transition amplitude imposes, $tr(\rho_\psi P_\phi) = |(\phi|\psi)|^2$, where $\rho_\psi := (\psi|\ )\psi$ and $P_\phi = (\phi|\ )\phi$. Using (4) we conclude:
$|(\phi|\psi)|^2$ *is the probability that the state, given (at time t) by the vector* ψ*, following a measurement (at time t) on the system becomes determined by* ϕ*.*

Notice $|(\phi|\psi)|^2 = |(\psi|\phi)|^2$, so the probability transition of the state determined by ψ on the state determined by ϕ concides with the analogous probability where the vectors are swapped. This fact is, *a priori*, highly non-evident in physics. ■

7.4.4 Axiom A3*: post-measurement states and preparation of states*

The standard formulation of QM assumes the following axiom (introduced by von Neumann and generalised by Lüders) about what occurs to the physical system S, in state $\rho \in \mathfrak{S}(\mathsf{H}_S)$ at time t, when subjected to the measurement of proposition $P \in \mathfrak{P}(\mathsf{H}_S)$, if the latter is true (so in particular $tr(\rho P) > 0$, prior to the measurement).

A3. *If the quantum system S is in state* $\rho \in \mathfrak{S}(\mathsf{H}_S)$ *at time t and proposition* $P \in \mathfrak{P}(\mathsf{H}_S)$ *is true after a measurement at time t, the system's state immediately afterwards is:*

$$\rho_P := \frac{P\rho P}{tr(\rho P)} .$$

In particular, if ρ *is pure and determined by the unit vector* ψ*, the state immediately after measurement is still pure, and determined by:*

$$\psi_P = \frac{P\psi}{||P\psi||} .$$

Obviously, in either case ρ_P and ψ_P define states. In the former, in fact, ρ_P is positive of trace class, with unit trace, while in the latter $||\psi_P|| = 1$.

Remark 7.33. (1) As already highlighted, measuring a property of a physical quantity goes through the interaction between the system and an instrument (supposed macroscopic and obeying the laws of classical physics). Quantum Mechanics, in its

[12] We cannot but notice how this interpretation muddles the semantic and syntactic levels. Although this could be problematic in a formulation within formal logic, the use physicists make of the interpretation eschews the issue.

standard formulation, does not establish what a measuring instrument is, it only says they exist; nor is it capable of describing the interaction of instrument and quantum system beyond the framework set in **A3**. Several viewpoints and conjectures exist on how to complete the physical description of the measuring process; these are called, in the slang of QM, **collapse**, or **reduction**, **of the state** or **of the wavefunction**. For various reasons, though, none of the current proposals is entirely satisfactory [Des99, Bon97, Ghi97, Alb94]. A very interesting proposal was put forward in 1985 by G.C.Girardi, T. Rimini and A. Weber (*Physical Review* D34, 1985 p. 470), who described in a dynamically nonlinear way the measuring process and assumed it is due to a self-localisation process, *rather than* to an instrument. This idea, alas, still has several weak points, in particular it does not allow – at least in an obvious manner – for a relativistic description.
(2) Axiom **A3** refers to *non-destructive* testing, also known as *indirect measurement* or *first-kind measurement* [BrKh95], where the physical system examined (typically a particle) is not absorbed/annihilated by the instrument. They are idealised versions of the actual processes used in labs, and only in part they can be modelled in such a way.
(3) Measuring instruments are commonly employed to *prepare a system in a certain state*. Theoretically-speaking the preparation of a *pure* state is carried out like this. A finite collection of *compatible* propositions $P_1,\ldots,P_n$ is chosen so that the projection subspace of $P_1 \wedge \cdots \wedge P_n = P_1 \cdots P_n$ is *one-dimensional*. In other words $P_1 \cdots P_n = (\psi|\ )\psi$ for some vector with $||\psi|| = 1$. The existence of such propositions is seen in practically all quantum systems used in experiments. (From a theoretical point of view these are *atomic* propositions in the sense of Remark 7.21(3), and must exist because of the Hilbert space.) Then propositions P_i are simultaneously measured on several identical copies of the physical system of concern (e.g., electrons), whose initial states, though, are unknown. If for one system the measurements of all propositions are successful, the post-measurement state is determined by the vector ψ, and the system was **prepared** in that particular state.

Normally each projector P_i is associated to a measurable quantity A_i on the system, and P_i defines the proposition "the quantity A_i belongs to the set E_i". In practice, thus, to prepare a sysytem (available in arbitrarily many identical copies) in the pure state ψ one measures simultaneously a number of *compatible* quantities A_i and selects the systems for which the readings of the A_i belong to the given sets E_i.
(4) Let us explain how to obtain nonpure states from pure ones. Consider q_1 identical copies of system S prepared in the pure state associated to ψ_1, q_2 copies of S prepared in the pure state associated to ψ_2 and so on, up to ψ_n. If we mix these states each one will be in the nonpure state:

$$\rho = \sum_{i=1}^{n} p_i(\psi_i|\ )\psi_i\,,$$

where $p_i := q_i / \sum_{i=1}^{n} q_i$. In general, $(\psi_i|\psi_j)$ is not zero if $i \neq j$, so the above expression for ρ is not the decomposition with respect to an eigenvector basis for ρ. This procedure hints at the existence of two different types of probability, one intrinsic and due to the quantum nature of state ψ_i, the other epistemic, and encoded in the probability p_i. But this is not true: once a nonpure state has been created, as above, there is no way, within QM, to distinguish the states forming the mixture. For example, the same ρ could have been obtained mixing other pure states than those determined by the ψ_i. In particular, one could have used those in the decomposition of ρ into a basis of its eigenvectors. For physics, no kind of measurement (under the axioms of QM stated so far) would distinguish the two mixtures. ■

7.4.5 Superselection rules and coherent sectors

We need a general definition of purely mathematical flavour.

Definition 7.34. *If* H *is a Hilbert space and* $\{\mathsf{H}_\alpha\}_{\alpha\in A}$ *a collection of arbitrary cardinality of closed subspaces, we shall write* $\mathsf{H} = \oplus_{\alpha\in A}\mathsf{H}_\alpha$*, and say* H *is the* **Hilbert (direct) sum** *of the* H_α*, if the latter are mutually orthogonal and* $\mathsf{H} = \overline{< \{\mathsf{H}_\alpha\}_{\alpha\in A} >}$.

Remarks 7.35. Relatively to the *orthogonal decomposition* $\mathsf{H} = \oplus_{\alpha\in A}\mathsf{H}_\alpha$ in the sense of Definition 7.34, we leave the reader to prove the following identities, which descend from the observation that the union of bases chosen in each H_α is a basis of H.

For each vector $\psi \in \mathsf{H}$

$$||\psi||^2 = \sum_{\alpha\in A} ||P_\alpha\psi||^2 \tag{7.26}$$

(in the sense of Definition 3.19), where P_α is the projector on H_α for any $\alpha \in A$. Moreover (remembering Lemma 3.25)

$$\psi = \sum_{\alpha\in A} P_\alpha\psi \tag{7.27}$$

where the series may be rearranged arbitrarily. The sum is a series or a finite sum since only a countable, at most, number of $P_\alpha\psi$ does not vanish (Proposition 3.21(b)). ■

For the known quantum systems not all normalised ψ determine states that are physically admissible. There are various theoretical reasons (which we shall return to in the sequel) that force the existence of so-called **superselection rules**. According to these rules the system's Hilbert space H is a Hilbert sum – at most countable if the Hilbert space is assumed separable – of closed subspaces called **coherent sectors**:

$$\mathsf{H} = \mathsf{H}_1 \oplus \mathsf{H}_2 \oplus \cdots .$$

The only physically admissible states, given by single vectors, are represented by vectors in H_1, or H_2, or H_3,, ... *States given by linear combinations over distinct coherent sectors are not admissible.*

Physics views coherent sectors as subspaces of H associated to a collection of mutually-exclusive propositions – i.e. orthogonal projectors P_1, P_2, ... onto the orthogonal coherent sectors, with $\sum_i P_i = I$ (the sum, if infinite, is meant in strong sense). The proposition associated to P_i corresponds to the assertion that the quantity determining the superselection rule has a certain value. More generally, the quantity is not required to take a specific value on each subspace, but only to range over a certain set specified by the proposition. Let us see this with two examples.

(1) As first example we recall the **superselection rule of the electric charge** for charged quantum system. This demands that each vector ψ, determining a system's state, satisfies a proposition P_Q of the type: "the system's charge equals Q" for some value Q. Mathematically, then, $tr(P_Q(\psi|\ )\psi) = 1$ for some value Q, which amounts to say $P_Q\psi = \psi$ for some Q. In other words: *states, determined by single vectors, whose charge is not a definite value are not admissible*. This demand is obvious in classical physics, not in QM, where an electrically charged system, *a priori*, could admit states with indefinite charge. Imposing the Hilbert space is separable requires that the number of values Q of the charge, i.e. the coherent sectors with given charge, is at most countable, so that the electric charge cannot vary with continuity.[13]

(2) Another superselection rule concerns the angular momentum of any physical system. From QM we know the squared modulus J^2 of the angular momentum, when in a definite state, can only take values $j(j+1)$ with j integer or semi-integer (in $\hbar = \frac{h}{2\pi}$ units, where h is the usual Planck constant). Then there is a Hilbert space decomposition of the physical system in two orthogonal closed subspaces, one with integer-valued j, the other with semi-integer j. The **superselection rule of the angular momentum** dictates that vectors representing states are not linear combinations over both subspaces. It is important to remark that a pure state can have an *undefined* angular momentum, since the state/associated vector is a linear combination of pure states/vectors with distinct angular momentum; by superselection, however, these values must be either all integer or all semi-integer. We w ill return to this point in Chapter 12.3.2.

In presence of superselection rules associated to the coherent decomposition of H into:

$$\mathsf{H} = \bigoplus_{k \in K} \mathsf{H}_k\,, \tag{7.28}$$

we can define spaces of states $\mathfrak{S}(\mathsf{H}_k)$ and pure states $\mathfrak{S}_p(\mathsf{H}_k)$ of each sector. These can be identified with subsets in $\mathfrak{S}(\mathsf{H})$ and $\mathfrak{S}_p(\mathsf{H})$ respectively, by the following obvious argument: if M is a closed subspace in the Hilbert space H, $A \in \mathfrak{B}(\mathsf{M})$ is identified with an operator of $\mathfrak{B}(\mathsf{H})$ simply by extending it as the null operator on $\mathsf{M}^\perp$. If A is positive and of trace class, the extension is positive, trace class, and the trace preserved. If A is of the form $(\psi|\ )\psi$, $\psi \in \mathsf{M}$, $||\psi|| = 1$, the extension is alike. In the case considered we identify every $\mathfrak{S}(\mathsf{H}_k)$ and $\mathfrak{S}_p(\mathsf{H}_k)$ with a subset in $\mathfrak{S}(\mathsf{H})$ and $\mathfrak{S}_p(\mathsf{H})$ respectively, extending each state ρ, nonpure or pure, to the zero operator on $\mathsf{H}_k^\perp$. Thereby $\mathfrak{S}(\mathsf{H}_k) \cap \mathfrak{S}(\mathsf{H}_j) = \varnothing$ and $\mathfrak{S}_p(\mathsf{H}_k) \cap \mathfrak{S}_p(\mathsf{H}_j) = \varnothing$ if $k \neq j$. Physically-

[13] If the charge is taken to be continuous and H_q is the subspace where it equals $q \in \mathbb{R}$, i.e. the eigenspace relative to eigenvalue q for a self-adjoint operator Q, the Hilbert space (non-separable) is still a direct sum $\oplus_{q\in\mathbb{R}}\mathsf{H}_q$. $\mathbb{R}$ coincides then with the *point* spectrum of Q. Some authors, instead, prefer to think the Hilbert space as a *direct integral*, thereby preserving its separability, and in this case $bR = \sigma_c(Q)$.

admissible pure states for the physical system described on H are precisely those in:

$$\mathfrak{S}_p(\mathsf{H})_{adm} := \bigsqcup_{k\in K} \mathfrak{S}_p(\mathsf{H}_k)\,. \tag{7.29}$$

It is reasonable to assume that physically-admissible nonpure states for the physical system described on H are then those that can be built as mixtures of admissible pure states. Hence physically-admissible mixed states will be finite convex combinations of elements of

$$\bigsqcup_{k\in K} \mathfrak{S}(\mathsf{H}_k)\,. \tag{7.30}$$

One could allow for infinite convex combinations, for some operator topology, but we will not do this now.

Remark 7.36. (1) Asking for admissible pure states to be elements of (7.29), and mixed states to belong in the convex hull of (7.30), implies immediately that each admissible state ρ (pure or nonpure) satisfies the constraints:

$$\rho P_k = P_k \rho \quad \text{for any } k \in K \tag{7.31}$$

where P_k is the orthogonal projector on the coherent sector H_k. Actually, the converse is true too, provided $Ran(\rho) \subset \oplus_{k\in F}\mathsf{H}_k$, for some $F \subset K$ *finite*, or $\rho \in \mathfrak{S}_p(\mathsf{H})$. For mixed ρ, in fact, since $\sum_{k\in K} P_k = I$ (in the strong topology), and $P_k P_h = 0$ if $h \neq k$, asking (7.31) forces:

$$\rho = s\text{-}\sum_{k\in K} P_k \rho \left(s\text{-}\sum_{h\in K} P_h \right) = s\text{-}\sum_{k\in K} P_k \rho P_k = \sum_{k\in F} P_k \rho P_k = \sum_{k\in F_*} p_k \frac{P_k \rho P_k}{tr(P_k \rho P_k)} = \sum_{k\in F_*} p_k \rho_k\,,$$

where $F_* \subset F$ is the subset of $k \in F$ for which $p_k := tr(P_k \rho P_k) \neq 0$, so $p_k > 0$. Note that by construction $\rho_k := \frac{P_k \rho P_k}{tr(P_k \rho P_k)} \in \mathfrak{S}(\mathsf{H}_k)$ if $k \in F_*$. Beside $p_k \geq 0$,

$$1 = tr\rho = tr\left(\sum_{k\in F} p_k \rho_k \right) = \sum_{k\in F} p_k\, tr\rho_k = \sum_{k\in F} p_k$$

as F is finite (this hypothesis is used only here). Hence if $\rho \in \mathfrak{S}(\mathsf{H})$ satisfies $Ran(\rho) \subset \oplus_{k\in F}\mathsf{H}_k$ for some finite F, and (7.31) hold, then $\rho = \sum_{k\in F_*} p_k \rho_k$ is a convex combination of elements $\rho_k \in \mathfrak{S}(\mathsf{H}_k)$, as we wanted. Let now $\rho = (\psi|\ )\psi$ be pure. By the orthogonal decomposition of H in coherent spaces H_k we have:

$$1 = ||\psi||^2 = \sum_{k\in K} ||P_k \psi||^2\,. \tag{7.32}$$

Equations (7.31) imply $P_k \psi = (\psi|P_k \psi)\psi$, and substituting above gives, using $P_k P_k = P_k$ and $P_k = P_k^*$:

$$1 = ||\psi||^2 = \sum_{k\in K} ||P_k \psi||^2 = \sum_{k\in K} |(\psi|P_k \psi)|^2 1 = \sum_{k\in K} |(P_k \psi|P_k \psi)|^2 = \sum_{k\in K} ||P_k \psi||^4\,. \tag{7.33}$$

But $0 \leq ||P_k\psi||^4 < ||P_k\psi||^2$ if $||P_k\psi|| < 1$, so (7.32) and (7.33) are simultaneously valid only if $||P_k\psi|| = 0$, $k \in F \setminus \{k_0\}$ and $||P_{k_0}\psi|| = 1$. This means $\rho \in \mathfrak{S}_p(H_{k_0})$.
(2) The existence of coherent sectors H_k also implies that *only certain orthogonal projectors have physical meaning as elementary propositions on the physical system; these are projectors Q satisfying $P_kQ = QP_k$ for every k labelling a coherent sector H_k.*

Let us prove this assertion. If $\psi \in \mathsf{H}_k \setminus \{0\}$ is a pure state and we measure the pair of mutually exclusive, compatible elementary observables $Q, Q' := I - Q$, one of them must be true. The post-measurement state, up to renormalisation, is given by vectors $Q\psi$ or $Q'\psi$. The corresponding states must abide by the superselection rule, so $Q\psi \in \mathsf{H}_r$ and $Q'\psi \in \mathsf{H}_s$ for some r, s. If $Q\psi = 0$, we automatically think $Q\psi$ in H_k. But $\mathsf{H}_k \ni \psi = Q\psi + Q'\psi \in \mathsf{H}_r \oplus \mathsf{H}_s$. As $\mathsf{H}_k, \mathsf{H}_r, \mathsf{H}_s$ are pairwise orthogonal if k, r, s are distinct, $k = r$ cannot be if $Q\psi \neq 0$. In conclusion $Q\psi \in \mathsf{H}_k$ if $\psi \in \mathsf{H}_k$ (for $\psi = 0$ this is trivial). If $\phi \in \mathsf{H}$, we can say $QP_k\phi \in \mathsf{H}_k$, and so $P_hQP_k\phi = \delta_{hk}QP_k\phi = P_kQP_h\phi$. Using $I = s\text{-}\sum_h P_h$ and the uniform continuity of Q, then $P_hQP_k\phi = P_kQP_h\phi$ implies $QP_k = P_kQ$. The results generalises straighforward to self-adjoint operators representing observables, as we will see. Abstractly, relating to Remark 7.21(2), the presence of a superselection rule compels to assume that the von Neumann algebra $\mathfrak{R}_S$ of (bounded) observables on the system cannot be a factor, because its centre must contain the projectors P_k. We could envisage that the quantities defining a physical system, those "always true" (e.g., the mass), are actually associated the propositions in the centre of $\mathfrak{R}_S$. They are true simply because we work in a specific sector.
(3) Actually, if one accepts as principle that elementary propositions Q that are admissible under a superselection rule are those that commute with orthogonal projectors P_k on coherent sectors, then the demand that vectors representing states are only those belonging to coherent sectors H_k can be dropped. A linear combination $\Psi = \sum_k \psi_k$ with $\psi_k \in \mathsf{H}_k \setminus \{0\}$ and $||\Psi|| = 1$, theoretically-speaking forbidden, behaves, in relationship to the axioms **A1**, **A2**, as the nonpure state $\rho = \sum_k \psi_k(\psi_k|\cdot)$, as long as we consider only elementary propositions Q compatible with the superselection rule. We leave the proof and the mathematical details to the reader. ■

7.4.6 Algebraic characterisation of a state as a noncommutative Riesz theorem

In this last part of the section we describe a characterisation of the space of states $\mathfrak{S}(\mathsf{H})$ on the complex Hilbert space H, in the simple case where superselection sectors are absent. The characterisation has a certain interest for the *algebraic formulation* of QM [Str05a] and quantum theories in general, which we will briefly see in Chapter 14.1. The mathematical relevance resides in that it implies a *noncommutative* version of Riesz's Theorem 1.58 on (finite) positive Borel measures, where the word *noncommutative* refers to the measures on the lattice of projectors $\mathfrak{P}(\mathsf{H})$ in the sense of axiom **A2** (preliminary form), rather than on a σ-algebra.

First observe that a positive trace-class operator T determines a linear functional of the C^*-algebra of compact operators $\mathfrak{B}_\infty(\mathsf{H})$, given by $\omega_T : \mathfrak{B}_\infty(\mathsf{H}) \to \mathbb{C}$,

$\omega_T(A) = tr(TA)$. This is *positive*:

$$\omega_T(A^*A) \geq 0 \quad \text{for any } A \in \mathfrak{B}_\infty(\mathsf{H}). \tag{7.34}$$

In fact, $tr(A^*TA) = tr(A^*T^{1/2}T^{1/2}A) = tr((T^{1/2}A)^*T^{1/2}A) \geq 0$. The last inequality comes from expanding the trace with respect to any basis of H. Viewing ω_T as linear operator on the Banach space $\mathfrak{B}_\infty(\mathsf{H})$ (with the norm of $\mathfrak{B}(\mathsf{H})$), we have

$$||\omega_T|| = trT\,. \tag{7.35}$$

In fact, if $A \in \mathfrak{B}_\infty(\mathsf{H})$ and $||A|| \leq 1$, taking the trace in a basis $\{\psi_j\}_{j\in J}$ of eigenvectors of $T = T^*$ gives

$$|\omega_T(A)| \leq \sum_{j\in J} |p_j(\psi_j|A\psi_j)| \leq \sum_{j\in J} p_j ||A\psi_j|| \leq \sum_{j\in J} p_j = trT\,.$$

Eventually $\omega_T(A_N) \to trT$, as $N \to +\infty$, if $A_N := \sum_{0\leq p_j<N} \psi_j(\psi_j|\ )$, where $||A_N|| \leq 1$. We also have $A_N \in \mathfrak{B}_\infty(\mathsf{H})$ because the latter's range is finite-dimensional (see Example 4.16(1)).

Positive, *unit* linear functionals $\omega : \mathfrak{B}_\infty(\mathsf{H}) \to \mathbb{C}$ are called **algebraic states** on the C^*-algebra $\mathfrak{B}_\infty(\mathsf{H})$. Their set will be denoted $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$.

Therefore every state $T \in \mathfrak{S}(\mathsf{H})$ determines an algebraic state ω_T of $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$. This can be accompanied by the following characterisation.

Theorem 7.37. *If* H *is a complex Hilbert space, let* $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$ *be the set of bounded, positive and unit linear functionals* $\omega : \mathfrak{B}_\infty(\mathsf{H}) \to \mathbb{C}$ *(*$\omega(A^*A) \geq 0$ *if* $A \in \mathfrak{B}_\infty(\mathsf{H})$*). Then the mapping* $\mathfrak{S}(\mathsf{H}) \ni T \mapsto \omega_T \in \mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$*, with* $\omega_T(A) = tr(TA)$*,* $A \in \mathfrak{B}_\infty(\mathsf{H})$*, is well defined and bijective. Equivalently: states of* $\mathfrak{S}(\mathsf{H})$ *are in one-to-one correspondence with algebraic states of* $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$.

Proof. The map $T \mapsto \omega_T$ is well defined by the above arguments, but also one-to-one. If $\omega_T = \omega_{T'}$ in fact, $tr((T-T')A) = 0$ for any compact operator A. Decomposing the self-adjoint, trace-class operator $T - T'$ over an eigenvector basis $\{\phi_i\}_{i\in I}$ and choosing $A = \phi_i(\phi_i|\ )$ for any $i \in I$, we conclude the eigenvalues of $T - T'$ must all vanish, so $T - T' = 0$ by (6) in Theorem 4.17(b).

Let us prove the surjectivity of $T \mapsto \omega_T$. Considering $\omega \in \mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$ we try to find $T \in \mathfrak{S}(\mathsf{H})$ such that $\omega = \omega_T$. If $\psi, \phi \in \mathsf{H}$, define $A_{\psi,\phi} := \psi(\phi|\ ) \in \mathfrak{B}_\infty(\mathsf{H})$. By definition of norm $||A_{\psi,\phi}|| = ||\psi||\,||\phi||$. The coefficients $\omega(\phi,\psi) := \omega(A_{\psi,\phi})$ define a function from $\mathsf{H}\times\mathsf{H}$ to $\mathbb{C}$, linear in the right argument and antilinear in the left one. Further, $|\omega(\phi,\psi)| = |\omega(A_{\psi,\phi})| \leq 1||A_{\psi,\phi}|| = ||\psi||\,||\phi||$. Then Riesz's representation Theorem 3.16 for Hilbert spaces implies there is a linear $T' : \mathsf{H} \to \mathsf{H}$ such that $\omega(A_{\psi,\phi}) = (T'\psi|\phi)$, for any $\psi, \phi \in \mathsf{H}$. As $||T'\psi||^2 = |(T'\psi|T'\psi)| = |\omega(A_{\psi,T'\psi})| \leq ||\psi||\,||T'\psi||$, we conclude $||T'|| \leq 1$. Setting $T := T'^*$ we have a $T \in \mathfrak{B}(\mathsf{H})$ with $||T|| \leq 1$ and $\omega(A_{\psi,\phi}) = (\psi|T\phi)$ for any $\psi, \phi \in \mathsf{H}$. As ω is positive, taking $\psi = \phi$ implies $T \geq 0$, so in particular $T = T^*$ and $|T| = T$. Now take a basis N of H. If $F \subset N$ is finite define $L_F := \sum_{z\in F} z(z|\ )$. By construction $L_F \in \mathfrak{B}_\infty(\mathsf{H})$ and $||L_F|| \leq 1$ (orthogonal projector). Therefore

$$0 \leq \sum_{z\in F}(z||T|z) = \sum_{z\in F}(z|Tz) = \sum_{z\in F}\omega(A_{z,z}) = \omega(L_F) = |\omega(L_F)| \leq ||\omega||\,.$$

But F is arbitrary, so $\sum_{z\in N}(z||T|z) \leq 1 < +\infty$, and by definition of trace class $T \in \mathfrak{B}_1(\mathsf{H})$. Splitting T over an eigenvector basis, $T = \sum_{i\in I} p_i \psi_i(\psi_i|\ )$ (by construction $p_i \geq 0$, $trT = \sum_i p_i \leq ||\omega||$), and taking the trace, by linearity we have

$$|\omega(A)| = |tr(TA)| \leq \sum_{i\in I} p_i |(\psi_i|A\psi_i)| \leq (trT)||A||$$

if $A \in \mathfrak{B}_\infty(\mathsf{H})$ is a finite combination of $A_{\psi,\phi}$. Since the above A are dense in $\mathfrak{B}_\infty(\mathsf{H})$ in the uniform topology (Theorem 4.21), by continuity and linearity $\omega(A) = tr(TA)$ and $|\omega(A)| \leq trT\,||A||$ for any $A \in \mathfrak{B}_\infty(\mathsf{H})$. The latter tells $||\omega|| \leq tr(T)$; but since we know $trT \leq ||\omega||$, then $trT = ||\omega||$. In particular $trT = 1$, for $||\omega|| = 1$ by assumption. Hence we have $\omega = \omega_T$ for some $T \in \mathfrak{S}(\mathsf{H})$, rendering the map onto. □

Remarks 7.38. One fact becomes evident from the proof: among the theorem's assumptions we may drop the hypothesis that ω has unit norm, and demand, more weakly, that the norm be finite. Then the positive operator $T_\omega \in \mathfrak{B}_1(\mathsf{H})$ corresponding to ω will satisfy $tr(T_\omega) = ||\omega||$. ■

We wish to interpret the result in the light of the theory of the probability measure ρ on the lattice of projectors $\mathfrak{P}(\mathsf{H})$, in the sense of axiom **A2** (preliminary form). To this end recall *Riesz's theorem on positive Borel measures* 1.58 on the locally compact Hausdorff space X. Consider, slightly modifying the theorem's hypotheses, positive linear functionals $\Lambda : C_0(\mathsf{X}) \to \mathbb{C}$, where $C_0(\mathsf{X})$ is the space of continuous complex functions on X that vanish at infinity with norm $||\ ||_\infty$ (Example 2.26(4)).

Proposition 7.39. *If* X *is locally compact, Hausdorff and* $\Lambda : C_0(\mathsf{X}) \to \mathbb{C}$ *is a bounded positive linear functional (with norm* $||\ ||_\infty$ *on the domain), there exists a unique positive and* σ*-additive regular measure* μ_Λ *on* $\mathscr{B}(\mathsf{X})$*, such that* $\Lambda(f) = \int_{\mathsf{X}} f d\mu_\Lambda$ *for any* $f \in C_0(\mathsf{X})$*. Moreover* μ_Λ *is finite and* $||\Lambda|| = \mu_\Lambda(\mathsf{X})$*.*

Proof. The restriction $\Lambda\restriction_{C_c(\mathsf{X})}$ gives a positive functional as in Riesz's Theorem 1.58. Applying the theorem produces a measure $\mu_\Lambda : \mathscr{B}(\mathsf{X}) \to [0,+\infty]$ mapping compact sets to a finite measure, uniquely determined by $\Lambda(f) = \int_{\mathsf{X}} f d\mu_\Lambda$, $f \in C_c(\mathsf{X})$, if we impose μ_Λ is regular. So we assume regularity from now on. If Λ is bounded, easily $\mu_\Lambda(\mathsf{X}) = ||\Lambda||$, so $\mu_\Lambda(\mathsf{X})$ is finite. (For any $f \in C_0(\mathsf{X})$ we have $|\Lambda(f)| \leq \int_X |f| d\mu_\lambda \leq ||f||_\infty \mu_\Lambda(\mathsf{X})$, hence $||\Lambda|| \leq ||\mu_\Lambda(X)||$. For any compact set $K \subset \mathsf{X}$, by local compactness and Hausdorff's property, *Urysohn's lemma* (Theorem 1.24) gives $f_K \in C_c(\mathsf{X})$ such that $f : \mathsf{X} \to [0,1]$ with $f_K(K) = \{1\}$, so $\mu_\Lambda(K) \leq \int_{\mathsf{X}} f_K d\mu_\Lambda \leq ||\Lambda||||f_K||_\infty = ||\Lambda||$ and then $\mu_\Lambda(\mathsf{X}) \leq ||\Lambda||$, because $\mu_\Lambda(\mathsf{X}) = \sup\{\mu_\Lambda(K)\,|\,K \subset \mathsf{X}\,,\text{compact}\}$ by inner regularity of μ_Λ.) μ_Λ finite implies any map of $C_0(\mathsf{X})$ can be integrated. Then the above constraint on the integral in $d\mu_\Lambda$, fixing the regular measure μ_Λ on $\mathscr{B}(\mathsf{X})$, becomes $\Lambda(f) = \int_{\mathsf{X}} f d\mu_\Lambda$ for any $f \in C_0(\mathsf{X})$. □

Knowing that any positive operator $T \in \mathfrak{B}_1(\mathsf{H})$ gives a generalised measure on $\mathfrak{P}(\mathsf{H})$ (a probability measure if $trT = 1$) in the sense of Proposition 7.23, Theorem 7.37 implies a *noncommutative version* of Riesz's representation theorem for *finite measures*, stated in Proposition 7.39. This comes about as follows: think the lattice of orthogonal projectors $\mathfrak{P}(\mathsf{H})$ on the Hilbert space H as the noncommutative variant of the Borel σ-algebra $\mathscr{B}(\mathsf{X})$, and the C^*-algebra of compact operators $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$ as the noncommutative correspondent to the commutative C^*-algebra $C_0(\mathsf{X})$. (Both algebras are without unit if H is infinite-dimensional and X non-compact, respectively.) In that case the bounded positive functional Λ on $C_0(\mathsf{X})$ becomes the bounded positive functional ω on $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$. In either case the existence of positive functionals ω, Λ implies the existence of corresponding finite measures on $\mathfrak{P}(\mathsf{H})$, $\mathscr{B}(\mathsf{X})$ respectively. The latter is what we denoted μ_Λ above, whilst the former is simply defined as $\rho_\omega(P) := tr(T_\omega P)$ for any $P \in \mathfrak{P}(\mathsf{H})$, where $T_\omega \in \mathfrak{B}_1(\mathsf{H})$, ω correspond as in Theorem 7.37 (where T_ω is called T and ω is ω_T). The requests fixing μ_Λ (assumed regular) and ρ_ω are

$$\Lambda(f) = \int_{\mathsf{X}} f(x)\, d\mu(x) \quad \forall f \in C_0(\mathsf{X}) \quad \text{and} \quad \omega(A) = tr(T_\omega A) \quad \forall A \in \mathfrak{B}_\infty(\mathsf{H})$$

respectively. The identity $||\Lambda|| = \mu_\Lambda(\mathsf{X})$ now is $||\omega|| = trT_\omega$.

Remarks 7.40. This discussion serves to explain that the generalisation of the integral of maps in $C_0(\mathsf{X})$ with respect to μ_Λ should be viewed, in the noncommutative setting, as the trace $tr(T_\omega \cdot)$ acting on $\mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H}))$. Hence if $T_\rho \in \mathfrak{B}_1(\mathsf{H})$ is the operator associated (by Remark 7.26(3) only) to a probability measure $\rho : \mathfrak{P}(\mathsf{H}) \to [0,1]$ (fulfilling (1) and (2) of axiom **A2** (preliminary form)) by Gleason's theorem, we will use the writing

$$\int_{\mathfrak{P}(\mathsf{H})} A d\rho := tr(T_\rho A)\,. \tag{7.36}$$

■

Now we can prove the noncommutative version of Proposition 7.39.

Theorem 7.41. *If H is a complex Hilbert space, separable or of finite dimension ≥ 3, and $\omega : \mathfrak{C}(\mathfrak{B}_\infty(\mathsf{H})) \to \mathbb{C}$ is a bounded positive linear functional with unit norm, there exists a unique probability measure $\rho_\omega : \mathfrak{P}(\mathsf{H}) \to [0,1]$ (satisfying (1), (2) in axiom* **A2** *(preliminary form)), such that:*

$$\omega(A) = \int_{\mathfrak{P}(\mathsf{H})} A d\rho_\omega \quad \forall A \in \mathfrak{B}_\infty(\mathsf{H})\,.$$

Suppose further $||\omega|| < +\infty$, but not necessarily one. Then $\rho : \mathfrak{P}(\mathsf{H}) \to [0, ||\omega||]$ and $\rho_\omega(I) = ||\omega||$ instead of $\rho(I) = 1$.

Proof. Define $\rho_\omega(P) := tr(T_\omega P)$ for any $P \in \mathfrak{P}(\mathsf{H})$, where ω and $T_\omega \in \mathfrak{B}_1(\mathsf{H})$ correspond bijectively as in Theorem 7.37. Then $\omega(A) = tr(T_\omega A) =: \int_{\mathfrak{P}(\mathsf{H})} A d\rho_\omega$ for any $A \in \mathfrak{B}_\infty(\mathsf{H})$, because of (7.36) and T_ω is by construction associated to the measure ρ_ω by Gleason's theorem. (Proposition 7.23 ensures ρ_ω fulfills (1), (2) in axiom **A2** (preliminary form)). Let us prove uniqueness. By Gleason's Theorem 7.24 and Remark 7.26(3), every probability measure ρ on $\mathfrak{P}(\mathsf{H})$ satisfies $\rho(P) = tr(T_\rho P)$ for any $P \in \mathfrak{P}(\mathsf{H})$ and a unique positive operator T_ρ of trace class with unit trace. If $\omega(A) =$

$\int_{\mathfrak{P}(\mathsf{H})} A d\rho := tr(T_\rho A)$ for any compact operator A, since we saw $\omega(A) = tr(T_\omega A)$, choosing $A = \psi(\psi|\ )$ will give $(\psi|(T_\omega - T_\rho)\psi) = 0$ for any $\psi \in \mathsf{H}$. Hence $T_\rho = T_\omega$, and consequently $\rho(P) = tr(T_\rho P) = tr(T_\omega P) = \rho_\omega(P)$ for any $P \in \mathfrak{P}(\mathsf{H})$. All this extends to the case $0 < ||\omega|| \neq 1$, by using the functional $\omega' := ||\omega||^{-1}\omega$. If $||\omega|| = 0$ then $\omega = 0$. Therefore a possible measure ρ compatible with $0 = \omega(A) = Tr(T_\rho A)$ for any $A \in \mathfrak{B}_\infty(\mathsf{H})$ is the null measure. It is unique by the same argument. □

7.5 Observables as projector-valued measures on $\mathbb{R}$

Now we want to define *observables* by *projector-valued measures* (PVM). This notion lies at the heart of the mathematical formulation of standard QM. In ensuing chapters the notion will be generalised, and made more precise mathematically; the culmination will be the proof of the spectral decomposition theorem for unbounded self-adjoint operators, whose statement brings PVMs to the fore.

7.5.1 *Axiom* A4*: the notion of observable*

In Quantum Mechanics, physical quantities testable on physical systems and whose behaviour is described by **QM1** and **QM2**, are called **observables**. Now we shall discuss them.

As seen in Chapter 7.2.4, it is reasonable to label physical quantities' measurement readings by Borel sets of $\mathbb{R}$. From the physical point of view it is natural to assume that if the orthogonal projectors $P_E^{(A)}$ associated to the observable A commute with each other then $E \in \mathscr{B}(\mathbb{R})$ (the Borel σ-algebra of $\mathbb{R}$), since we expect propositions like

$$P_E^{(A)} :=$$

The value of A on the state of the system belongs to the Borel set $E \subset \mathbb{R}$,

to be all compatible for $E \in \mathscr{B}(\mathbb{R})$. If it were not so we would not have an observable, but distinct incompatible quantities. Since the outcome belongs to both E and E' if and only if it belongs to $E \cap E'$, we take $P_E^{(A)} \wedge P_{E'}^{(A)} = P_{E\cap E'}^{(A)}$. Assume also $P_{\mathbb{R}}^{(A)} = I$, because the result certainly belongs to $\mathbb{R}$, so $P_{\mathbb{R}}^{(A)}$ is a tautology, independent of the state on which the measurement is done. Eventually, for physically self-evident reasons and because of the *logical* meaning of $\vee$, it is reasonable to suppose

$$\vee_{n\in\mathbb{N}} P_{E_n}^{(A)} = P_{\cup_{n\in\mathbb{N}} E_n}^{(A)}$$

for any finite or countable collection $\{E_n\}_{n\in\mathbb{N}}$ of Borel sets of $\mathbb{R}$. Although one could also take sets of arbitrary cardinality, we will stop at countable, as we did in the classical case.

Definition 7.42. *If* H *is a Hilbert space, a function A mapping* $E \in \mathscr{B}(\mathbb{R})$ *to an orthogonal projector* $P_E^{(A)} \in \mathfrak{P}(\mathsf{H})$ *is called an* **observable** *if:*

(a) $P_E^{(A)} P_{E'}^{(A)} = P_{E'}^{(A)} P_E^{(A)}$ *for any* $E, E' \in \mathscr{B}(\mathbb{R})$.
(b) $P_E^{(A)} \wedge P_{E'}^{(A)} = P_{E \cap E'}^{(A)}$ *for any* $E, E' \subset \mathscr{B}(\mathbb{R})$.
(c) $P_{\mathbb{R}}^{(A)} = I$.
(d) *For any family* $\{E_n\}_{n \in \mathbb{N}} \subset \mathscr{B}(\mathbb{R})$*:*

$$\vee_{n\in\mathbb{N}} P_{E_n}^{(A)} = P_{\cup_{n\in\mathbb{N}} E_n}^{(A)} .$$

Remark 7.43. (1) It is straightforward to see that $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$ is a Boolean σ-algebra for the usual partial order relation $\leq$ of projectors. In general $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$ is not maximal commutative.
(2) Bearing in mind Definition 7.11 it is easy to prove any observable is nothing but a *homomorphism of Boolean* σ*-algebras*, mapping the Borel σ-algebra $\mathscr{B}(\mathbb{R})$ to the Boolean σ-algebra of projectors $\{P_E^{(A)}\}_{E\subset\mathscr{B}(\mathbb{R})}$. It can be proved ([Jau73, Chapter 6-5]) that if H is a separable Hilbert space, any subset of projectors in $\mathfrak{P}(\mathsf{H})$ forming a Boolean σ-algebra is automatically an observable, i.e. of the form $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$, and satisfies Definition 7.42. ■

Observables may be redefined, in an equivalent way but mathematically simpler, as the next proposition shows.

Proposition 7.44. *Let* H *be a Hilbert space. A map* $P : \mathscr{B}(\mathbb{R}) \to \mathfrak{B}(\mathsf{H})$ *is an observable if and only if the following hold.*

(a) $P(B) \geq 0$ *for any* $B \in \mathscr{B}(\mathbb{R})$.
(b) $P(B)P(B') = P(B \cap B')$ *for any* $B, B' \in \mathscr{B}(\mathbb{R})$.
(c) $P(\mathbb{R}) = I$.
(d) *for any family* $\{B_n\}_{n\in\mathbb{N}} \subset \mathscr{B}(\mathbb{R})$ *with* $B_n \cap B_m = \varnothing$ *if* $n \neq m$*:*

$$s\text{-}\sum_{n=0}^{+\infty} P(B_n) = P(\cup_{n\in\mathbb{N}} B_n) .$$

Proof. If $P : \mathscr{B}(\mathbb{R}) \to \mathfrak{B}(\mathsf{H})$ is an observable properties (a), (b), (c), (d) in Proposition 7.44 are trivially true. So we have to prove any $P : \mathscr{B}(\mathbb{R}) \to \mathfrak{B}(\mathsf{H})$ satisfying them is an observable.

Let us put all operators $P(B)$ with $B \in \mathscr{B}(\mathbb{R})$ in a maximal set of commuting projectors $\mathfrak{P}_0(\mathsf{H})$ (which exists by Zorn's lemma), and from now on we shall work in it without loss of generality.
(a) Implies every $P(B)$ is self-adjoint by Proposition 3.54(f), so (b) implies $P(B)P(B) = P(B \cap B) = P(B)$, whence every $P(B)$ is an orthogonal projector. Moreover (b) implies, if $P(B)P(B') = P(B \cap B') = P(B' \cap B) = P(B')P(B)$, that all projectors commute with one another. Using the first identity in (i) of Theorem 7.18(b), condition (b) above reads $P(B) \wedge P(B') = P(B \cap B')$. To finish we need to show property (d) of Definition 7.42. Consider countably many sets $\{E_n\}_{n\in\mathbb{N}} \subset \mathscr{B}(\mathbb{R})$, not disjoint in general. We want to find $\vee_{n\in\mathbb{N}} P(E_n)$ and prove that

$$\vee_{n\in\mathbb{N}} P(E_n) = P(\cup_{n\in\mathbb{N}} E_n) .$$

To do so define a collection $\{B_n\}_{n\in\mathbb{N}}$ of pairwise disjoint Borel sets: $B_0 := E_0$ and

$$B_n = E_n \setminus (E_1 \cup \cdots \cup E_{n-1})$$

for $n > 0$. Then

$$\cup_{n=0}^{p} E_n = \cup_{n=0}^{p} B_n \quad \text{for any } p \in \mathbb{N} \cup \{+\infty\}.$$

From this, using $I - P(B) = P(\mathbb{R} \setminus B)$ and recursively the second identity in (i) of Theorem 7.18(b), we find

$$\vee_{n=0}^{p} P(E_n) = \vee_{n=0}^{p} P(B_n) \quad \text{for any } n \in \mathbb{N}.$$

By the fact that part (d) of the present proposition implies

$$\vee_{n=0}^{p} P(B_n) = \sum_{n=0}^{p} P(B_n)$$

for finitely many disjoint B_n (this collection may be made countable by adding infinitely many empty sets), we have

$$\vee_{n=0}^{p} P(E_n) = \sum_{n=0}^{p} P(B_n)\,. \tag{7.37}$$

To conclude we take the strong limit as $p \to +\infty$ in (7.37). This exists by Theorem 7.18(b), and we also have

$$\vee_{n\in\mathbb{N}} P(E_n) = \text{s-}\lim_{p\to+\infty} \sum_{n=0}^{p} P(B_n) = P(\cup_{n\in\mathbb{N}} B_n) = P(\cup_{n\in\mathbb{N}} E_n)\,. \qquad \square$$

Remark 7.45. (1) Notice that (c) and (d) alone imply $I = P(I \cup \varnothing) = I + P(\varnothing)$, so

$$P(\varnothing) = 0\,.$$

(2) If $B \in \mathscr{B}(\mathbb{R})$ then $\mathbb{R} \setminus B \in \mathscr{B}(\mathbb{R})$ and $\mathbb{R} = B \cup (\mathbb{R} \setminus B)$. By (d), taking $B_0 = B$, $B_1 = \mathbb{R} \setminus B$ and all remaining $B_k = \varnothing$, we obtain $I = P(B) + P(\mathbb{R} \setminus B)$. Therefore

$$\neg P(B) = P(\mathbb{R} \setminus B)\,. \qquad \blacksquare$$

The above proposition allows to identify observables bijectively with well-known objects in mathematics, namely *projector-valued measures* on $\mathbb{R}$. The latter will be generalised in the next chapter.

Definition 7.46. *A map $P : \mathscr{B}(\mathbb{R}) \to \mathfrak{B}(\mathsf{H})$, H Hilbert space, satisfying (a), (b), (c) and (d) in Proposition 7.44 is called* **projector-valued measure (PVM) on** $\mathbb{R}$ *or* **spectral measure on** $\mathbb{R}$.

We are in the position to disclose the fourth axiom of the general mathematical formulation of Quantum Mechanics.

A4. *Every observable A on the quantum system S is described by a projector-valued measure $P^{(A)}$ on $\mathbb{R}$ in the Hilbert space H_S of the system, in such a way that if E is a Borel set in $\mathbb{R}$, the projector $P^{(A)}(E)$ corresponds to the proposition "the reading of a measurement of A falls in the Borel set E".*

Remark 7.47. (1) Let us suppose that, owing to a superselection rule, the Hilbert space splits into coherent sectors $\mathsf{H}_S = \oplus_{k\in K}\mathsf{H}_{Sk}$. Call P_k the orthogonal projector on H_{Sk}. Recalling Remark 7.36(2), every projector $P_E^{(A)}$ of an observable A satisfies $P_k P_E^{(A)} = P_E^{(A)} P_k$ for any $k \in K$ and any Borel set $E \subset \mathbb{R}$.
(2) The observable B **is a function of the observable** A, $B = f(A)$, when there is a measurable map $f : \mathbb{R} \to \mathbb{R}$ such that $P^{(B)}(E) = P^{(A)}(f^{-1}(E))$ for any Borel set $E \subset \mathbb{R}$. This is totally natural: if "$B = f(A)$" then to measure B we can measure A and use f on the reading. In this sense the outcome of measuring B belongs to E iff the outcome of A belongs in $f^{-1}(E)$. In particular, the elementary propositions (orthogonal projectors) $P_E^{(B)}$ and $P_F^{(A)}$ are always compatible, and $\{P_E^{(B)}\}_{E\in\mathscr{B}(\mathbb{R})} \subset \{P_F^{(A)}\}_{F\in\mathscr{B}(\mathbb{R})}$. It is possible to prove [Jau73] that for given observables A, B in a separable Hilbert space, the previous inclusion is equivalent to the existence of a measurable map f such that $B = f(A)$. More important is a result of von Neumann and Varadarajan [Jau73, Chapter 6-7] (valid for any orthocomplemented, σ-complete, separable lattice, not necessarily the lattice of projectors on a Hilbert space):

Theorem 7.48. *If $\{A_j\}_{j\in J}$ is a family of pairwise-compatible observables (that is, $P^{(j)}(E)P^{(i)}(F) = P^{(i)}(F)P^{(j)}(E)$ if $P^{(j)}(E) \in A_j$, $P^{(i)}(F) \in A_i$) on a separable Hilbert space, there exists an observable A and a corresponding family of measurable maps $f_j : \mathbb{R} \to \mathbb{R}$, $j \in J$ such that $A_j = f_j(A)$ for any $j \in J$.* ∎

7.5.2 Self-adjoint operators associated to observables: physical motivation and basic examples

This section contains the idea underlying the correspondence between self-adjoint operators and observables. We will, in other words, provide the physical motivation for the spectral theorems of Chapter 8, 9.

For classical systems, at time t on phase space $\mathscr{F}$, we know that observables correspond to what have been called physical quantities, i.e. measurable maps $f : \mathscr{F} \to \mathbb{R}$. To any physical quantity f we can associate the collection of all propositions/Borel sets of the form:

$$P_E^{(f)} :=$$

The value of f on the system's state belongs to the Borel set $E \subset \mathbb{R}$,

or, set-theoretically,

$$P_E^{(f)} = \text{“}f^{-1}(E) \subset \mathscr{F}_t\text{”}$$

Propositions 7.14 and 7.15 explained that $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$ is a (Boolean) σ-algebra and the map $\mathscr{B}(\mathbb{R}) \ni E \mapsto P_E^{(f)} \in \mathscr{B}(\mathscr{F})$ a Boolean σ-algebra homomorphism. The picture is the same in the quantum case when we look at the class $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$ of propositions/projectors associated to an observable A: they form a Boolean σ-algebra and $\mathscr{B}(\mathbb{R}) \ni E \mapsto P_E^{(A)} \in \mathfrak{P}(\mathsf{H})$ a homomorphism of Boolean σ-algebras. If we restrict to comparing $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$ and $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$ the situation is analogous. In the classical case, though, *there exists a function* f consenting to *build* the collection $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$: this map retains, alone, all possible information about the propositions $P_E^{(f)}$. This is no surprise since we defined propositions/sets starting from f! In the quantum case, when an observable $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$ is given, there is nothing, at least at present, that may correspond to a function f "generating" the PVM $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$.
Is there a quantum analogue to f?
In trying to aswer let us dig deeper into the relationship between f and the associated family $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$. We know how to get the latter out of the former, but now we are interested in recovering the map from the family, because in the quantum formulation one starts from the analogue of $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$. As a matter of fact the σ-algebra $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$ allows to *reconstruct* f by means of a certain limiting process reminiscent of *integration*.

To explain this point we need a technical result. Recall that if (X,Σ) is a measure space, a Σ-measurable map $s : \mathsf{X} \to \mathbb{C}$ is **simple** if its range is finite.

Proposition 7.49. *Let (X,Σ) be a measure space, $S(\mathsf{X})$ the space of complex-valued simple functions with respect to Σ, $M(\mathsf{X})$ the space of $\mathbb{C}$-valued, Σ-measurable maps, and $M_b(\mathsf{X}) \subset M(\mathsf{X})$ the subspace of bounded maps. Then:*

(a) *$S(\mathsf{X})$ is dense in $M(\mathsf{X})$ pointwise.*
(b) *$S(\mathsf{X})$ is dense in $M_b(\mathsf{X})$ in norm $|| \; ||_\infty$.*
(c) *If $f \in M(\mathsf{X})$ ranges on non-negative reals, there is a sequence $\{s_n\}_{n\in\mathbb{N}} \subset S(\mathsf{X})$ with:*

$$0 \le s_0 \le s_1 \le \cdots \le s_n(x) \to f(x) \quad \text{as } n \to +\infty, \text{for any } x \in \mathsf{X},$$

and provided $f \in M_b(\mathsf{X})$, convergence is in norm $|| \; ||_\infty$ as well.

Proof. It is enough to prove the claim for real-valued maps, for the complex case is a consequence of decomposing complex functions into real and imaginary parts. Define $f_+(x) := \sup\{0, f(x)\}$ and $f_-(x) := \inf\{0, f(x)\}$, $x \in \mathsf{X}$; then $f = f_+ + f_-$, where $f_+ \ge 0$, $f_- \le 0$ are known to be measurable since f is. Now we construct a sequence of simple maps converging to f_+ (whence part (c) is proven, as $f = f_+$ if $f \ge 0$). For given $0 < n \in \mathbb{N}$ let us partition the real semi-axis $[0, +\infty)$ into Borel sets $E_{n,i}$, E_n where:

$$E_{n,i} := \left[\frac{i-1}{2^n}, \frac{i}{2^n}\right), \quad E_n := [n, +\infty),$$

$1 \leq i \leq n2^n$. For each n let

$$P_{n,i}^{(f)} := f^{-1}(E_{n,i}) , \quad P_n^{(f)} := f^{-1}(E_n)$$

be subsets in Σ. Then define $s_0(x) := 0$ if $x \in \mathsf{X}$, and

$$s_n := \sum_{i=1}^{n2^n} \frac{i-1}{2^n} \chi_{P_{n,i}^{(f)}} + n\chi_{P_n^{(f)}} \tag{7.38}$$

for any $n \in \mathbb{N} \setminus \{0\}$. By construction $0 \leq s_n \leq s_{n+1} \leq f$, $n = 1, 2, \ldots$. Furthermore, for any given x we have $|f_+(x) - s_n(x)| \leq 1/2^n$ definitely. Evidently, then, $s_n \to f_+$ pointwise if $n \to +\infty$. The estimate $|f_+(x) - s_n(x)| \leq 1/2^n$ is uniform in x if f_+ is bounded (take $n > \sup f_+$), and then $s_n \to f_+$ also uniformly. Similarly, by decomposing the negative semi-axis we can construct another simple sequence $\{s_n^{(-)} \leq 0\}$ pointwise tending to f_-. Overall the simple sequence $s_n^{(-)} + s_n$ converges to f pointwise, and uniformly if f is additionally bounded. □

Remarks 7.50. If f is non-negative, part (a) still holds even when $f : \mathsf{X} \to [0, +\infty]$, by taking simple maps that can equal $+\infty$. ■

It is thus clear that a given classical quantity $f : \mathscr{F} \to \mathbb{R}$ (measurable) can be recovered using a sequence of maps that are constant non-zero only on sets in $\{P_E^{(f)}\}_{E \in \mathscr{B}(\mathbb{R})}$. Without loss of generality we focus on the situation $f : \mathscr{F} \to \mathbb{R}_+$ and further suppose f bounded; this entitles us to forget, in (7.38) and for n large enough: (i) *all intervals* E_n and (ii) the $E_{n,i}$ with left endpoint $((i-1)2^{-n})$ bigger, say, than $(\sup f) + 1/2^n$, for the pre-image of these sets under f is empty. If we do so the sum in (7.38) can be truncated:

$$f = \lim_{n \to +\infty} \sum_{i=1}^{2+2^n \sup f} \frac{i-1}{2^n} \chi_{P_{n,i}^{(f)}} . \tag{7.39}$$

This limit may be understood as an *integration* of sorts with respect to a "measure valued on characteristic functions":

$$\nu^{(f)} : \mathscr{B}(\mathbb{R}) \ni E \mapsto \chi_{f^{-1}(E)} \in S(\mathsf{X}),$$

associating to a Borel subset $E \subset \mathbb{R}$ (in the range space of the map) a characteristic function $\chi_{f^{-1}(E)} : \mathsf{X} \to \mathbb{C}$. Observe, in fact, that $\frac{i-1}{2^n}$ is approximately the value f assumes at $P_{n,i}^{(f)}$ – and the estimate becomes more accurate as n increases – and the right-hand side in (7.39) is just a "Cauchy sum". Equation (7.39) might be formally written as:

$$f = \int_{\mathbb{R}} \lambda d\nu^{(f)}(\lambda) . \tag{7.40}$$

But as we are concerned with the quantum setting, we will not push the analogy further, though doing so would give a rigorous meaning to the above integral. In such case the similar formula to (7.40) is:

$$A = \int_{\mathbb{R}} \lambda dP^{(A)}(\lambda)$$

where the characteristic functions $\chi_{f^{-1}(E)}$ are replaced by the orthogonal projectors $P_E^{(A)}$ of the observable A. This relation defines a *self-adjoint operator* A associated to an observable, that was called A and that corresponds to the classical quantity f. From such an operator the observable $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$ can be recovered, *a posteriori*, in a similar manner to what we do to get $\{P_E^{(f)}\}_{E\in\mathscr{B}(\mathbb{R})}$ out of f. We will see all this in full generality, and rigorously, in the sequel. At this juncture we describe an elementary example of observable and show how to associate to it a self-adjoint operator.

Examples 7.51. (1) Consider a quantum system, described on a Hilbert space H, and take a physical quantity ranging, from the point of view of physics, over a discrete and finite set of *distinct* values $\{a_n\}_{n=1,\cdots,N} \subset \mathbb{R}$. We first show how to find an observable A given by a family of orthogonal projectors $P_E^{(A)}$, $E \in \mathscr{B}(\mathbb{R})$. We posit there are *non-null* orthogonal projectors labelled by a_n, $\{P_{a_n}\}_{n=1,\cdots,N}$, such that $P_{a_n}P_{a_m} = 0$ if $n \neq m$ (i.e., taking adjoints, $P_{a_m}P_{a_n} = 0$ if $n \neq m$), and moreover:

$$\sum_{n=1}^{N} P_{a_n} = I\,. \tag{7.41}$$

The meaning of P_{a_n}, clearly, is:

The value of A, read by a measurement on the system, is precisely a_n.

Obviously the equations $P_{a_n}P_{a_m} = P_{a_m}P_{a_n} = 0$, i.e. $P_{a_n} \wedge P_{a_m} = 0$ for $n \neq m$ correspond to two physical requirements: (a) propositions P_{a_n}, P_{a_m} are physically compatible, but (b) the observable's measurement cannot produce distinct values a_n and a_m simultaneously (the proposition associated to the null projector is contradictory). Demanding $\sum_{n=1}^{N} P_{a_n} = I$, i.e. $P_{a_1} \vee \cdots \vee P_{a_n} = I$, amounts to asking that at least one proposition P_{a_n} is true when measuring A. The observable $A : \mathscr{B}(\mathbb{R}) \to \mathfrak{P}(\mathsf{H})$ is built as follows: for any Borel set $E \subset \mathbb{R}$

$$P_E^{(A)} := \sum_{a_n \in E} P_{a_n}\,, \quad \text{with } P_{\varnothing}^{(A)} := 0\,. \tag{7.42}$$

Properties (a), (b), (c) and (d) in Proposition 7.44 are immediate.

(2) Referring to example (1), to the observable A we can associate an operator still called A. Here is how:

$$A := \sum_{n=0}^{N} a_n P_{a_n}\,. \tag{7.43}$$

A is bounded and self-adjoint by construction, being a real linear combination of self-adjoint operators. It has another interesting property: *the eigenvalue set $\sigma_p(A)$ of A coincides with the values the observable A can assume.*

The proof is direct: if $0 \neq u \in P_{a_n}(\mathsf{H})$ then $P_{a_m}u = P_{a_m}P_{a_n}u = u$ if $n = m$ or 0 if $n \neq m$. Inserting this in (7.43) gives $Au = a_n u$, so $a_n \in \sigma_p(A)$. Conversely, if $u \neq 0$ is an eigenvector of A with eigenvalue λ (real since $A = A^*$), then (7.43) implies

$$\lambda u = \sum_{n=0}^{N} a_n P_{a_n} u\,.$$

On the other hand, since $\sum_{a_n} P_{a_n} = I$ and modifying the left-hand side of the above identity, we obtain

$$\sum_{n=0}^{N} \lambda P_{a_n} u = \sum_{n=0}^{N} a_n P_{a_n} u \,,$$

hence

$$\sum_{n=0}^{N} (\lambda - a_n) P_{a_n} u = 0 \,. \tag{7.44}$$

Now apply P_m and recall $P_m P_n = \delta_{m,n} P_n$, resulting in N identities:

$$(\lambda - a_m) P_{a_m} u = 0 \,.$$

If all of them were solved by $P_m u = 0$ for any m, we would obtain a contradiction, because

$$0 \neq u = Iu = \sum_{n=0}^{N} P_{a_n} u \,.$$

Therefore there must be some n in (7.44) for which $\lambda = a_n$. This can happen for one value n only, since by assumption the a_n are distinct. Overall the eigenvalue λ of A must be one a_n. So we proved that the eigenvalue set of A coincides with the values A can assume. The self-adjoint operator A here plays a role similar to that of f relatively to the classical quantity $\{P_E^{(f)}\}_{E \in \mathscr{T}(\mathbb{R})}$.

(3) To conclude suppose A is the operator of an observable in the sense of (7.43), and $g : \mathbb{R} \to \mathbb{R}$ is a function. We can define a new observable, thought as if it were function of the previous one, determined entirely by the self-adjoint operator

$$C := g(A) := \sum_{n=0}^{N} g(a_n) P_{a_n} \,. \tag{7.45}$$

By construction the possible values of the new observable are the images $g(a_n)$, that in turn determine the eigenvalues of C. ■

In the next chapters we will develop a procedure for associating, uniquely, to each observable A (i.e. a projector-valued measure on $\mathbb{R}$) a self-adjoint operator (typically unbounded) denoted by the same letter A, thereby generalising the previous examples. The values the observable can take will be elements in the *spectrum* $\sigma(A)$, which we will explain is normally larger than the set $\sigma_p(A)$ of eigenvalues. The major tool will be integration with respect to a projector-valued measure, corresponding to a generalisation of

$$\sum_{\lambda \in \sigma_p(A)} h(\lambda) P_\lambda =: \int_{\sigma(A)} h(\lambda)\, dP^{(A)}(\lambda)$$

to the case when the values of λ can be infinite (cf. (7.41) for $h : \lambda \to 1$ and (7.43) for $h : \lambda \to 1$). In particular

$$A = \int_{\sigma(A)} \lambda \, dP^{(A)}(\lambda) \,, \quad I = \int_{\sigma(A)} 1 \, dP^{(A)}(\lambda) ,$$

whose interpretation befits the theory of spectral measures.

7.5.3 Probability measures associated to state/observable couples

Here is yet another remarkable property about PVMs on $\mathbb{R}$, with important consequences in physics.

Proposition 7.52. *Let* H *be a Hilbert space and* $A = \{P_E\}_{E\in\mathscr{B}(\mathbb{R})}$ *a projector-valued measure on* $\mathbb{R}$*. If* $\rho \in \mathfrak{S}(\mathsf{H})$ *is a state, the map* $\mu_\rho^{(A)} : E \mapsto tr(\rho P_E)$ *is a Borel probability measure on* $\mathbb{R}$*.*

Proof. The proof is elementary. It suffices to show $\mu_\rho^{(A)}$ is positive, σ-additive and $\mu_\rho^{(A)}(\mathbb{R}) = 1$. $\mathbb{R}$ is Hausdorff and locally compact, so every positive σ-additive measure on the Borel algebra is a Borel measure. Decompose ρ in the usual way with an eigenvector basis:

$$\rho = \sum_{j\in\mathbb{N}} p_j(\psi_j|\;)\psi_j\,,$$

where the p_j are non-negative and their sum is 1. Then $\mu_\rho^{(A)}(E) = tr(\rho P_E) \geq 0$ because orthogonal projectors are positive, $P_j \geq 0$ and $tr(\rho P_E) = \sum_{j\in\mathbb{N}} p_j(\psi_j|P_E\psi_j)$. Moreover $\mu_\rho^{(A)}(\mathbb{R}) = 1$, since $P_\mathbb{R} = I$ implies

$$\sum_{j\in\mathbb{N}} p_j(\psi_j|I\psi_j) = tr\rho = 1\,.$$

Let us show σ-additivity. If $\{E_n\}_{n\in\mathbb{N}}$ are pairwise disjoint Borel sets and $E := \cup_{n\in\mathbb{N}}E_n$, by Proposition 7.44(d):

$$+\infty > tr(\rho P_E) = \sum_{j=0}^{+\infty} p_j\left(\psi_j\middle|\sum_{i=0}^{+\infty} P_{E_i}\psi_j\right) = \sum_{j=0}^{+\infty}\sum_{i=0}^{+\infty} p_j(\psi_j|P_{E_i}\psi_j)\,.$$

Since $p_j \geq 0$ and $(\psi_j|P_{E_i}\psi_j) \geq 0$, Fubini's theorem allows to swap the series:

$$tr(\rho P_E) = \sum_{i=0}^{+\infty}\sum_{j=0}^{+\infty} p_j(\psi_j|P_{E_i}\psi_j) = \sum_{i=0}^{+\infty} tr(\rho P_{E_i})\,.$$

That is to say, if $\{E_n\}_{n\in\mathbb{N}}$ are pairwise disjoint Borel sets then

$$\mu_\rho^{(A)}(\cup_{n\in\mathbb{N}}E_n) = \sum_{n=0}^{+\infty}\mu_\rho^{(A)}(E_n)\,,$$

ending the proof. □

Examples 7.53. (1) The observable A of examples (1) and (2) in 7.51 assumes a finite number N of discrete values a_n. Let A (7.43) be the self-adjoint operator of the observable. Fix a state $\rho \in \mathfrak{S}(\mathsf{H})$ and consider its probability measure relative to the observable $\{P_E^{(A)}\}_{E\in\mathscr{B}(\mathbb{R})}$. By construction, if $E \in \mathscr{B}(\mathbb{R})$:

$$\mu_\rho^{(A)}(E) := tr(\rho P_E^{(A)}) = \sum_{a_n\in E} tr(\rho P_{a_n}) = \sum_{a_n} p_n\delta_{a_n}(E)$$

with

$$p_n := tr(\rho P_{a_n}) .$$

Hence

$$\mu_\rho^{(A)} = \sum_{a_n} p_n \delta_{a_n} , \tag{7.46}$$

where δ_{a_n} are Dirac measures centred at a_n: $\delta_a(E) = 1$ if $a \in E$, $\delta_a(E) = 0$ if $a \notin E$. Note $0 \leq p_n \leq 1$ and $\sum_n p_n = 1$ by construction. Thus the probability measure associated to the state ρ and relative to A is a convex combination of Dirac measures.
(2) The mean value of A on state ρ, $\langle A \rangle_\rho$, and its standard deviation ΔA_ρ^2 on ρ, can be written succinctly using the associated operator A of (7.43). By definition of mean value

$$\langle A \rangle_\rho = \int_{\mathbb{R}} a \, d\mu_\rho^{(A)}(a) .$$

On the other hand, by (7.46) we have

$$\int_{\mathbb{R}} a \, d\mu_\rho^{(A)}(a) = \sum_n p_n a_n = \sum_n a_n tr(\rho P_{a_n}) .$$

Using (7.43) and the linearity of the trace, we conclude

$$\langle A \rangle_\rho = tr(A\rho) . \tag{7.47}$$

In case ρ is pure, i.e. $\rho = \psi(\psi|\cdot)$, $||\psi|| = 1$, (7.47) implies

$$\langle A \rangle_\psi = (\psi|A\psi) \tag{7.48}$$

if $\langle A \rangle_\psi$ indicates the mean value of A on the state of the vector ψ. By definition the deviation equals

$$\Delta A_\rho^2 = \int_{\mathbb{R}} a^2 \, d\mu_\rho^{(A)}(a) - \langle A \rangle_\rho^2 .$$

Proceeding as before,

$$\int_{\mathbb{R}} a^2 \, d\mu_\rho^{(A)}(a) = \sum_n p_n a_n^2 = \sum_n a_n^2 tr(\rho P_{a_n}) = tr\left(\rho \sum_n a_n^2 P_{a_n}\right) .$$

Now observe

$$A^2 = \sum_n a_n P_{a_n} \sum_m a_m P_{a_m} = \sum_{n,m} a_n a_m P_{a_n} P_{a_m} = \sum_n a_n^2 P_{a_n} ,$$

where we used $P_{a_n} P_{a_m} = \delta_{n,m} P_n$. Therefore

$$\Delta A_\rho^2 = tr\left(\rho A^2\right) - \left(tr\left(\rho A\right)\right)^2 . \tag{7.49}$$

If ρ is a pure state, i.e. $\rho = \psi(\psi|\cdot)$, $||\psi|| = 1$, we have, from (7.49),

$$\Delta A_\psi^2 = \left(\psi|A^2\psi\right) - (\psi|A\psi)^2 = \left(\psi|\left(A^2 - \langle A \rangle_\psi^2\right)\psi\right) \tag{7.50}$$

if ΔA_ψ^2 is the standard deviation of A on the state determined by the vector ψ. ■

The formulas in the above examples about mean values and standard deviations of observables in given states, are actually valid more generally (with suitable technical assumptions). This will be proved in Proposition 11.8, after showing in full generality the procedure for associating self-adjoint operators to observables.

Exercises

7.1. Prove that in a Boolean algebra X, for any $a \in \mathsf{X}$ there exists a unique element, written $\neg a$, that satisfies properties (i), (ii) in Definition 7.8(c).

7.2. Keeping in mind, in particular, condition (iv) in the definition of orthocomplementation, prove that every orthocomplemented lattice X satisfies De Morgan's laws (7.5).

7.3. Show that an orthocomplemented lattice is σ-complete iff every countable set $\{a_n\}_{n\in\mathbb{N}} \subset \mathsf{X}$ admits greatest lower bound.

7.4. Prove Propositions 7.13 and 7.14.

7.5. Prove Theorem 7.20.

Solution. $Q(PQ)^n = (QPQ)^n = (QP)^n Q$. The sequence $A_n = (QPQ)^n \in \mathfrak{B}(\mathsf{H})$ satisfies $A_n \geq A_{n+1} \geq 0$: $||\sqrt{QPQ}||^2 = ||QPQ|| \leq ||Q||^2||P|| \leq 1$ and $0 \leq (x|A_{n+1}x) = ||\sqrt{QPQ}(QPQ)^{n/2}x||^2 \leq ||\sqrt{QPQ}||^2||(QPQ)^{n/2}x||^2 = ||\sqrt{QPQ}||^2(x|A_n x)$. By Proposition 3.65, $\text{s-}\lim_{n\to+\infty}(QPQ)^n = R \in \mathfrak{B}(\mathsf{H})$. Immediately, $RR = R$ and $(Rx|y) = \lim_n (x|(Q(PQ)^n)^* y) = \lim_n (x|(QP)^n Qy) = (x|Ry)$, so $R = R^*$. By construction $PR = \text{s-}\lim_n PQ(PQ)^n = R$. Therefore $(PQ)^n \to R \in \mathfrak{P}(\mathsf{H})$ in the strong topology. Analogously $(QP)^n \to R' \in \mathfrak{P}(\mathsf{H})$ in the same topology. However, $(x|(PQ)^n y) \to (x|Ry)$ is equivalent to $((QP)^n x|y) \to (x|Ry)$, i.e. $(R'x|y) = (x|Ry)$. Since $R' = R'^*$ we have $R' = R$. Clearly $RP = R = RQ$, so $R(\mathsf{H}) \supset P(\mathsf{H}) \cap Q(\mathsf{H})$. The orthogonal to the latter space is generated (by De Morgan's laws) by $(\neg P)(\mathsf{H})$ and $(\neg Q)(\mathsf{H})$. As $R(\neg P) = R(\neg Q) = 0$, we conclude $R(\mathsf{H}) = P(\mathsf{H}) \cap Q(\mathsf{H})$.

7.6. Consider two self-adjoint operators

$$A = \sum_{n=1}^{N} a_n P_{a_n} \quad \text{and} \quad B = \sum_{m=1}^{M} b_m Q_{b_m}$$

that represent, as in Examples 7.51, observables in the d-dimensional Hilbert space H (d finite). Show A, B commute iff the orthogonal projectors P_{a_n}, Q_{b_m} commute, irrespective of how we choose the eigenvalues a_n, b_m.

Hint. Identify H with $\mathbb{C}^d$ and diagonalise simultaneously the matrices representing A and B.

7.7. Consider to self-adjoint operators A, B representing, as in the previous exercise, observables in a Hilbert space H of finite dimension d. Prove that if A and B commute, there exists a third observable (self-adjoint operator) C such that: $A = f(C)$ and $B = g(C)$ in the sense of (7.45), for some real-valued maps on $\mathbb{R}$. Show that C, f and g can be chosen in infinitely many ways.

Hint. If $\{\psi_n\}_{n=1,\ldots,d}$ is an orthonormal basis of H of eigenvectors for both A and B, define $C := \sum_{k=1}^d k\psi_k(\psi_k|\cdot)$. We must find f, g such that $A = \sum_{k=1}^d f(k)\psi_k(\psi_k|\cdot)$ and $B = \sum_{k=1}^d g(k)\psi_k(\psi_k|\cdot)$. At this point the choice for f, g should be patent.

7.8. Prove that two mixed states ρ_1, ρ_2 on the Hilbert space H satisfy $\overline{Ran(\rho_1)} \perp \overline{Ran(\rho_2)}$ iff there exists an orthogonal projector $P \in \mathfrak{P}(\mathsf{H})$ with $tr(\rho_1 P) = 1$, $tr(\rho_2 P) = 0$.

Solution. If $\overline{Ran(\rho_1)} \perp \overline{Ran(\rho_2)}$, the orthogonal projector onto $\overline{Ran(\rho_1)}$ solves the problem. Conversely, if $tr(\rho_1 P) = 1$ and $tr(\rho_2 P) = 0$ for some $P \in \mathfrak{P}(\mathsf{H})$, let $P' := I - P$. Then $1 = tr(\rho_1) = tr(P\rho_1 P) + tr(P'\rho_1 P') + tr(P'\rho_1 P) + tr(P\rho_1 P')$. But $tr(P\rho_1 P) = tr(\rho_1 P) = 1$, $tr(P'\rho_1 P) = tr(\rho_1 PP') = 0$, $tr(P\rho_1 P') = tr(\rho_1 P'P) = 0$, and therefore $tr(P'\rho_1 P') = 0$. Since $P'\rho_1 P'$ is positive, self-adjoint and of trace class, and the trace equals the sum of the eigenvalues, the latter all vanish. By the spectral decomposition Theorem 4.18 we have $P'\rho_1 P' = 0$, so $\rho_1 = P\rho_1 P + P'\rho_1 P + P\rho_1 P'$. From this identity we easily see that $P\rho_1 P' \neq 0$ implies $(x + ay|\rho_1(x + ay)) < 0$ for some $x \in P(\mathsf{H})$, $y \in P'(\mathsf{H})$, with $a \in \mathbb{R}$ or $a \in i\mathbb{R}$ of large enough modulus. Hence $P\rho_1 P' = P'\rho_1 P = 0$ and $\rho_1 = P\rho_1(P + P') = P\rho_1$, and then $Ran(\rho_1) \subset P(\mathsf{H})$. A similar reasoning gives $P'\rho_2 P = P\rho_2 P' = 0$, whence $\rho_2 = P'\rho_2(P + P') = P'\rho_2$. This implies $Ran(\rho_2) \subset P'(\mathsf{H})$, and therefore $\overline{Ran(\rho_1)} \perp \overline{Ran(\rho_2)}$.

This page intentionally left blank

8

Spectral Theory I: generalities, abstract C^*-algebras and operators in $\mathfrak{B}(\mathsf{H})$

A mathematician plays a game and invents the rules. A physicist plays a game whose rules are dictated by Nature. As time goes by it is more and more evident that the rules the mathematician finds appealing are precisely those Nature has chosen.

P.A.M. Dirac

In this purely mathematically-flavoured chapter we introduce the basic spectral theory for operators on normed spaces, up to the notion of spectral measure and the spectral decomposition theorem for normal operators in $\mathfrak{B}(\mathsf{H})$, with H a Hilbert space. (The spectral theorem for *unbounded* self-adjoint operators will be discussed in the next chapter.) Here we present a number of general result in the abstract theory of C^*-algebras and *-homomorphisms.

The first part is devoted to the *resolvent set* and *spectrum* of an operator, and more generally of an element in a Banach algebra with unit, of which we will study general properties. Given a normal element in a C^*-algebra with unit, possibly a normal operator when the C^*-algebra is a concrete algebra of bounded operators over a Hilbert space, we shall prove there exists a *-homomorphism mapping algebra elements, i.e. operators, to continuous maps defined on a compact set of $\mathbb{C}$ (the spectrum of the element). In case we are dealing with operators we will show that such *-homomorphism extends to the C^*-algebra of bounded measurable functions defined on the compact set.

The spectrum of an operator is a collection of complex numbers that generalise eigenvalues. The spectral theorem, proved afterwards, warrants any operator – in this chapter always bounded and normal – a decomposition by integrating the spectrum with respect to a suitable "projector-valued" measure. Altogether the theorem may be viewed as a generalisation to Hilbert spaces of the diagonalisation of complex-valued normal matrices. The necessary mathematical tools to establish the spectral theorem are useful also for other reasons. Through them, namely, we will be able to define "operators depending on operators", a notion with several applications in mathematical physics.

The relationship between spectral theory and Quantum Mechanics lies in the fact that projector-valued measures are nothing but the observables defined in the previous chapter. Via the spectral theorem observables are in one-to-one correspondence to self-adjoint operators (typically unbounded), and the latter's spectra are the sets of possible measurements of observables. The correspondence observables/self-adjoint

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_8,

operators will allow us to develop the quantum-theory formulation in tight connection to classical mechanics, for which observables are the physical quantities represented by real functions.

Let us present the contents in better detail.

In section one we define the notions of *spectrum*, *resolvent set* and *resolvent operator*, establish their main properties and in particular discuss the *spectral radius* formula. All this will be generalised to abstract Banach or C^*-algebras, thus including the proof of the *Gelfand-Mazur theorem*, and a brief overview of the major features of C^*-algebra representations. We state the important *Gelfand–Najmark theorem*, that establishes any C^*-algebra with unit is a concrete C^*-algebra of operators on a Hilbert space. The proof will be given in Chapter 14 after the GNS theorem has been discussed.

In section two we construct continuous *-homomorphisms of C^*-algebras of functions induced either by normal elements in an abstract C^*-algebra, or by bounded self-adjoint operators on Hilbert spaces. Such homomorphisms represent the primary tool towards the spectral theorem. We discuss also the general properties of *-homomorphisms of C^*-algebras with unit and positive elements of C^*-algebras. Then we introduce the *Gelfand transform* to study commutative C^*-algebras with unit, proving the *commutative Gelfand–Najmark theorem*.

In the fourth section we introduce *spectral measures*, also known as *projector-valued measures* (PVMs), and define the integral of a bounded function with respect to a projector-valued measure.

The statement and proof of the spectral theorem for normal bounded operators (in particular self-adjoint or unitary) and some related technical facts are dealt with in section five.

The final section is devoted to *Fuglede's theorem* and some consequences.

8.1 Spectrum, resolvent set and resolvent operator

In this section we study the structural notions and results of spectral theory in normed, Banach and Hilbert spaces, but also in the more general context of Banach and C^*-algebras.

We shall make use of *analytic functions* defined on open subsets of $\mathbb{C}$ with values in a complex Banach space [Rud82], rather than in $\mathbb{C}$.

Definition 8.1. *Let* $(\mathsf{X}, || \, ||)$ *be a Banach space over* $\mathbb{C}$ *and* $\Omega \subset \mathbb{C}$ *a non-empty open set. A function* $f : \Omega \to \mathsf{X}$ *is called* **analytic** *if for any* $z_0 \in \Omega$ *there exists* $\delta > 0$ *such that*

$$f(z) = \sum_{n=0}^{+\infty} (z - z_0)^n a_n \quad \textit{for any } z \in B_\delta(z_0),$$

where $B_\delta(z_0) \subset \Omega$, $a_n \in \mathsf{X}$ *for any* $n \in \mathbb{N}$ *and the series converges in norm* $|| \, ||$.

The theory of analytic functions in Banach spaces is essentially the same as that of complex-valued analytic functions, which we take for granted; the only difference is that on the range the Banach norm replaces the modulus of complex numbers. With this proviso, definitions, theorems and proofs are the same as in the holomorphic case.

8.1.1 Basic notions in normed spaces

We begin with operators on normed spaces, and recall that if X is a vector space, "A is a operator on X" (Definition 5.1) means $A : D(A) \to \mathsf{X}$, where the domain $D(A) \subset \mathsf{X}$ is a subspace, usually not closed, in X.

Definition 8.2. *Let A be an operator on the normed space X.*

(a) *One calls* **resolvent set** *of A the set $\rho(A)$ of $\lambda \in \mathbb{C}$ such that:*
(i) $\overline{Ran(A-\lambda I)} = \mathsf{X}$;
(ii) $(A-\lambda I) : D(A) \to \mathsf{X}$ is injective;
(iii) $(A-\lambda I)^{-1} : Ran(A-\lambda I) \to \mathsf{X}$ is bounded.
(b) *If $\lambda \in \rho(A)$, the* **resolvent** *of A is the operator*

$$R_\lambda(A) := (A-\lambda I)^{-1} : Ran(A-\lambda I) \to D(A)\,.$$

(c) *The* **spectrum** *of A is the set $\sigma(A) := \mathbb{C} \setminus \rho(A)$.*

The spectrum of A is the disjoint union of the following three subsets:

(i) *the* **point spectrum** *of A, $\sigma_p(A)$, made by complex numbers λ for which $A-\lambda I$ is not injective;*
(ii) the **continuous spectrum** *of A, $\sigma_c(A)$, made by complex numbers λ such that $A-\lambda I$ is injective and $\overline{Ran(A-\lambda I)} = \mathsf{X}$, but $(A-\lambda I)^{-1}$ is not bounded;*
(iii) *the* **residual spectrum** *of A, $\sigma_r(A)$, made by complex numbers λ for which $A-\lambda I$ is injective, but $\overline{Ran(A-\lambda I)} \neq \mathsf{X}$.*

Remark 8.3. (1) It is clear that $\sigma_p(A)$ consists precisely of the eigenvalues of A (see Definition 3.53). In case $\mathsf{X} = \mathsf{H}$ is a Hilbert space and the eigenvectors of A form a basis in H one says A has a **purely point spectrum**. This does not mean, generally speaking, that $\sigma_p(T) = \sigma(T)$; for example compact self-adjoint operators have purely point spectrum, but 0 can still belong to the continuous spectrum.
(2) There exist other decompositions of the spectrum [ReSi80, AbCi97] in the case that $\mathsf{X} = \mathsf{H}$ is a Hilbert space and A is normal in $\mathfrak{B}(\mathsf{H})$, or self-adjoint in H. We shall consider some alternative splittings in the following chapter, after proving the spectral theorem for unbounded self-adjoint operators. An in-depth analysis of these classifications, relative to important operators in QM, can be found in [ReSi80, AbCi97]. ■

We start making a few precise assumptions, like taking X a Banach space and working with closed operators. In particular the next result holds if $T \in \mathfrak{B}(\mathsf{X})$ or, in a Hilbert space $\mathsf{X} = \mathsf{H}$, if $T : D(\mathsf{H}) \to \mathsf{H}$ is self-adjoint or an adjoint operator in H, bearing in mind Theorem 5.10.

Theorem 8.4. *Let T be a closed operator in the Banach space $\mathsf{X} \neq \{0\}$. Then:*

(a) $\lambda \in \rho(T) \;\Leftrightarrow\;$ *$T-\lambda I$ is a bijection from $D(T)$ to X.*

(b) (i) *$\rho(T)$ is open;*
(ii) *$\sigma(T)$ is closed;*
(iii) *if $\rho(T) \neq \varnothing$, the map $\rho(T) \ni \lambda \mapsto R_\lambda(T) \in \mathfrak{B}(\mathsf{X})$ is analytic.*

(c) *If $D(T) = \mathsf{X}$ (hence $T \in \mathfrak{B}(\mathsf{X})$):*

(i) *$\rho(T) \neq \varnothing$;*
(ii) *$\sigma(T) \neq \varnothing$ and is compact;*
(iii) *$|\lambda| \leq ||T||$ for any $\lambda \in \sigma(T)$.*

(d) *For any $\lambda, \mu \in \rho(T)$ the* **resolvent identity** *holds:*

$$R_\lambda(T) - R_\mu(T) = (\lambda - \mu)R_\lambda(T)R_\mu(T)\,.$$

Remark 8.5. (1) A comment on (c): if X is a Banach space and $D(T) = \mathsf{X}$, then $T : D(T) \to \mathsf{X}$ is closed iff $T \in \mathfrak{B}(\mathsf{X})$, by the closed graph Theorem 2.95.
(2) Part (a) is technically rather useful for deciding whether $\lambda \in \rho(T)$: it is not necessary to consider the topology, i.e. the density of $Ran(T - \lambda I)$ and the boundedness of $(T - \lambda I)^{-1}$; for that it is enough to check $T - \lambda I : D(T) \to \mathsf{X}$ is bijective, a set-theoretical property. ■

Proof of Theorem 8.4. (a) If $\lambda \in \rho(T)$, it suffices to show $Ran(T - \lambda I) = \mathsf{X}$. Since $(T - \lambda I)^{-1}$ is continuous, there exists $K \geq 0$ such that $||(T - \lambda I)^{-1}x|| \leq K||x||$ for any $x = (T - \lambda I)y \in Ran(T - \lambda I)$. Consequently, for any $y \in D(T)$:

$$||y|| \leq K||(T - \lambda I)y||\,. \tag{8.1}$$

Because $\overline{Ran(T - \lambda I)} = \mathsf{X}$, if $x \in \mathsf{X}$ there is a sequence $\{y_n\}_{n\in\mathbb{N}} \subset D(T)$ for which $(T - \lambda I)y_n \to x$, as $n \to +\infty$. From (8.1) we conclude $\{y_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence and so it admits a limit $y \in \mathsf{X}$. T being a closed operator, $y \in D(T)$ and $(T - \lambda I)y = x$, hence $x \in Ran(T - \lambda I)$. Thus $Ran(T - \lambda I) = \mathsf{X}$, as claimed.

Suppose now $T - \lambda I$ is a bijection from $D(T)$ to X; to prove the claim we need to show $(T - \lambda I)^{-1}$ is continuous. Since T is closed, then also $T - \lambda I$ is closed, i.e. its graph is closed. But $T - \lambda I$ is a bijection, so $(T - \lambda I)^{-1}$ has a closed graph and is then closed. Being $(T - \lambda I)^{-1}$ defined on X by assumption, Theorem 2.95 implies $(T - \lambda I)^{-1}$ is bounded.
(b) If $\mu \in \rho(T)$, the series

$$S(\lambda) := \sum_{n=0}^{+\infty} (\lambda - \mu)^n R_\mu(T)^{n+1}$$

converges absolutely in operator norm (hence in the uniform topology) provided

$$|\lambda - \mu| < 1/||R_\mu(T)||\,. \tag{8.2}$$

In fact,

$$\sum_{n=0}^{+\infty} |\lambda - \mu|^n ||R_\mu(T)^{n+1}|| \leq \sum_{n=0}^{+\infty} |\lambda - \mu|^n ||R_\mu(T)||^{n+1}$$
$$= ||R_\mu(T)|| \sum_{n=0}^{+\infty} |\,(\lambda - \mu)\,||R_\mu(T)||\,|^n\,.$$

The last series is geometric of reason $|(\lambda-\mu)\,||R_\mu(T)||\,|$, and converges because $|(\lambda-\mu)\,||R_\mu(T)||\,|<1$ by (8.2).

If λ satisfies the above condition, applying $T-\lambda I=(T-\mu I)+(\mu-\lambda)I$ to the left and right of the series of $S(\lambda)$ gives (again using the definition $R_\mu(T)^0:=I$):

$$(T-\lambda I)S(\lambda)=I_{\mathsf{X}}$$

while:

$$S(\lambda)(T-\lambda I)=I_{D(T)}\,.$$

Hence if $\mu\in\rho(T)$ there is an open neighbourhood of μ such that, for any λ in that neighbourhood, the left and right inverses of $T-\lambda I$, from X to $D(T)$, exist and are finite. By (a) then, the neighbourhood is contained in $\rho(T)$, and so $\rho(T)$ is open and $\sigma(T)=\mathbb{C}\setminus\rho(T)$ closed. Moreover $R_\lambda(T)$ has a Taylor series around any point of $\rho(T)$ in uniform topology, so by definition $\rho(T)\ni\lambda\mapsto R_\lambda(T)$ is analytic and maps $\rho(T)$ to the Banach space $\mathfrak{B}(\mathsf{X})$.
(c) In case $D(T)=\mathsf{X}$, since T is closed and X Banach, the closed graph theorem makes T bounded. If $\lambda\in\mathbb{C}$ satisfies $|\lambda|>||T||$, the series

$$S(\lambda)=\sum_{n=0}^{+\infty}(-\lambda)^{-(n+1)}T^n$$

($T^0:=I$), converges absolutely in operator norm. A direct computation, as before, gives the identities

$$(T-\lambda I)S(\lambda)=I$$

and

$$S(\lambda)(T-\lambda I)=I\,,$$

hence $S(\lambda)=R_\lambda(T)$ by (a). Hence by (a) every $\lambda\in\mathbb{C}$ with $|\lambda|>||T||$ belongs to $\rho(T)$, which is thus non-empty. Furthermore, if $\lambda\in\sigma(T)$, $|\lambda|\le||T||$, and $\sigma(T)$ is compact if non-empty as closed and bounded. Let us show $\sigma(T)\neq\varnothing$. By contradiction assume $\sigma(T)=\varnothing$. Then $\lambda\mapsto R_\lambda(T)$ is defined on $\mathbb{C}$. Fix $f\in\mathsf{X}'$ (dual to X) and $x\in\mathsf{X}$, and consider the complex-valued function $\rho(T)\ni\lambda\mapsto g(\lambda):=f(R_\lambda(T)x)$. It is certainly analytic on $\mathbb{C}$, because if $\mu\in\rho(T)$, on a neighbourhood of μ contained in $\rho(T)$ we have a Taylor expansion

$$f(R_\lambda(T)x):=\sum_{n=0}^{+\infty}(\lambda-\mu)^n f(R_\mu(T)^{n+1}x)\,.$$

We have used the continuity of the linear functional f, and the fact the series converges uniformly (and so weakly). Hence assuming $\sigma(T)=\varnothing$, g is analytic on $\mathbb{C}$. We notice that for $|\lambda|>||T||$ we have

$$g(\lambda):=f(R_\lambda(T)x)=\sum_{n=0}^{+\infty}(-\lambda)^{-(n+1)}f(T^nx)\,.$$

This series converges absolutely (Abel's theorem on power series), so we can wirte, for $|\lambda| \geq 1 + ||T||$:

$$|g(\lambda)| \leq \frac{1}{|\lambda|}\sum_{n=0}^{+\infty} \frac{|f(T^n x)|}{|\lambda|^n} \leq \frac{||f||||x||}{|\lambda|}\sum_{n=0}^{+\infty}\left(\frac{||T||}{|\lambda|}\right)^n = \frac{||f||||x||}{|\lambda|}\frac{|\lambda|}{|\lambda|-||T||} \leq \frac{K}{|\lambda|}$$

with $K > 0$. Thus $|g|$, everywhere continuous and bounded from above by $K|\lambda^{-1}|$, when $|\lambda| \geq \Lambda$ for some constant Λ, must be bounded on the *entire complex plane*. Being analytic on $\mathbb{C}$, g is constant by Liouville's theorem. As $|g(\lambda)|$ vanishes at infinity, g is the null map. Then $f(R_\lambda(T)x) = 0$. But the result holds for any $f \in \mathsf{X}'$, so Corollary 2.52 to Hahn–Banach (where $\mathsf{X} \neq \{0\}$), implies $||R_\lambda(T)x|| = 0$. As $x \in \mathsf{X} \neq \{0\}$ was arbitrary, we have to conclude $R_\lambda(T) = 0$ for any $\lambda \in \rho(T)$. Therefore $R_\lambda(T)$ cannot invert $T - \lambda I$, and the contradiction disproves the assumption $\sigma(T) = \varnothing$.
(d) The resolvent identity is proved as follows. First, we have

$$(T-\lambda I)R_\lambda(T) = I \quad \text{and} \quad (T-\mu I)R_\mu(T) = I\,.$$

Consider the products $TR_\lambda(T) - \lambda R_\lambda(T) = I_\mathsf{X}$ and $TR_\mu(T) - \mu R_\mu(T) = I_\mathsf{X}$, multiply the first by $R_\mu(T)$ on the left and the second by $R_\lambda(T)$ on the right, and then substract them. Recalling $R_\mu(T)R_\lambda(T) = R_\lambda(T)R_\mu(T)$ and $R_\mu(T)TR_\lambda(T) = R_\lambda(T)TR_\mu(T)$, we obtain the resolvent equation. The first commutation relation used above follows from the evident

$$(T-\mu I)(T-\lambda I) = (T-\lambda I)(T-\mu I),$$

which also gives a similar equation for inverses. The other relationship is explained as follows:

$$\begin{aligned} R_\mu(T)TR_\lambda(T) &= R_\mu(T)(T-\lambda I)R_\lambda(T) + R_\mu(T)\lambda I R_\lambda(T) \\ &= R_\mu(T)I + \lambda R_\mu(T)R_\lambda(T) = R_\mu(T) + \lambda R_\lambda(T)R_\mu(T) \\ &= (I + \lambda R_\lambda(T))R_\mu(T) = (R_\lambda(T)(T-\lambda I) + \lambda R_\lambda(T))R_\mu(T) \\ &= R_\lambda(T)TR_\mu(T)\,. \end{aligned}$$

This ends the proof. □

A useful corollary is worth citing that descends immediately from the resolvent identity and (a), (b) of Proposition 4.9.

Corollary 8.6. *Let $T : D(T) \to \mathsf{X}$ be a closed operator on the Banach space X. If for one $\mu \in \rho(T)$ the resolvent $R_\mu(T)$ is compact, then $R_\lambda(T)$ is compact for any $\lambda \in \rho(T)$.*

8.1.2 The spectrum of special classes of normal operators in Hilbert spaces

Let us focus on unitary operators and self-adjoint operators in Hilbert spaces, and discuss the structure of their spectrum. Using Definition 8.2 we work in full generality and consider unbounded operators with non-maximal domains.

Proposition 8.7. *Let* H *be a Hilbert space.*

(a) *If A is self-adjoint in* H *(not necessarily bounded, nor defined on the whole* H *in general):*

(i) $\sigma(A) \subset \mathbb{R}$;
(ii) $\sigma_r(A) = \varnothing$;
(iii) *the eigenspaces of A with distinct eigenvalues (points in $\sigma_p(A)$) are orthogonal*[1].

(b) *If $U \in \mathfrak{B}(\mathsf{H})$ is unitary:*

(i) $\sigma(U)$ *is a non-empty compact subset of* $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$;
(ii) $\sigma_r(U) = \varnothing$.

(c) *If $T \in \mathfrak{B}(\mathsf{H})$ is normal:*

(i) $\sigma_r(T) = \sigma_r(T^*) = \varnothing$;
(ii) $\sigma_p(T^*) = \overline{\sigma_p(T)}$;
(iii) $\sigma_c(T^*) = \overline{\sigma_c(T)}$, *where the bar denotes complex conjugation.*

Proof. (a) Let us begin with (i). Suppose $\lambda = \mu + i\nu$, $\nu \neq 0$ and let us prove $\lambda \in \rho(A)$. If $x \in D(A)$,

$$((A - \lambda I)x|(A - \lambda I)x) = ((A - \mu I)x|(A - \mu I)x) + \nu^2(x|x) + i\nu[(Ax|x) - (x|Ax)] .$$

The last summand vanishes for A is self-adjoint. Hence

$$||(A - \lambda I)x|| \geq |\nu| \, ||x|| .$$

With a similar arguement we obtain

$$||(A - \overline{\lambda} I)x|| \geq |\nu| \, ||x|| .$$

The operators $A - \lambda I$ and $A - \overline{\lambda} I$ are then one-to-one, and $||(A - \lambda I)^{-1}|| \leq |\nu|^{-1}$, where $(A - \lambda I)^{-1} : Ran(A - \lambda I) \to D(A)$. Notice

$$\overline{Ran(A - \lambda I)}^{\perp} = [Ran(A - \lambda I)]^{\perp} = Ker(A^* - \overline{\lambda} I) = Ker(A - \overline{\lambda} I) = \{0\} ,$$

where the last equality makes use of the injectivity of $A - \overline{\lambda} I$. Summarising: $A - \lambda I$ in injective, $(A - \lambda I)^{-1}$ bounded and $\overline{Ran(A - \lambda I)}^{\perp} = \{0\}$, i.e. $Ran(A - \lambda I)$ is dense in H; therefore $\lambda \in \rho(A)$, by definition of resolvent set.

Now to (ii) Suppose $\lambda \in \sigma(A)$, but $\lambda \notin \sigma_p(A)$. Then $A - \lambda I$ must be one-to-one and $Ker(A - \lambda I) = \{0\}$. Since $A = A^*$ and $\lambda \in \mathbb{R}$ by (i), we have $Ker(A^* - \overline{\lambda} I) = \{0\}$, so $[Ran(A - \lambda I)]^{\perp} = Ker(A^* - \overline{\lambda} I) = \{0\}$ and $\overline{Ran(A - \lambda I)} = \mathsf{H}$. Consequently $\lambda \in \sigma_c(A)$.

Proving (iii) is easy: if $\lambda \neq \mu$ and $Au = \lambda u$, $Av = \mu v$, then

$$(\lambda - \mu)(u|v) = (Au|v) - (u|Av) = (u|Av) - (u|Av) = 0 ;$$

from $\lambda, \mu \in \mathbb{R}$ and $A = A^*$. But $\lambda - \mu \neq 0$, so $(u|v) = 0$.

[1] The analogous property for normal operators (hence unitary or self-adjoint too) in $\mathfrak{B}(\mathsf{H})$ is contained in Proposition 3.54(b).

(b) (i) The closure of $\sigma(U)$ is a consequence of Theorem 8.4(b), because any unitary operator is defined on H, bounded and so closed. As $||U|| = 1$, part (c) of that theorem implies $\sigma(U)$ is a compact non-empty subset in $\{\lambda \in \mathbb{C} \mid |\lambda| \le 1\}$. To finish, consider the series

$$S(\lambda) = \sum_{n=0}^{+\infty} \lambda^n (U^*)^{n+1}$$

with $|\lambda| < 1$. Since $||U|| = ||U^*|| = 1$, the series converges absolutely in operator norm, so it defines an operator in $\mathfrak{B}(\mathsf{H})$. Because $U^*U = UU^* = I$,

$$(U - \lambda I)S(\lambda) = S(\lambda)(U - \lambda I) = I\,.$$

By Theorem 8.4(a) $\lambda \in \rho(U)$. To sum up: $\sigma(U)$ is compact and non-empty inside $\{\lambda \in \mathbb{C} \mid |\lambda| = 1\}$.
(ii) This follows from part (i) of (c), for every unitary operator is normal.
(c) Recall that $T \in \mathfrak{B}(\mathsf{H})$ normal implies $\lambda \in \mathbb{C}$ is an eigenvalue iff $\overline{\lambda}$ is an eigenvalue of T^* ((i) in Proposition 3.54(b)). This is enough to give (ii). The three parts of the spectrum are disjoint, and $\sigma(T) = \overline{\sigma(T^*)}$ (by Proposition 8.14(b)), so to prove (iii) it is enough to show (i). Assume $\lambda \in \sigma(T)$, but $\lambda \notin \sigma_p(T)$. Since $\sigma(T) = \overline{\sigma(T^*)}$ and $\sigma_p(T) = \overline{\sigma_p(T^*)}$, the hypothesis is equivalent to $\overline{\lambda} \in \sigma(T^*)$, but $\overline{\lambda} \notin \sigma_p(T^*)$. Then $T^* - \overline{\lambda} I$ must be one-to-one and $Ker(T^* - \overline{\lambda} I) = \{0\}$. Now Proposition 3.38(d) tells $[Ran(T - \lambda I)]^\perp = Ker(T^* - \overline{\lambda} I) = \{0\}$, hence $\overline{Ran(T - \lambda I)} = \mathsf{H}$ (here the bar denotes the closure). Therefore $\lambda \in \sigma_c(T)$, i.e. $\sigma_r(T) = \varnothing$. The proof for T^* is the same, because $(T^*)^* = T$ (Proposition 3.38(b)). □

8.1.3 Abstract C^*-algebras: Gelfand-Mazur theorem, spectral radius, Gelfand's formula, Gelfand–Najmark theorem

Now we consider, more abstractly, Banach algebras with unit and C^*-algebras with unit (Definitions 2.23 and 3.40). Recall that $\mathfrak{B}(\mathsf{X})$ is a Banach algebra with unit if X is normed, by (i) in Theorem 2.41(c). If H is a Hilbert space, $\mathfrak{B}(\mathsf{H})$ is a C^*-algebra with unit, whose involution is the Hermitian conjugation by Theorem 3.45.

First of all we generalise the notions of resolvent set and spectrum to an abstract setting, using $\mathfrak{B}(\mathsf{X})$ as model, with X Banach, so we have Theorem 8.4 in action. Recall that in an algebra $\mathfrak{A}$ with unit $\mathbb{I}$ the inverse a^{-1} to $a \in \mathfrak{A}$ is defined as the unique element, if present, such that $a^{-1}a = aa^{-1} = \mathbb{I}$.

Definition 8.8. *Let $\mathfrak{A}$ be a Banach algebra with unit $\mathbb{I}$ and take $a \in \mathfrak{A}$.*

(a) *The* **resolvent set** *of a is the set:*

$$\rho(a) := \{\lambda \in \mathbb{C} \mid \exists (a - \lambda\mathbb{I})^{-1} \in \mathfrak{A}\}\,.$$

(b) *The* **spectrum** *of a is the set $\sigma(a) := \mathbb{C} \setminus \rho(a)$.*

The following fact generalises the assertion in Theorem 8.4 about operators of $\mathfrak{B}(\mathsf{X})$.

Theorem 8.9. *Let $\mathfrak{A} \neq \{0\}$ be a Banach algebra with unit $\mathbb{I}$, $a \in \mathfrak{A}$ an arbitrary element.*

(a) *$\rho(a) \neq \varnothing$ is open, $\sigma(a) \neq \varnothing$ is compact and:*

$$|\lambda| \leq ||a||\,, \text{ for any } \lambda \in \sigma(a).$$

(b) *The map $\rho(a) \ni \lambda \mapsto R_\lambda(a) := (a - \lambda\mathbb{I})^{-1} \in \mathfrak{A}$ is analytic.*
(c) *If $\lambda, \mu \in \rho(a)$ the* **resolvent identity** *holds:*

$$R_\lambda(a) - R_\mu(a) = (\lambda - \mu) R_\lambda(a) R_\mu(a)\,.$$

Proof. The argument is the same of properties (b), (c), (d) of Theorem 8.4, because of Remark 8.5(1) and replacing, in the proof of (c), $f(R_\lambda(T)x)$ by $f(R_\lambda(a))$, where $f \in \mathfrak{A}'$ (Banach dual to $\mathfrak{A}$). □

A straightforward, yet important corollary is known as *Gelfand-Mazur theorem*, and tells that every normed division algebra over $\mathbb{C}$ is isomorphic to $\mathbb{C}$, so in particular is commutative.

Theorem 8.10 (Gelfand-Mazur). *A complex Banach algebra $\mathfrak{B} \neq \{0\}$ with unit, in which every non-zero element is invertible, is naturally isomorphic to $\mathbb{C}$. (In particular $\mathfrak{B}$ is commutative.)*

Proof. Take $x \in \mathfrak{B}$, so $\sigma(x) \neq \varnothing$ by part (a) in the previous theorem. Then $x - c\mathbb{I}$ is not invertible for some $c \in \mathbb{C}$ by definition of spectrum. In our case $x - c\mathbb{I} = 0$, so $x = c\mathbb{I}$. c is completely determined by x, for $c\mathbb{I} \neq c'\mathbb{I}$ if $c \neq c'$. The map $\mathfrak{B} \ni x \mapsto c \in \mathbb{C}$ is a Banach algebra isomorphism, as is easy to see. □

According to Theorem 8.9(a), the spectrum of $a \in \mathfrak{A}$ is contained in the disc of radius $||a||$ centred at the origin of $\mathbb{C}$. Yet there might be a disc of smaller radius at the origin enclosing $\sigma(a)$. In this respect see the next definition.

Remarks 8.11. The assumption that the field is $\mathbb{C}$ is crucial. There exist Banach division algebras that are not commutative, like the algebra $\mathbb{H}$ of *quaternions* introduced in Example 3.44(4). The latter, though, is a real algebra. The only algebras with an associative product on $\mathbb{R}$ that are normed and without zero-divisors are $\mathbb{R}$, $\mathbb{C}$ and $\mathbb{H}$ up to isomorphisms, as Hurwitz established in 1898. Dropping associativity there remains only one other instance, the *Cayley numbers*, also knowns as *octonions*. ■

Definition 8.12. *Let $\mathfrak{A}$ be a Banach algebra with unit. The* **spectral radius** *of $a \in \mathfrak{A}$ is the non-negative real number*

$$r(a) := \sup\{|\lambda| \mid \lambda \in \sigma(a)\}\,.$$

This applies in particular when $\mathfrak{A} = \mathfrak{B}(\mathsf{X})$, X Banach space.

Remarks 8.13. Any element a in a Banach algebra with unit $\mathfrak{A}$ satisfies the elementary (yet fundamental) property:

$$0 \leq r(a) \leq ||a||\,, \tag{8.3}$$

immediately ensuing Theorem 8.9(a). ■

There is a known formula for the spectral radius, due to the mathematician I.Gelfand. We shall recover Gelfand's formula using a property of the spectrum of polynomials over $\mathfrak{A}$.

Proposition 8.14. *Let $\mathfrak{A}$ be an Banach algebra with unit $\mathbb{I}$, $a \in \mathfrak{A}$ and $p = p(z)$ a complex-valued polynomial in $z \in \mathbb{C}$.*

(a) *Let $p(a)$ be the element in $\mathfrak{A}$ obtained by formally substituting the element a to z in $p(z)$ and interpreting powers a^n in the obvious way ($a^0 := \mathbb{I}$); then*

$$\sigma(p(a)) = p(\sigma(a)) := \{p(\lambda) \mid \lambda \in \sigma(a)\} . \tag{8.4}$$

(This holds in particular for $\mathfrak{A} = \mathfrak{B}(\mathsf{X})$, X a Banach space.)
(b) *If $\mathfrak{A}$ is additionally a $*$-algebra, the spectrum of a^* satisfies*

$$\sigma(a^*) = \overline{\sigma(a)} := \{\overline{\lambda} \mid \lambda \in \sigma(a)\} . \tag{8.5}$$

(This holds in particular for $\mathfrak{A} = \mathfrak{B}(\mathsf{H})$ with H a Hilbert space.)

Proof. (a) If $\alpha_1, \ldots, \alpha_n$ denote the roots of a polynomial q (not necessarily distinct), $q(z) = c\prod_{i=1}^n (z - \alpha_i)$ for some complex number c. Hence $q(a) = c\prod_{i=1}^n (a - \alpha_i\mathbb{I})$. Let $\lambda \in \sigma(a)$, so $(a - \lambda\mathbb{I})$ *is not invertible* by definition; set $\mu := p(\lambda)$. Consider now the polynomial $q := p - \mu$. As $q(\lambda) = 0$, one factor in the above decomposition of q will be $(z - \lambda)$, and so choosing the root order appropriately, and recalling that the $a - \alpha_i\mathbb{I}$ commute, we have:

$$p(a) - \mu\mathbb{I} = c\left[\prod_{i=1}^{n-1}(a - \alpha_i\mathbb{I})\right](a - \lambda\mathbb{I}) = c(a - \lambda\mathbb{I})\prod_{i=1}^{n-1}(a - \alpha_i\mathbb{I}) .$$

Thus $p(a) - \mu\mathbb{I}$ cannot be invertible, for $a - \lambda\mathbb{I}$ is not. (Were $p(a) - \mu\mathbb{I}$ invertible, we whould have

$$\mathbb{I} = (a - \lambda\mathbb{I})\left[\left(\prod_{i=1}^{n-1} c(a - \alpha_i\mathbb{I})\right)(p(a) - \mu\mathbb{I})^{-1}\right] ,$$

$$\mathbb{I} = \left((p(a) - \mu\mathbb{I})^{-1}\prod_{i=1}^{n-1} c(a - \alpha_i\mathbb{I})\right)(a - \lambda\mathbb{I}) ,$$

implying $(a - \lambda\mathbb{I})$ invertible. Applying the first factor on the right in the second line to the first equation would say the right and left inverses of $(a - \lambda\mathbb{I})$ coincide:

$$(p(a) - \mu\mathbb{I})^{-1}\prod_{i=1}^{n-1} c(a - \alpha_i\mathbb{I}) = \left(\prod_{i=1}^{n-1} c(a - \alpha_i\mathbb{I})\right)(p(a) - \mu\mathbb{I})^{-1} ,$$

as it should be.) By definition we must have $\mu \in \sigma(p(a))$, hence we proved $p(\sigma(a)) \subset \sigma(p(a))$. Now we go for the other inclusion. Let $\mu \in \sigma(p(a))$, set $q = p - \mu$ and decompose q as $q(z) = c\prod_{i=1}^n (z - \alpha_i)$. Therefore

$$p(a) - \mu\mathbb{I} = c\prod_{i=1}^{n}(a - \alpha_i\mathbb{I})$$

as before. If all roots α_i belonged to $\rho(a)$ every $(a-\alpha_i\mathbb{I}) : \mathsf{X} \to \mathsf{X}$ would be invertible, so $p(a)-\mu\mathbb{I}$ would become invertible, which is excluded by assumption. Therefore there is a root α_k such that $(a-\alpha_k\mathbb{I})$ is not invertible, so $\alpha_k \in \sigma(a)$. But then $p(\alpha_k)-\mu = 0$, so $\mu \in p(\sigma(a))$, and hence $p(\sigma(a)) \supseteq \sigma(p(a))$.
(b) $(a-\lambda\mathbb{I})$ is invertible if and only if $(a-\lambda\mathbb{I})^* = a^* - \overline{\lambda}\mathbb{I}$ by Proposition 3.42(c), hence the claim. □

Theorem 8.15. *Let $\mathfrak{A}$ be a Banach algebra with unit and $a \in \mathfrak{A}$.*

(a) *The spectral radius of a can be computed with* **Gelfand's formula***:*

$$r(a) = \lim_{n\to+\infty} ||a^n||^{1/n},$$

where the limit always exists. (This holds in particular when $\mathfrak{A} = \mathfrak{B}(\mathsf{X})$ with X Banach space.)
(b) *If $\mathfrak{A}$ is a C^*-algebra with unit and a is normal ($a^*a = aa^*$), then*

$$r(a) = ||a||, \tag{8.6}$$

and consequently:

$$||a|| = r(a^*a)^{1/2} \quad \text{for any } a \in \mathfrak{A}. \tag{8.7}$$

(Valid in particular for $\mathfrak{A} = \mathfrak{B}(\mathsf{H})$, H Hilbert space.)

Proof. (a) By Proposition 8.14(a) $(\sigma(a))^n = \sigma(a^n)$, so $r(a)^n = r(a^n) \leq ||a^n||$, and then

$$r(a) \leq \liminf_n ||a^n||^{1/n}. \tag{8.8}$$

(In contrast to the limit infimum, which always exists, the limit might not.) If $|\lambda| > r(a)$,

$$R_\lambda(a) = \sum_{n=0}^{+\infty} (-\lambda)^{-(n+1)} a^n, \tag{8.9}$$

because a *theorem of Hadamard* guarantees that the convergence disc of the Laurent series of an analytic function touches the closest singularity to the point at infinity. In our case all singularities belong to the spectrum $\sigma(a)$, so the boundary consists of points $\lambda \in \mathbb{C}$ with $|\lambda| > r(a)$. Therefore the above series converges for any $\lambda \in \mathbb{C}$ such that $|\lambda| > r(a)$, hence it converges absolutely on any disc, centred at infinity, passing through such λ. In particular

$$|\lambda|^{-(n+1)} ||a^n|| \to 0,$$

as $n \to +\infty$, for any $\lambda \in \mathbb{C}$ with $|\lambda| > r(a)$. Hence for any $\varepsilon > 0$

$$||a^n||^{1/n} < \varepsilon^{1/n} |\lambda|^{(n+1)/n} = (\varepsilon|\lambda|)^{1/n} |\lambda|$$

definitely. Since $(\varepsilon|\lambda|)^{1/n} \to 1$ for $n \to +\infty$, we have $\limsup_n ||a^n||^{1/n} \leq |\lambda|$ for every $\lambda \in \mathbb{C}$ with $|\lambda| > r(a)$. We can get as close as we want to $r(a)$ with $|\lambda|$, so $\limsup_n ||a^n||^{1/n} \leq r(a)$. Finally, by (8.8),

$$r(a) \leq \liminf_n ||a^n||^{1/n} \leq \limsup_n ||a^n||^{1/n} \leq r(a).$$

This shows the limit of $||a^n||^{1/n}$ exists as $n \to +\infty$, and it concides with $r(a)$.
(b) By Proposition 3.42(a) we have $||a^n|| = ||a||^n$ if a is normal. Gelfand's formula gives

$$r(a) = \lim_{n\to+\infty} ||a^n||^{1/n} = \lim_{n\to+\infty} (||a||^n)^{1/n} = ||a|| .$$

Equation (8.7) follows from a property of C^*-algebras, i.e. $||a^*a|| = ||a||^2$ for any a, because a^*a is self-adjoint hence normal. □

Identity (8.7) explains that the norm of a given C^*-algebra is *uniquely determined* by algebraic properties, because the spectral radius is obtainable from the spectrum, and this is built by algebraic means entirely.

Corollary 8.16. *A *-algebra with unit $\mathfrak{A}$ admits one norm at most that makes it a C^*-algebra with unit.*

Notation 8.17. Let $\mathfrak{A}$ and $\mathfrak{A}_1$ be C^*-algebras with unit and take $a \in \mathfrak{A}_1 \cap \mathfrak{A}$. *A priori*, the element a could have two different spectra if thought of as element of $\mathfrak{A}_1$ or of $\mathfrak{A}$. That is why here, and in other similar situations where confusion might arise, we will label spectra: $\sigma_{\mathfrak{A}}(a)$ or $\sigma_{\mathfrak{A}_1}(a)$. ■

There is another important consequence of (8.7) concerning algebra homomorphisms, to which we will return later with a general theorem. Remarkably enough, *-homomorphisms of C^*-algebras with unit are continuous. Subsequently we will see something stronger: if injective, namely, they are automatically isometric.

Corollary 8.18. *Let $\phi : \mathfrak{A} \to \mathfrak{B}$ be a *-homomorphism between C^*-algebras with unit.*

(a) *ϕ is continuous, for $||\phi(a)||_{\mathfrak{B}} \leq ||a||_{\mathfrak{A}}$ for any $a \in \mathfrak{A}$.*
(b) *For every $a \in \mathfrak{A}$, $\sigma_{\mathfrak{B}}(\phi(a)) \subset \sigma_{\mathfrak{A}}(a)$.*
(c) *If ϕ is additionally a *-isomorphism, it is also isometric: $||\phi(a)||_{\mathfrak{B}} = ||a||_{\mathfrak{A}}$ for any $a \in \mathfrak{A}$, and $\sigma_{\mathfrak{B}}(\phi(a)) = \sigma_{\mathfrak{A}}(a)$ for any $a \in \mathfrak{A}$.*

Proof. (a) If $\lambda \in \rho(a)$, $\lambda \in \rho(\phi(a))$ because ϕ is a *-homomorphism. Thus $\sigma(\phi(a)) \subset \sigma(a)$, and $r(\phi(a)) \leq r(a)$. Equation (8.7) implies $||\phi(a)||^2_{\mathfrak{B}} = r_{\mathfrak{B}}(\phi(a)^*\phi(a)) = r_{\mathfrak{B}}(\phi(a^*a)) \leq r_{\mathfrak{A}}(a^*a) = ||a||^2_{\mathfrak{A}}$.
(b) If a' exists such that $(a - \lambda\mathbb{I}_{\mathfrak{A}})a' = a'(a - \lambda\mathbb{I}_{\mathfrak{A}}) = \mathbb{I}_{\mathfrak{A}}$, appying the *-homomorphism ϕ we conclude $(\phi(a) - \lambda\mathbb{I}_{\mathfrak{B}})\phi(a') = \phi(a')(\phi(a) - \lambda\mathbb{I}_{\mathfrak{B}}) = \mathbb{I}_{\mathfrak{B}}$, so $\rho(\phi(a)) \subset \rho(a)$ and the claim follows.
(c) is obvious from (a) and (b), replicating the argument for ϕ^{-1}. □

To end the abstract considerations we are making, let us present the next result on C^*-algebras in relationship to the classification of Definition 3.40.

Proposition 8.19. *Let $\mathfrak{A}$ be a C^*-algebra with unit (in particular $\mathfrak{A}$ can be $\mathfrak{B}(\mathsf{H})$, H Hilbert space).*

(a) *If $a \in \mathfrak{A}$ admits a left inverse, $\sigma(a) = \sigma(a)^{-1} := \{\lambda^{-1} \,|\, \lambda \in \sigma(A)\}$.*
(b) *If $a \in \mathfrak{A}$ is isometric, i.e. $a^*a = \mathbb{I}$, then $r(a) = 1$.*
(c) *If $a \in \mathfrak{A}$ is unitary, i.e. $a^*a = aa^* = \mathbb{I}$, then $\sigma(a) \subset \mathbb{S}^1 \subset \mathbb{C}$.*

(d) *If $a \in \mathfrak{A}$ is self-adjoint, i.e. $a = a^*$, then $\sigma(a) \subset \mathbb{R}$. More precisely, $\sigma(a) \subset [-||a||, ||a||]$, and $\sigma(a^2) \subset [0, ||a||^2]$.*
(e) *If $a, b \in \mathfrak{A}$ then $\sigma(ab) \setminus \{0\} = \sigma(ba) \setminus \{0\}$.*

Proof. (a) If a is left-invertible, $0 \notin \sigma(a) \cup \sigma(a^{-1})$. If $\lambda \neq 0$ then

$$\lambda \mathbb{I} - a = \lambda a(a^{-1} - \lambda^{-1}\mathbb{I}) \quad \text{and} \quad \lambda^{-1}\mathbb{I} - a^{-1} = \lambda^{-1}a^{-1}(a - \lambda \mathbb{I}).$$

Thus $a - \lambda\mathbb{I}$ is invertible iff $a^{-1} - \lambda^{-1}\mathbb{I}$ is.
(b) If $a^*a = \mathbb{I}$ then $||a^n||^2 = ||(a^n)^*a^n|| = ||(a^*)^n a^n|| = ||\mathbb{I}|| = 1$. Gelfand's formula implies $r(a) = 1$.
(c) By (b) and the definition of spectral radius we infer $\sigma(a) \subset \{\lambda \in \mathbb{C} \,|\, |\lambda| \leq 1\}$. On the other hand we know from Proposition 8.14 that $\sigma(a) = \overline{\sigma(a^*)}$. As $a^* = a^{-1}$ and using part (a) we have $\sigma(a) = \overline{\sigma(a)}^{-1}$. Hence any element $\lambda \in \sigma(a)$ satsfies $|\lambda| \leq 1$ and can be written as $\lambda = \overline{\mu}^{-1}$, $|\mu| \leq 1$. This implies $|\lambda| = 1$.
(d) First of all we prove $\sigma(a) \subset \mathbb{R}$. Fix $\lambda \in \mathbb{R}$, $\lambda^{-1} > ||a||$, so that $|-i\lambda^{-1}| = \lambda^{-1} > r(a)$ and consequently $\mathbb{I} + i\lambda a = i\lambda(-i\lambda^{-1}\mathbb{I} + a)$ is invertible. Define $b := (\mathbb{I} - i\lambda a)(\mathbb{I} + i\lambda a)^{-1}$. Then $b^* = (\mathbb{I} - i\lambda a)^{-1}(\mathbb{I} + i\lambda a)$, and since the terms in brackets trivially commute,

$$b^*b = (\mathbb{I} - i\lambda a)^{-1}(\mathbb{I} + i\lambda a)(\mathbb{I} - i\lambda a)(\mathbb{I} + i\lambda a)^{-1} = \mathbb{I}.$$

A similar computation gives $bb^* = \mathbb{I}$, making b unitary. We may then invoke part (c), so that $\sigma(b) \subset \mathbb{S}^1$. Directly, $|(1 - i\lambda\mu)(1 + i\lambda\mu)^{-1}| = 1$ iff $\mu \in \mathbb{R}$. Therefore

$$z := (1 - i\lambda\mu)(1 + i\lambda\mu)^{-1}\mathbb{I} - b$$

is invertible when $\mu \in \mathbb{C} \setminus \mathbb{R}$. Solving the expression of b for a gives

$$z = 2i\lambda(\mathbb{I} + i\lambda\mu)^{-1}(a - \mu\mathbb{I})(\mathbb{I} + i\lambda a)^{-1},$$

hence $a - \mu\mathbb{I}$ is invertible for any $\mu \in \mathbb{C} \setminus \mathbb{R}$. It follows $\sigma(a) \subset \mathbb{R}$. But $r(a) = ||a||$, so $\sigma(a) \subset [-||a||, ||a||]$ is immediate by definition of spectral radius.
(d) follows from Proposition 8.14(a, b).
(e) If c is the inverse of $\mathbb{I} - ab$, then $(\mathbb{I} + bca)(\mathbb{I} - ba) = \mathbb{I} - ba + bc(\mathbb{I} - ab)a = \mathbb{I}$ and $(\mathbb{I} - ba)(\mathbb{I} + bca) = \mathbb{I} - ba + b(\mathbb{I} - ab)ca = \mathbb{I}$. Hence $\mathbb{I} + bca$ inverts $\mathbb{I} - ba$, implying (e). □

We could ask ourselves whether there exist C^*-algebras that cannot be realised as algebras of operators on Hilbert spaces. The answer is no, even if the identification between the C^*-algebra and a C^*-algebra of operators is not fixed uniquely. In fact, the following truly paramount result holds, which we shall prove as Theorem 14.23 after the *GNS theorem* (Chapter 14).

Theorem (Gelfand–Najmark). *If $\mathfrak{A}$ is a C^*-algebra with unit, there exist a Hilbert space H and an isometric *-isomorphism $\phi : \mathfrak{A} \to \mathfrak{B}$, where $\mathfrak{B} \subset \mathfrak{B}(\mathsf{H})$ is a C^*-subalgebra of $\mathfrak{B}(\mathsf{H})$.*

8.2 Functional calculus: representations of commutative C^*-algebras of bounded maps

This section aims to show how to represent an algebra of bounded measurable functions f by an algebra of *functions* $f(T,T^*)$ *of a bounded normal operator* T. We shall construct a continuous map

$$\widehat{\Phi}_T : \mathsf{M}_b(K) \ni f \mapsto f(T,T^*) \in \mathfrak{B}(\mathsf{H}),$$

perserving the structure of commutative C^*-algebra with unit, from bounded measurable functions defined on a compact set K, to bounded operators on a Hilbert space H (see Examples 2.26(4) and 3.44(1)). This will be a *representation* (Definition 3.48) of the commutative C^*-algebra with unit $\mathsf{M}_b(K)$ on H. It will be "generated" by a normal operator $T \in \mathfrak{B}(\mathsf{H})$ and $K = \sigma(T)$. Viewing $\widehat{\Phi}_T(f)$ as $f(T)$, when $T = T^*$, arises also from the physical interpretation related to the notion of observable, as we shall see. The theory we are talking about goes under the name of *functional calculus*. In a subsequent section we will show how the operator $f(T,T^*)$ can be understood as an integral of f with respect to an operator-valued measure. For the time being we shall construct $f(T,T^*)$ with no mention to spectral measures.

The first part of the construction involves only continuous maps f, and one speaks about *continuous functional calculus*. Continuous functional calculus overlooks the concrete C^*-algebra of bounded operators, and is valid more abstractly if we replace T by a normal element a in a given C^*-algebra. Therefore we shall work first in an abstract setting, and build first a continuous functional calculus for self-adjoint elements, and only after for normal elements in a general C^*-algebra with unit, owing to the *Gelfand transform*. Eventually, when dealing with measurable functions, we will return to operator algebras. By the way, continuous functional calculus touches upon *-homomorphisms of C^*-algebras, and allows to characterise positive elements of a C^*-algebra, as explained in a moment.

8.2.1 Abstract C^*-algebras: functional calculus for continuous maps and self-adjoint elements

Let us put ourselves in a general case where $\mathfrak{A}$ is a C^*-algebra with unit $\mathbb{I}$. We may think rather concretely that $\mathfrak{A} = \mathfrak{B}(\mathsf{H})$ for some Hilbert space H, although the following considerations transcend this case.

The first step to build the aforementioned *-homomorphisms is to study *polynomial functions* of a *self-adjoint element*: $a^* = a \in \mathfrak{A}$.

Define the function ϕ_a that maps a polynomial with complex coefficients $p = p(x)$, $x \in \mathbb{R}$, to the normal element $p(a)$ of $\mathfrak{A}$, in the obvious way: i.e., evaluating at a and interpreting the product in the algebra. Set also $a^0 := \mathbb{I}$.
ϕ_a has interesting features, of immediate proof:

(a) it is linear: $\phi_a(\alpha p + \beta p') = \alpha\phi_a(p) + \beta\phi_a(p')$ for any $\alpha, \beta \in \mathbb{C}$;
(b) it transforms products of polynomials into composite elements in the algebra: $\phi_a(p \cdot p') = \phi_a(p)\phi_a(p')$;

(c) it maps the constant polynomial 1 to the neutral element: $\phi_a(1) = \mathbb{I}$.

By Definition 2.23 these properties make ϕ_a a *homomorphism of algebras with unit*, from the commutative *-algebra with unit of complex polynomials to the C^*-algebra with unit $\mathfrak{A}$.

Here are other properties:

(d) ϕ_a maps the polynomial $\mathbb{R} \ni x \mapsto x$ (denoted x, inappropriately) to a, i.e. $\phi_a(x) = a$;
(e) if $\overline{p}$ is the conjugate polynomial to p ($\overline{p}(x) = \overline{p(x)}$, $x \in \mathbb{R}$), then $\phi_a(p)^* = \phi_a(\overline{p})$;
(f) if $ba = ab$ for some $b \in \mathfrak{A}$, $b\phi_a(p) = \phi_a(p)b$ for any polynomial p.

Property (e) establishes ϕ_a is a *-homomorphism (Definition 3.40) from the *-algebra with unit of polynomials to the C^*-algebra with unit $\mathfrak{A}$.

There is a further property if we restrict polynomials to the compact set $\sigma(a) \subset \mathbb{R}$. Since $\phi_a(p) = p(a)$ is self-adjoint and hence normal, by virtue of Theorem 8.15(b)

$$||p(a)|| = r(p(a)) = \sup\{|\mu| \mid \mu \in \sigma(p(a))\} .$$

The fact that $\sigma(p(a)) = p(\sigma(a))$ (Proposition 8.14(a)) implies

$$||\phi_a(p)|| = \sup\{|p(x)| \mid x \in \sigma(a)\} . \tag{8.10}$$

That is to say if the algebra of polynomials on $\sigma(a)$ is endowed with norm $||\ ||_\infty$, ϕ_a is an *isometry*. As we shall see, this fact can be generalised beyond polynomials.

Remarks 8.20. Assuming $\sigma(a)$ is *not* a *finite* set, with a minor reinterpretation of the symbols we denote, henceforth, by ϕ_a the map sending a function p to $p(a) \in \mathfrak{A}$, given by the restriction to $\sigma(a)$ of a polynomial p. Thus $||p||_\infty$ will for instance indicate the least upper bound of the absolute value of p over the compact set $\sigma(a)$. Properties (a)-(f) still hold, because a polynomial's restriction to an infinite set determines the polynomial: the difference of two polynomials (in $\mathbb{R}$ or $\mathbb{C}$, with complex coefficients) with infinite zero set is a polynomial (in $\mathbb{R}$ or $\mathbb{C}$, with complex coefficients) with infinitely many zeroes. Therefore it is the null polynomial.

In the case $\sigma(a)$ is *finite*, the matter is more delicate, because the restriction $q\restriction_{\sigma(a)}$ of a polynomial q does not determine the polynomial completely. However,

$$||q(a)|| = \sup\{|q(x)| \mid x \in \sigma(a)\}$$

implies immediately that if $q\restriction_{\sigma(a)} = q'\restriction_{\sigma(a)}$ for two q, q', then $q(a) = q'(a)$. Therefore everything we say will work for $\sigma(a)$ finite as well, even though we will not separate the cases $\sigma(a)$ finite vs infinite. ■

Recall that the space $C(\mathsf{X})$ of complex-valued continuous maps on a compact space X (cf. Examples 2.26(4), 3.44(1) in Chapter 2, 3), is a commutative C^*-algebra with unit: the norm is $||\ ||_\infty$, sum and product are the standard pointwise operations, the involution is the complex conjugation and the unit is the constant map 1.

Theorem 8.21 (Functional calculus for continuous maps and self-adjoint elements). *Let $\mathfrak{A}$ be a C^*-algebra with unit $\mathbb{I}$ and $a \in \mathfrak{A}$ a self-adjoint element.*

(a) *There exists a unique* **-homomorphism defined on the commutative* C^**-algebra with unit* $C(\sigma(a))$*:*

$$\Phi_a : C(\sigma(a)) \ni f \mapsto f(a) \in \mathfrak{A}\,,$$

such that

$$\Phi_a(x) = a, \tag{8.11}$$

x being the map $\sigma(a) \ni x \mapsto x$*.*
(b) *The following properties hold:*

(i) Φ_a *is isometric: for any* $f \in C(\sigma(a))$, $||\Phi_a(f)|| = ||f||_\infty$*;*
(ii) *if* $ba = ab$ *with* $b \in \mathfrak{A}$*, then* $bf(a) = f(a)b$ *for any* $f \in C(\sigma(a))$*;*
(iii) Φ_a *preserves involutions:* $\Phi_a(\overline{f}) = \Phi_a(f)^*$ *for any* $f \in C(\sigma(a))$*.*

(c) $\sigma(f(a)) = f(\sigma(a))$ *for any* $f \in C(\sigma(a))$*.*
(d) *If* $\mathfrak{B}$ *is a* C^**-algebra with unit and* $\pi : \mathfrak{A} \to \mathfrak{B}$ *a* **-homomorphism:*

$$\pi(f(a)) = f(\pi(a)) \quad \textit{for any } f \in C(\sigma_{\mathfrak{A}}(a)).$$

Proof. (a) In the sequel we assume the spectrum of a is infinite; the finite case must be treated separately by keeping in account the previous remark.

Let us show existence. The spectrum $\sigma(a) \subset \mathbb{C}$ is compact by Theorem 8.4(c), and $C(\sigma(a))$ Hausdorff because $\mathbb{C}$ is, so we can use Stone–Weierstrass (Theorem 2.27). The space $P(\sigma(a))$ of polynomials $p = p(x)$, $x \in \sigma(a)$, with complex coefficients is a subalgebra in $C(\sigma(a))$ that contains the unit (the function 1), separates points in $\sigma(a)$ and is closed under complex conjugation. Hence Theorem 2.27 guarantees it is dense in $C(\sigma(a))$. Consider the map

$$\phi_a : P(\sigma(a)) \ni p \mapsto p(a) \in \mathfrak{A}\,,$$

and refer to properties (a)–(f). We know ϕ_a is linear and $||\phi_a(p)|| = ||p||_\infty$ by (8.10), which implies continuity. By Proposition 2.44 there is a unique bounded linear operator $\Phi_a : C(\sigma(a)) \to \mathfrak{A}$ extending ϕ_a to $C(\sigma(a))$ maintaining the norm. This must be a homomorphism of algebras with unit because: (a) it is linear, (b) $\Phi_a(f \cdot g) = \Phi_a(f)\Phi_a(g)$ by continuity (it is true on the subalgebra of polynomials, by definition of ϕ_a), (c) it maps the constant function $1 \in P(\sigma(a))$ to the identity $\mathbb{I} \in \mathfrak{A}$, by definition of ϕ_a. Equation (8.11) holds trivially by property (d). That Φ_a is a *-homomorphism is due to this argument: if $\{p_n\}$ are polynomials uniformly converging on $\sigma(a)$ to the continuous f, $\{\overline{p_n}\}$ tends uniformly on $\sigma(a)$ to the continuous $\overline{f}$; as seen above, though (cf. property (e)), $\Phi_a(\overline{p_n}) = \phi_a(\overline{p_n}) = \phi_a(p_n)^* = \Phi_a(p_n)^*$ and Hermitian conjugation is continuous in the uniform topology. By continuity of Φ_a, $\Phi_a(\overline{f}) = \Phi_a(f)^*$.
Now to uniqueness. Any *-homomorphism χ_a of C^*-algebras with unit, fulfilling (8.11), must agree with Φ_a on integer powers of x, hence on any polynomials by definition of *-homomorphism. Moreover χ_a must be continuous by Corollary 8.18(a). Since χ_a and Φ_a are linear, by Proposition 2.44 χ_a coincides with Φ_a.
(b) Property (iii) was proved above. (i) and (ii) are immedate for polynomials, so they extend by continuity to $C(\sigma(a))$.

(c) Observe first that the set of non-invertible elements in $\mathfrak{A}$ is closed under the norm because its complement is open (Remark 2.24(2)). Consider a polynomial sequence $p_n \to f$ converging to some $f \in C(\sigma(a))$ uniformly on $\sigma(a)$. Then $p_n(\lambda) \in \sigma(p_n(a))$ by Proposition 8.14(a), i.e. $p_n(a) - p_n(\lambda)\mathbb{I}$ is not invertible. The set of non-invertible elements is closed in $\mathfrak{A}$, so we can take the limit and obtain $f(a) - f(\lambda)\mathbb{I}$ is not invertible. Hence $f(\lambda) \in \sigma(f(a))$ and then $f(\sigma(a)) \subset \sigma(f(a))$. Conversely, if $\mu \notin f(\sigma(a))$, then $g : \sigma(a) \ni \lambda \mapsto (f(\lambda) - \mu)^{-1}$ is in $C(\sigma(a))$. That is because f is continuous and $f(\sigma(a))$ closed (continuous image in $\mathbb{C}$ of a compact set). By construction $g(a)(f(a) - \mu\mathbb{I}) = (f(a) - \mu\mathbb{I})g(a) = \mathbb{I}$, so $f(a) - \mu\mathbb{I}$ is invertible, hence $\mu \notin \sigma(f(a))$.
(d) The statement is true if f is a polynomial. By continuity of π (Corollary 8.18(a)) it stays true when passing to continuous maps. □

What we would like to do now is generalise the above theorem to normal elements, not necessarily self-adjoint, in a C^*-algebra with unit $\mathfrak{A}$. We want to define an element $f(a, a^*) \in \mathfrak{A}$ for f an arbitrary continuous map defined on the spectrum $\sigma(a) \subset \mathbb{C}$ of a, so that its norm is $||f||_\infty$.

One possibility is to do as follows:

(1) start from polynomials $p(z, \overline{z})$ (dense in $C(\sigma(a))$ by the Stone–Weierstrass theorem) defined on the spectrum of a;
(2) associate to each $p(z, \overline{z})$ the operator polynomial $p(a, a^*) \in \mathfrak{A}$;
(3) show the above correspondence is a continuous *-homomorphism of *-algebras with unit.

Yet a problem arises when passing from polynomials to continuous maps by a limiting procedure. We should prove $||p(a, a^*)|| \leq ||p||_\infty$. In case a is self-adjoint the equality was proved using 'spectral invariance' (Proposition 8.14(a), i.e. $\sigma(p(a)) = p(\sigma(a))$) and Theorem 8.15(b), for which $||p(a)|| = r(p(a)) = \sup\{|\mu| \in \mathbb{C} \mid \mu \in \sigma(p(a))\}$. In the case at stake there is nothing guaranteeing $\sigma(p(a, a^*)) = p(\sigma(a, a^*))$. The failure of the fundamental theorem of algebra for complex polynomials in the variables z and $\overline{z}$ is the main cause of the lack of a direct proof of the above fact, and the reason why we have to look for an alternative, albeit very interesting, way.

8.2.2 Key properties of *-homomorphisms of C^*-algebras, spectra and positive elements

This section is devoted to a series of technical corollaries to Theorem 8.21, essential to extend continuous functional calculus to normal, not self-adjoint, elements. A number of results are nonetheless interesting on their own.

Corollary 8.18(c) tells a *-homomorphism $\phi : \mathfrak{A} \to \mathfrak{B}$ between C^*-algebras with unit is isometric if one-to-one and onto. But surjectivity is not necessary, for a consequence of the previous theorem is that injectivity is equivalent to norm preservation. We encapsulate in the next statement also Corollary 8.18(a), proved earlier.

Theorem 8.22 (On *-homomorphisms of C^*-algebras with unit). *A *-homomorphism $\pi : \mathfrak{A} \to \mathfrak{B}$ of C^*-algebras with unit is continuous, for*

$$||\pi(a)||_{\mathfrak{B}} \leq ||a||_{\mathfrak{A}} \quad \textit{for any } a \in \mathfrak{A}.$$

Furthermore

(a) *π is one-to-one iff isometric, i.e. $||\pi(a)|| = ||a||$ for any $a \in \mathfrak{A}$.*
(b) *$\pi(\mathfrak{A})$ is a C^*-subalgebra with unit inside $\mathfrak{B}$.*

Proof. As mentioned, the first statement is Corollary 8.18(a).
(a) If π is isometric it is obviously injective, so we prove the converse. We have $||\pi(a)|| \leq ||a||$ by Corollary 8.18, so it suffices to prove that injectivity forces $||\pi(a)|| \geq ||a||$. If that is true for self-adjoint elements in a C^*-algebra with unit, it holds for any element:

$$||\pi(a)||^2 = ||\pi(a)^*\pi(a)||^2 = ||\pi(a^*a)|| \geq ||a^*a|| = ||a||^2 .$$

So assume there is a self-adjoint $a \in \mathfrak{A}$ with $||\pi(a)|| < ||a||$. Then Proposition 8.19 says $\sigma_{\mathfrak{A}}(a) \subset [-||a||, ||a||]$ and $r(a) = ||a||$, so $||a|| \in \sigma_{\mathfrak{A}}(a)$ or $-||a|| \in \sigma_{\mathfrak{A}}(a)$. Similarly $\sigma_{\mathfrak{B}}(\pi(a)) \subset [-||\pi(a)||, ||\pi(a)||]$. Then choose a continuous map $f : [-||a||, ||a||] \to \mathbb{R}$ that vanishes on $[-||\pi(a)||, ||\pi(a)||]$ and such that $f(-||a||) = f(||a||) = 1$. Theorem 8.21(d) implies $\pi(f(a)) = f(\pi(a)) = 0$, for $f\restriction_{\sigma_{\mathfrak{B}}(\pi(a))} = 0$ and $||f(a)|| = ||f||_{\infty, C(\sigma_{\mathfrak{A}}(a))} \geq 1$. Then $f(a) \neq 0$, contradicting the injectivity of π.
(b) The claim is immediate because π is isometric and by definition of C^*-subalgebra with unit. In particular π isometric warrants $\pi(\mathfrak{A})$ is closed in $\mathfrak{B}$, hence complete as normed space. □

A second result shows that the spectrum of C^*-subalgebras or *-(iso)morphic images does not change.

Theorem 8.23 (Invariance of spectrum). *Let $\mathfrak{A}$ and $\mathfrak{B}$ be C^*-algebras with unit.*
(a) *If $\mathfrak{A}$ is C^*-subalgebra with unit in $\mathfrak{B}$,*

$$\sigma_{\mathfrak{A}}(a) = \sigma_{\mathfrak{B}}(a) \quad \textit{for any } a \in \mathfrak{A}.$$

(b) *If $\pi : \mathfrak{A} \to \mathfrak{B}$ is a *-homomorphism,*

$$\sigma_{\mathfrak{B}}(\pi(a)) = \sigma_{\pi(\mathfrak{A})}(\pi(a)) \subset \sigma_{\mathfrak{A}}(a) \quad \textit{for any } a \in \mathfrak{A}.$$

The last inclusion is an equality if π is one-to-one.

Proof. (a) Let us observe, preliminarily, that the unit $\mathbb{I}$ is the same in $\mathfrak{A}$ and $\mathfrak{B}$. If $a \in \mathfrak{A}$ moreover, a^* is the same in $\mathfrak{A}$ and $\mathfrak{B}$. It is clear that $\rho_{\mathfrak{A}}(a) \subset \rho_{\mathfrak{B}}(a)$, or equivalently $\sigma_{\mathfrak{B}}(a) \subset \sigma_{\mathfrak{A}}(a)$. Thus it is enough to prove, for any $a \in \mathfrak{A}$, that $(a - \lambda\mathbb{I})$ has inverse $(a - \lambda\mathbb{I})^{-1} \in \mathfrak{B}$ belonging to $\mathfrak{A}$. This is the same as demanding that the possible inverse $a^{-1} \in \mathfrak{B}$ to $a \in \mathfrak{A}$ is in $\mathfrak{A}$. Let us consider tha subcase where $a = a^*$ is inveritble in $\mathfrak{B}$. Then $\sigma_{\mathfrak{B}}(a) \subset \mathbb{R}$, and since $\rho_{\mathfrak{B}}(a)$ is open and $0 \in \rho_{\mathfrak{B}}(a)$ there is a disc $D \subset \mathbb{C}$

of radius $r > 0$ at the origin that does not intersect $\sigma_{\mathfrak{B}}(a)$. Hence $f : x \mapsto 1/x$ is continuous and bounded on $\sigma_{\mathfrak{B}}(a)$, and we can define $f(a) = \Phi_a(f)$ using Theorem 8.21 on $a = a^* \in \mathfrak{B}$. By construction $af(a) = f(a)a = \mathbb{I}$, i.e. $f(a) = a^{-1}$ in $\mathfrak{B}$. If $f(a) \in \mathfrak{A}$ the proof ends here. By construction of the one-to-one *-homomorphism Φ_a, we have $f(a) = \lim_{n\to+\infty} p_n(a)$, where the p_n are polynomials and the limit is understood in $\mathfrak{B}$. But $p_n(a) \in \mathfrak{A}$ by definition, for $\mathfrak{A}$ is closed under algebraic operations. Since $\mathfrak{A}$ has the induced topology of $\mathfrak{B}$ and $\mathfrak{A}$ is closed, $f(a) \in \mathfrak{A}$ as required, hence $a^{-1} \in \mathfrak{A}$.

Now consider tha case $a \in \mathfrak{A}$ not self-adjoint, such that $a^{-1} \in \mathfrak{B}$. Then also $(a^*)^{-1} = (a^{-1})^* \in \mathfrak{B}$ and we can write $a^{-1} = (a^*a)^{-1}a^*$. Notice $a^*a \in \mathfrak{A}$ is self-adjoint, so $(a^*a)^{-1} \in \mathfrak{A}$ by the previous argument. Trivially $a^* \in \mathfrak{A}$, so $a^{-1} = (a^*a)^{-1}a^* \in \mathfrak{A}$, thus ending part (a).
(b) The inclusion $\sigma_{\pi(\mathfrak{A})}(\pi(a)) \subset \sigma_{\mathfrak{A}}(a)$ was proved in Corollary 8.18(b). The equality $\sigma_{\mathfrak{B}}(\pi(a)) = \sigma_{\pi(\mathfrak{A})}(\pi(a))$ follows from part (a) and the fact $\pi(\mathfrak{A})$ is a C^*-subalgebra with unit in $\mathfrak{B}$ by Theorem 8.22(b). If π is one-to-one, $\sigma_{\pi(\mathfrak{A})}(\pi(a)) = \sigma_{\mathfrak{A}}(a)$ follows from (a) and the fact that $\pi : \mathfrak{A} \to \pi(\mathfrak{A})$ is a *-isomorphism of C^*-algebras with unit by Theorem 8.22(a). $\square$

Theorem 8.21 also permits to give a reasonable meaning to *positive elements* in a C^*-algebra with unit. The definition and characterisation play an important role in advanced formulations of quantum fields.

Definition 8.24. *An element a in a C^*-algebra with unit $\mathfrak{A}$ is* **positive** *if $a = a^*$ and $\sigma(a) \subset [0, +\infty)$. The set of positive elements of $\mathfrak{A}$ is denoted by $\mathfrak{A}^+$.*

We have arrived at the characterisation of positive elements, together with other properties, given in the next result.

Theorem 8.25 (On positive elements in a C^*-algebra with unit). *Let $\mathfrak{A}$ be a C^*-algebra with unit.*

(a) *If $\alpha_1, \ldots \alpha_n \in [0, +\infty)$ and $a_1, \ldots, a_n \in \mathfrak{A}$ are positive, then $\sum_{j=1}^n \alpha_j a_j$ is positive, so $\mathfrak{A}^+$ is a closed convex cone in $\mathfrak{A}$.*
(b) *The following assertions, for any $a \in \mathfrak{A}$, are equivalent.*

(i) *a is positive;*
(ii) *$a = a^*$ and $a = c^*c$ for some $c \in \mathfrak{A}$;*
(iii) *$a = a^*$ and $a = b^2$ for some self-adjoint $b \in \mathfrak{A}$.*

(c) *If $\mathfrak{A}_0 \subset \mathfrak{A}$ is a C^*-subalgebra with unit, then $\mathfrak{A}_0^+ = \mathfrak{A}_0 \cap \mathfrak{A}^+$, and $\mathfrak{A}_0 =< \mathfrak{A}_0^+ >$.*
(d) *If $\pi : \mathfrak{A} \to \mathfrak{B}$ is a *-homomorphism of C^*-algebras with unit, and $a \in \mathfrak{A}$ is positive, $\pi(a)$ is positive.*

Proof. (a) The claim is clearly true if $n = 1$, for $\sigma(\alpha_1 a_1) = \alpha_1 \sigma(a_1)$, and $\alpha_1 a$ is self-adjoint iff a_1 is and $\alpha_1 \geq 0$. So we will just prove the claim for $n = 2$ with α_1 and α_2 both non-zero. We will make use of the fact d is positive iff self-adjoint, plus

$$\left|\left|\mathbb{I} - ||d||^{-1}d\right|\right| \leq 1\,.$$

The above condition implies $\sigma\left(\mathbb{I}-||d||^{-1}d\right) \subset [-1,1]$ i.e. $1-||d||^{-1}\sigma(d) \subset [-1,1]$, by the properties of the spectral radius. This implies $\sigma(d) \subset [0, 2||d||]$, so d is positive. Conversely, if d is positive then $\sigma(d) \subset [0,||d||]$, so as before $||\mathbb{I}-||d||^{-1}d|| \leq 1$. If $d = d^*$ and

$$||\mathbb{I}-d|| \leq 1$$

then d is positive with $||d|| \leq 2$. The proof is the same as the previous one. All these facts in turn imply, if a_1 and a_2 are self-adjoint positive with $||a_1|| = ||a_2|| = 1$ and $\alpha_1, \alpha_2 \in (0,1)$, $\alpha_1 + \alpha_2 = 1$, that the self-adjoint element $\alpha_1 a_1 + \alpha_2 a_2$ is positive. In fact,

$$||\mathbb{I}-\alpha_1 a_1+\alpha_2 a_2|| \leq \alpha_1||\mathbb{I}-a_1||+\alpha_2||\mathbb{I}-a_2|| \leq \alpha_1+\alpha_2 = 1$$

so $\alpha_1 a_1 + \alpha_2 a_2$ is positive. Multiply by $\lambda > 0$, so (renaming constants) $\lambda\mu a_1 + \lambda(1-\mu)a_2$ is positive whichever $\mu \in (0,1)$ and $\lambda \in (0,+\infty)$ are chosen. If now we take $\alpha_1, \alpha_2 > 0$ without further conditions, $\lambda = \alpha_1 + \alpha_2 \in (0,+\infty)$ and $\mu = \alpha_1/(\alpha_1+\alpha_2) \in (0,1)$ immediately, and

$$\lambda\mu a_1 + \lambda(1-\mu)a_2 = \alpha_1 a_1 + \alpha_2 a_2\,.$$

But now $\alpha_1 a_1 + \alpha_2 a_2$ is positive for arbitrary $\alpha_1, \alpha_2 > 0$, so the claim is proved (note that the constraint $||a_1|| = ||a_2|| = 1$ has disappeared). Let us show $\mathfrak{A}^+$ is closed. If $\mathfrak{A}^+ \ni a_n \to a \in \mathfrak{A}$ then $||a_n - a|| \to 0$, so $||a_n|| - ||a|| \to 0$. That $a_n \in \mathfrak{A}^+$, by the properties of spectrum and spectral radius, implies $||\,||a_n||\mathbb{I}-a_n|| \leq ||a_n||$. In the limit $n \to +\infty$ we find $||\,||a_n||\mathbb{I}-a|| \leq ||a||$, hence $a \in \mathfrak{A}^+$.
(b) If (iii) holds, Proposition 8.19(d) gives $\sigma(a) = \sigma(b^2) = \{\lambda^2 \mid \lambda \in \sigma(a)\} \subset [0,+\infty)$, so (iii) implies (i). Now the converse. Using continuous functional calculus, and recalling $a = a^*$, the real continuous map $\sqrt{\cdot} : \sigma(a) \ni x \mapsto \sqrt{x}$ allows to define $\sqrt{a} := \Phi_a(\sqrt{\cdot})$. Set $b := \sqrt{a}$, so $b = b^*$ and $b^2 = a$, because Φ_a is a *-homomorphism. Hence (i) and (iii) are equivalent. That (iii) implies (ii) is obvious. So there remains to show (ii) gives (i). Let $a = a^*$, $a = c^*c$, and we claim $\sigma(a) \subset [0,+\infty)$. By contradiction assume $\sigma(-a) \subset (0,+\infty)$. Then Proposition 8.19(e) tells $\sigma(-cc^*) \setminus \{0\} = \sigma(-c^*c) \setminus \{0\} \subset (0,+\infty)$. Decomposing $c := c_1 + ic_2$, c_1, c_2 self-adjoint, we have

$$c^*c + cc^* = 2c_1^2 + 2c_2^2\,.$$

But c_1^2 and c_2^2 are positive by (iii), and $-cc^*$ is positive by assumption. Hence the linear combination with positive coefficients $2c_1^2 + 2c_2^2 - cc^* = c^*c$ is a positive operator by (a). Therefore $\sigma(c^*c) \subset [0,+\infty)$, but since $\sigma(-c^*c) \setminus \{0\} \subset (0,+\infty)$ as well, we have $\sigma(c^*c) = \{0\}$ i.e. $\sigma(a) = \sigma(-a) = \{0\}$, a contradiction. Hence $\sigma(-a) \subset (-\infty, 0]$, i.e. $\sigma(a) \subset [0,+\infty)$, so (ii) implies (i).
(c) If $a \in \mathfrak{A}_0$ is positive in $\mathfrak{A}$, it is positive in $\mathfrak{A}_0$ and conversely, for $\sigma_{\mathfrak{A}}(a) = \sigma_{\mathfrak{A}_0}(a)$ by Theorem 8.23(a), and also $a = a^*$ is invariant. Hence $\mathfrak{A}_0^+ = \mathfrak{A}_0 \cap \mathfrak{A}^+$. If $a \in \mathfrak{A}_0$, write $a = a_1 + ia_2$, with $a_1 := (a+a^*)/2$ and $a_2 := (a-a^*)/(2i)$ self-adjoint. If b is self-adjoint, we can define $b_+ := (|b|+b)/2$ and $b_- := (|b|-b)/2$, where $|b| = \Phi_b(|\cdot|)$ and $|\cdot| : \mathbb{C} \to [0,+\infty)$ is the modulus. Since Φ_b is a *-homomorphism and $|\cdot|$ is real-valued, b_+ and b_- are self-adjoint because b is (in particular $\sigma(b) \subset \mathbb{R}$). Property (c) in Theorem 8.21 says b_+ and b_- are positive, as $|x| \pm x \geq 0$ for any $x \in \sigma(b) \subset \mathbb{R}$. In

conclusion, every $a \in \mathfrak{A}_0$ is the complex linear combination of 4 positive elements in $\mathfrak{A}_0$, so $\mathfrak{A}_0 =< \mathfrak{A}_0^+ >$.
(d) This follows immediately from (b), using (iii) and bearing in mind π is a *-homomorphism. □

8.2.3 Commutative Banach algebras and the Gelfand transform

In order to generalise the isometric *-homomorphism $\Phi_a : C(\sigma(a)) \to \mathfrak{A}$ (defined for $a^* = a \in \mathfrak{A}$ in Chapter 8.2.1) for a normal, not self-adjoint, we introduce some technical results in the theory of commutative Banach (C^*-)algebras, due to Gelfand and interesting by their own means. We will prove a characterisation, the *commutative Gelfand–Najmark theorem*, according to which any commutative C^*-algebra with unit is canonically a C^*-algebra $C(\mathsf{X})$ of functions with norm $|| \, ||_\infty$ over the compact Hausdorff space X given by the algebra itself.

We need a technical result that explains the relationship between *maximal ideals* in Banach algebras and *multiplicative linear functionals*, after the mandatory definitions. In the sequel every Banach algebra will be complex.

Definition 8.26. *If $\mathfrak{A}$ is a Banach algebra with unit, a subset $I \subset \mathfrak{A}$ is a* **maximal ideal** *if*

(i) *I is a subspace in $\mathfrak{A}$;*
(ii) *$ba, ab \in I$ for any $a \in I$, $b \in \mathfrak{A}$;*
(iii) *$I \neq \mathfrak{A}$;*
(iv) *if $I \subset J$, with J as in (i), (ii), then either $J = I$ or $J = \mathfrak{A}$.*

Remarks 8.27. Conditions (i) and (ii) say I is an ideal, whereas (iii) prescribes the ideal must be *proper*. *Maximality* is expressed by (iv). ■

Definition 8.28. *If $\mathfrak{A}$ is a Banach algebra with unit, a multiplicative linear functional $\phi : \mathfrak{A} \to \mathbb{C}$ is called a* **character** *of $\mathfrak{A}$; i.e. $\phi(ab) = \phi(a)\phi(b)$. If $\mathfrak{A}$ is also commutative, the set of non-trivial characters is denoted by $\sigma(\mathfrak{A})$ and called the* **spectrum of the algebra**.

Now we can state and prove the advertised proposition.

Proposition 8.29. *Let $\mathfrak{A}$ be a Banach algebra with unit $\mathbb{I}$.*

(a) *A character χ of $\mathfrak{A}$ is non-zero iff $\chi(\mathbb{I}) = 1$.*
(b) *A maximal ideal $I \subset \mathfrak{A}$ is closed.*
(c) *If $\mathfrak{A}$ is commutative, the map $\mathscr{I} : \sigma(\mathfrak{A}) \ni \chi \mapsto Ker(\chi) \subset \mathfrak{A}$ is a bijection on the set of maximal ideals.*
(d) *If $\mathfrak{A}$ is commutative then $\sigma(\mathfrak{A}) \subset \mathfrak{A}'$, i.e. characters are continuous.*

Proof. Observe preliminarily that the existsence of the unit $\mathbb{I}$ in $\mathfrak{A}$, with $||\mathbb{I}|| = 1$, implies $\mathfrak{A} \neq \{0\}$.
(a) If χ is a character $\chi(a) = \chi(\mathbb{I}a) = \chi(\mathbb{I})\chi(a)$. If $\chi \neq 0$ then $\chi(a) \neq 0$ for some $a \in \mathfrak{A}$. Then $\chi(\mathbb{I}) = 1$. If $\chi(\mathbb{I}) \neq 0$, clearly $\chi \neq 0$.

(b) By assumption $I \neq \mathfrak{A}$, so $\mathbb{I} \notin I$ (otherwise $a = a\mathbb{I} \in I$ for any $a \in \mathfrak{A}$). Hence $\mathbb{I} \notin \overline{I}$. In fact, if $\mathbb{I} \in \overline{I}$, since the the set of invertible elements is open (Remark 2.24(2)), there would be an open neighbourhood B of $\mathbb{I}$ of invertible elements intersecting I. For any $a \in B \cap I$, then, $\mathbb{I} = a^{-1}a \in I$, which cannot be. Therefore $\overline{I} \neq \mathfrak{A}$, $\mathbb{I}$ being excluded. Since $\overline{I} \supset I$ and $\overline{I}$ satisfies (i), (ii), (iii) in Definition 8.26, we have $I = \overline{I}$ by Definition 8.26(iv).
(c) If $\chi \in \sigma(\mathfrak{A})$, $Ker(\chi)$ is a maximal ideal: (i),(ii) in Definition 8.26 are true as χ is linear and multiplicative, and (iii) holds for $\chi \neq 0$. Notice $\mathfrak{A} = Ker(\chi) \oplus V$, where $dim(V) = 1$, for this must be the dimension of the target space $\mathbb{C}$ of χ. Hence any subspace $J \subset \mathfrak{A}$ including $Ker(\chi)$ properly must be $\mathfrak{A}$ itself, so $Ker(\chi)$ is a maximal ideal. Therefore the map $\mathscr{I}$ sends characters to maximal ideals. Let us show it is one-to-one. If $\chi, \chi' \in \sigma(\mathfrak{A})$ and $Ker(\chi) = Ker(\chi') = N$, by $\mathfrak{A} = N \oplus V$ we have $\chi(a) = \chi(v_a)$ and $\chi'(a) = \chi'(v_a)$, where $n_a \in N$ and $v_a \in V$ are the projections of a on N and V. If e is a basis of V (1-dimensional), $v_a = c_a e$ for some complex number c_a determined by a. Hence $Ker(\chi) = Ker(\chi')$ implies $\chi(a) = a_v\chi(e)$ and $\chi'(a) = a_v\chi'(e)$. By (a) $\chi(\mathbb{I}) = \chi'(\mathbb{I}) = 1$, so $\chi(e) = \chi'(e)$ and $\chi = \chi'$, proving injectivity. Now surjectivity. If I is a maximal ideal, it is closed by (b). It is easy to show the quotient space $\mathfrak{A}/I$ of equivalence classes $[a]$ ($a \sim a' \Leftrightarrow a - a' \in I$) inherits a natural Banach space structure and a commutative Banach algebra structure with unit $[\mathbb{I}]$ from $\mathfrak{A}$. By construction $\mathfrak{A}/I$ does not contain ideals other than $\mathfrak{A}/I$ itself. So any non-null $[a] \in \mathfrak{A}/I$ is invertible, otherwise $[a]\mathfrak{A}/I$ would be a proper ideal in $\mathfrak{A}/I$. The Gelfand-Mazur theorem (8.10) guarantees the existence of a Banach space isomorphism $\psi : \mathfrak{A}/I \to \mathbb{C}$. If $\pi : \mathfrak{A} \ni a \to [a] \in \mathfrak{A}/I$ denotes the canonical projection (continuous by construction), $\chi := \psi \circ \pi : \mathfrak{A} \to \mathbb{C}$ is an element of $\sigma(\mathfrak{A})$ with trivial null space $Ker(\chi) = I$.
(d) The last argument above, being $\chi \mapsto Ker(\chi)$ a bijection, also tells that any character χ must look like $\psi \circ \pi$, for $Ker(\chi) = I$. Thus χ is continuous, because ψ and π are. □

Now it is time for the first theorem of Gelfand on commutative Banach algebras with unit. We refer to the **-weak topology* on the dual $\mathfrak{A}'$ of $\mathfrak{A}$ (seen as Banach space) introduced by Definition 2.68. More precisely, viewing $\sigma(\mathfrak{A})$ as subset of $\mathfrak{A}'$ with the induced topology, we consider the algebra with unit $C(\sigma(\mathfrak{A}))$ of continuous maps from $\sigma(\mathfrak{A})$ to $\mathbb{C}$ with norm $|| \, ||_\infty$. One result of the theorem establishes that $\sigma(\mathfrak{A})$ is a compact Hausdorff space; as we saw in Chapter 2 and 3 (Examples 2.26(4), 3.44(1)), in fact, $C(\sigma(\mathfrak{A}))$ is a Banach algebra with unit (and also a C^*-algebra).

Theorem 8.30. *Let $\mathfrak{A}$ be a commutative Banach algebra with unit $\mathbb{I}$ and let*

$$\mathscr{G} : \mathfrak{A} \ni x \mapsto \widehat{x} : \sigma(\mathfrak{A}) \to \mathbb{C}, \tag{8.12}$$

denote the **Gelfand transform**, *where*

$$\widehat{x}(\chi) := \chi(x), \quad x \in \mathfrak{A}, \, \chi \in \sigma(\mathfrak{A}). \tag{8.13}$$

Then

(a) $\sigma(\mathfrak{A})$ *is a *-weakly compact Hausdorff space, and $||\chi|| \leq 1$ if $\chi \in \sigma(\mathfrak{A})$ ($||.||$ is the strong norm on $\mathfrak{A}'$).*

(b) *If* $x \in \mathfrak{A}$*:*

$$\sigma(x) = \{\widehat{x}(\chi) \,|\, \chi \in \sigma(\mathfrak{A})\}\,.$$

(c) $\widehat{\mathfrak{A}} \subset C(\sigma(\mathfrak{A}))$*, and* $\mathscr{G} : \mathfrak{A} \to C(\sigma(\mathfrak{A}))$ *is a homomorphism of Banach algebras with unit.*
(d) $\mathscr{G} : \mathfrak{A} \to C(\sigma(\mathfrak{A}))$ *is continuous,* $||\widehat{x}||_\infty \leq ||x||$ *for any* $x \in \mathfrak{A}$*.*

Proof. (a) Consider $\chi \in \sigma(\mathfrak{A})$ and the associated maximal ideal $I = Ker(\chi)$ under Proposition 8.29(c). If $x \in \mathfrak{A}$, $\chi(x - \chi(x)\mathbb{I}) = 0$, so $x - \chi(x)\mathbb{I} \in I$ cannot be invertible (cf. propostion 8.29(b)). Then $\chi(x) \in \sigma(x)$, so $|\chi(x)| \leq ||x||$ by elementary properties of the spectral radius. Consequently $||\chi|| \leq 1$, where the norm defines the strong topology. Therefore $\sigma(\mathfrak{A})$ is contained in the unit ball of the dual $\mathfrak{A}'$. We know this set is *-weakly compact by Theorem 2.76 (Banach–Alaoglu). Since the *-weak topology is Hausdorff, to finish it suffices to show $\sigma(\mathfrak{A})$ is *-weakly closed. Saying $\sigma(\mathfrak{A}) \ni \chi_n \to \chi \in \mathfrak{A}'$ in that topology means $\chi_n(x) \to \chi$ for any $x \in \mathfrak{A}$. By continuity χ is a character if all χ_n are. Thus $\sigma(\mathfrak{A})$ is closed in the *-weak topology.
(b) Above we proved $\chi(x) \in \sigma(x)$, so $\{\widehat{x}(\chi) \,|\, \chi \in \sigma(\mathfrak{A})\} \subset \sigma(x)$. Let us prove the converse inclusion. If $\lambda \in \sigma(x)$ then $x - \lambda\mathbb{I}$ is not invertible, so $x\mathfrak{A} := \{(x-\lambda\mathbb{I})y \,|\, y \in \mathfrak{A}\}$ is a proper ideal. Zorn's lemma gives us a maximal ideal I containing $x\mathfrak{A}$. Let $\chi_I \in \sigma(\mathfrak{A})$ be the associated character by Proposition 8.29(c). Then $\widehat{x}(\chi_I) = \chi_I(x) = \lambda$ and so $\{\widehat{x}(\chi) \,|\, \chi \in \sigma(\mathfrak{A})\} \supset \sigma(x)$, as required.
(c) and (d) That $\mathscr{G}$ is a homomorphism of algebras with unit is straightforward, because $\widehat{x}$ acts on characters χ (linear and multiplicative, plus $\widehat{\mathbb{I}}(\chi) := \chi(\mathbb{I}) = 1$). Moreover, from (b) and the definition of spectral radius we have $||\widehat{x}||_\infty = r(x)$; on the other hand $r(x) \leq ||x||$ by elementary properties of the spectral radius. □

Example 8.31. Let $\ell^1(\mathbb{Z})$ be the Banach space of maps $f : \mathbb{Z} \to \mathbb{C}$ such that

$$||f||_1 := \sum_{n \in n\mathbb{Z}} |f(n)| < +\infty\,.$$

Equip $\ell^1(\mathbb{Z})$ with the structure of a Banach algebra with unit by defining the product using the *convolution*:

$$(f * g)(m) := \sum_{n \in \mathbb{Z}} f(m-n)g(n), \quad f, g \in \ell^1(\mathbb{Z}).$$

This product is well defined and satisfies $||f * g||_1 \leq ||f||_1 ||g||_1$, because:

$$\sum_{n\in\mathbb{Z}} |(f*g)(n)| = \sum_{n\in\mathbb{Z}} \left| \sum_{m\in\mathbb{Z}} f(n-m)g(m) \right| \leq \sum_{n\in\mathbb{Z}} \sum_{m\in\mathbb{Z}} |f(n-m)|\,|g(m)|$$
$$= \sum_{m\in\mathbb{Z}} \sum_{n\in\mathbb{Z}} |f(n-m)|\,|g(m)| \leq \sum_{m\in\mathbb{Z}} \left(|g(m)| \sum_{n\in\mathbb{Z}} |f(n-m)| \right) = \sum_{m\in\mathbb{Z}} |g(m)|\,||f||_1$$
$$= ||f||_1\,||g||_1\,.$$

There is a unit $\mathbb{I}$, the map $\mathbb{I}(n) = 1$ if $n = 0$ and $\mathbb{I}(n) = 0$ if $n \neq 0$. Since $f * g = g * f$, as is easy to see, $\ell^1(\mathbb{Z})$ becomes a commutative Banach algebra with unit, and we can apply Gelfand's theory.

Let $\mathbb{S}^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ and define characters χ_z associated to $z \in \mathbb{S}^1$:

$$\chi_z(f) := \sum_{n \in n\mathbb{Z}} f(n) z^n .$$

Trivially they are well-defined characters. Hence we have a function $\Gamma : \mathbb{S}^1 \ni z \mapsto \chi_z \in \sigma(\ell^1(\mathbb{Z}))$ easily seen to be invertible. Actually, it is a homeomorphism, for we shall prove it is a continuous bijection between compact Hausdorff spaces (Proposition 1.23). Continuity, using the *-weak topology on $\sigma(\ell^1(\mathbb{Z}))$, amounts to continuity of $\mathbb{S}^1 \ni z \mapsto \chi_z(f) \in \mathbb{C}$ with $f \in \ell^1(\mathbb{Z})$ fixed, because $z \mapsto \chi_z(f)$ is the uniform limit of the continuous $g_m(z) := \sum_{|n|<m} f(n)z^n$, for $\sum_{n\in\mathbb{Z}} |f(n)z^n| = ||f||_1 < +\infty$ with $|z| = 1$. Therefore we may identify the spectrum $\sigma(\ell^1(\mathbb{Z}))$ with $\mathbb{S}^1$ under the homeomorphism Γ. The Gelfand transform $\widehat{f}$ of $f \in \ell^1(\mathbb{Z})$ is thus continuous on $\mathbb{S}^1$, and defined by

$$\widehat{f}(z) := \sum_{n \in \mathbb{Z}} f(n) z^n .$$

The elementary theory of Fourier series forces $f(n)$ to be the Fourier coefficient

$$f(n) = \frac{1}{2\pi} \int_0^{2\pi} \widehat{f}(e^{i\theta}) e^{-in\theta} d\theta .$$

Therefore $\mathscr{G}(\ell^1(\mathbb{Z}))$ is the subset, in the Banach algebra with unit $(C(\mathbb{S}^1), || \, ||_\infty)$, of maps with absolutely convergent Fourier series. Gelfand observed there is an interesting consequence to that, corresponding to a classical statement due to Wiener (but proved by different means):

Proposition 8.32. *If $h \in C(\mathbb{S}^1)$ has absolutely convergent Fourier series and no zeroes, the map $\mathbb{S}^1 \ni z \mapsto 1/h(z)$ (belonging in $C(\mathbb{S}^1)$) has absolutely convergent Fourier series.*

Proof. First, $h = \widehat{f}$ for some $f \in \ell^1(\mathbb{Z})$. Since $\widehat{f}(z) \neq 0$, then $0 \notin \sigma(f)$ by Theorem 8.30(b). Hence f has inverse $g \in \ell^1(\mathbb{Z})$ and $\widehat{g} = 1/h$. We conclude that the Fourier series of $1/h$ must converge absolutely. □

To conclude we consider the more rigid case in which $\mathfrak{A}$ is a commutative C^*-algebra with unit. Then the Gelfand transform defines an honest *-isomorphism of C^*-algebras with unit, and must be isometric by Theorem 8.22(a). In fact we have the following commutative version of the Gelfand–Najmark theorem.

Theorem 8.33 (Commutative Gelfand–Najmark theorem). *Let $\mathfrak{A}$ be a commutative C^*-algebra with unit. If we think $C(\sigma(\mathfrak{A}))$ as a commutative C^*-algebra with unit (for the norm $|| \, ||_\infty$), the Gelfand transform*

$$\mathscr{G} : \mathfrak{A} \ni x \mapsto \widehat{x} \in C(\sigma(\mathfrak{A})) \quad \textit{where} \quad \widehat{x}(\chi) := \chi(x) \quad , \quad x \in \mathfrak{A}, \, \chi \in \sigma(\mathfrak{A}),$$

is an isometric *-isomorphism.

Proof. The only thing to prove is that the Gelfand transform defines a *-isomorphism, because the rest follows from Theorem 8.22(a). Knowing the Gelfand transform is an algebra homomorphism, though, requires we prove surjectivity and the involution property only. The first lemma is that $\widehat{x}$ is real if $x^* = x \in \mathfrak{A}$. If so, with $t \in \mathbb{R}$ we define

$$u_t := e^{itx} := \sum_{n=0}^{+\infty} \frac{(it)^n}{n!} x^n$$

with respect to the norm of $\mathfrak{A}$. Since $\mathfrak{A}$ is commutative, and working as we were in $\mathbb{C}$, we have $u_t^* = u_t$ and $u_t^* u_t = u_0 = \mathbb{I}$. Taking norms gives $||u_t|| = ||u_{-t}|| = 1$. If now χ is a character (continuous, linear and multiplicative), we see $\chi(u_t) = e^{it\chi(x)}$ and $\chi(u_{-t}) = e^{-it\chi(x)}$. So by Theorem 8.30(d):

$$|\chi(u_{\pm t})| = |\widehat{u_{\pm t}}(\chi)| \leq ||\widehat{u_{\pm t}}||_\infty \leq ||u_{\pm t}|| \leq 1 \,.$$

That is to say $|e^{\pm it\chi(x)}| \leq 1$, implying $\chi(x) \in \mathbb{R}$. Now if $x \in \mathfrak{A}$ we can decompose $x = a + ib$, $a = a^*$, $b = b^*$. Hence

$$\widehat{x^*}(\chi) = \chi(x^*) = \chi(a - ib) = \chi(a) - i\chi(b) = \overline{\chi(a) + i\chi(b)} = \overline{\chi(x)} = \overline{\widehat{x}(\chi)} \,.$$

Therefore the Gelfand transform preserves the involution.

To conclude we settle surjectivity, showing $\{\widehat{x} \,|\, x \in \mathfrak{A}\} = C(\sigma(\mathfrak{A}))$. The set on the left is closed as compact (continuous image of a compact set, Theorem 8.30) in a Hausdorff space. By construction, this set is a closed *-subalgebra of $C(\sigma(\mathfrak{A}))$ containing the identity ($\widehat{\mathbb{I}} = 1$, identity map). The elements of that algebra separate points of $\sigma(\mathfrak{A})$: if $\chi_1 \neq \chi_1$ then $\chi_1(x) \neq \chi_2(x)$ for some $x \in \mathfrak{A}$, so $\widehat{x}(\chi_1) \neq \widehat{x}(\chi_2)$. The Stone–Weiestrass theorem implies $\{\widehat{x} \,|\, x \in \mathfrak{A}\} = C(\sigma(\mathfrak{A}))$. □

Remark 8.34. (1) The commutative Gelfand–Najmark theorem proves that every commutative C^*-algebra $\mathfrak{A}$ with unit is canonically a C^*-algebra $C(\mathsf{X})$ of functions with norm $||\,||_\infty$ on a compact set $\mathsf{X} = \sigma(\mathfrak{A})$. The "points" of X are the characters of the C^*-algebra. Put equivalently, commutative C^*-algebras with unit are C^*-algebras of functions built in a canonical manner via the algebra's spectrum $\sigma(\mathfrak{A})$.
(2) If we start from a concrete C^*-algebra $C(\mathsf{X})$ of functions on a compact Hausdorff space X, Gelfand's procedure recovers exactly this algebraic construction, because characters, in the present case, are nothing but points in X. In fact, any $x \in \mathsf{X}$ can be mapped one-to-one to the corresponding character $\chi_x : C(\mathsf{X}) \to \mathbb{C}$, $\chi_x(f) := f(x)$ for any $f \in C(\mathsf{X})$. It can be proved that every character has this form by showing it is positive (by multiplicativity), and that it must be a positive Borel measure by the theorem of Riesz. Since the only multiplicative Borel measures are Dirac measures δ_x, we have $\chi(f) = \int_{\mathsf{X}} f d\delta_x = f(x)$ for some $x \in \mathsf{X}$ determined by χ. Observe that the topology on X coincides with the *-weak topology if we interpret points $x \in \mathsf{X}$ as characters χ_x, as is immediate to verify.

Naïvely speaking, a compact Hausdorff space can be fully described by the commutative C^*-algebra of its continuous complex functions. This remark can be taken, and indeed was by A. Connes, as a starting point to develop *noncommutative geometry*: instead of using a commutative C^*-algebra with unit one takes a noncommutative algebra, and the associated "space" is defined in terms of continuous linear functionals on the algebra. ■

8.2.4 Abstract C^*-algebras: functional calculus for continuous maps and normal elements

We wish to extend Chapter 8.2.1 to *normal elements* $a \in \mathfrak{A}$: $a^*a = aa^*$ in a C^*-algebra $\mathfrak{A}$ with unit $\mathbb{I}$. We want to make sense of the function $f(a,a^*) \in \mathfrak{A}$ of a, a^* when f is a continuous complex-valued map defined on the spectrum of a.

A few preliminary remarks and notational issues must be seen to before defining the functions $f(a,a^*)$.

We can always decompose a and a^* into linear combinations of two commuting *self-adjoint* elements x, y:

$$a = x_a + iy_a\,, \quad a^* = x_a - iy_a\,, \tag{8.14}$$

where by definition

$$x_a := \frac{a+a^*}{2}\,, \quad y_a := \frac{a-a^*}{2i}\,. \tag{8.15}$$

x_a and y_a are clearly self-adjoint. That they commute is also obvious, for a and a^* commute.

Decomposition (8.19) reminds of the analogue splitting of a complex number into real and imaginary parts

$$z = x + iy\,, \quad \overline{z} = x - iy\,, \tag{8.16}$$

where

$$x := \frac{z+\overline{z}}{2}\,, \quad y := \frac{z-\overline{z}}{2i}\,. \tag{8.17}$$

Remarks 8.35. The maps $f : \sigma(a) \to \mathbb{C}$ we shall deal with are to be thought of as functions in x and y, imagining $\sigma(a)$ as subset of $\mathbb{R}^2$ rather than $\mathbb{C}$. Equivalently, the variables may be taken to be z and $\overline{z}$, considered *independent*. They are bijectively determined by x, y, so maps in x, y are in one-to-one correspondence to maps in z, $\overline{z}$: to $f = f(z,\overline{z})$ we may associate $g = g(x,y)$:

$$g(x,y) := f(x+iy, x-iy)$$

and conversely, to $g = g(x,y)$ is associated $f_1 = f_1(z,\overline{z})$, where

$$f_1(z,\overline{z}) := g((z+\overline{z})/2, (z-\overline{z})/2i)\,.$$

Clearly, $f_1 = f$. This fact will be used often without further notice. ■

Now we are ready for the continuous functional calculus for normal elements. The proof will be substantially different from Theorem 8.21, in that it will involve the Gelfand transform of the previous section. We shall still use the name Φ_a for the *-isomorphism, because this generalises the one of Theorem 8.21.

Theorem 8.36 (Functional calculus for continuous maps and normal elements). *Let $\mathfrak{A}$ be a C^*-algebra with unit $\mathbb{I}$ and $a \in \mathfrak{A}$ a normal element. View f as a function of the independent variables z and $\overline{z}$.*

(a) *There exists a unique *-homomorphism on the commutative C^*-algebra with unit $C(\sigma(a))$:*

$$\Phi_a : C(\sigma(a)) \ni f \mapsto f(a,a^*) \in \mathfrak{A}\,,$$

such that

$$\Phi_a(z) = a \tag{8.18}$$

with z being the polynomial $\sigma(a) \ni (z,\overline{z}) \mapsto z$.
(b) *The following properties hold:*

(i) *Φ_a is isometric: for any $f \in C(\sigma(a))$, $||\Phi_a(f)|| = ||f||_\infty$;*
(ii) *if $ba = ab$ and $ba^* = a^*b$ for some $b \in \mathfrak{A}$, then $bf(a,a^*) = f(a,a^*)b$ for any $f \in C(\sigma(a))$;*
(iii) *Φ_a preserves the involution: $\Phi_a(\overline{f}) = \Phi_a(f)^*$ for any $f \in C(\sigma(a))$.*

(c) *$\sigma(f(a,a^*)) = f(\sigma(a),\overline{\sigma(a)})$ for any $f \in C(\sigma(a))$.*
(d) *If $\mathfrak{B}$ is a C^*-algebra with unit and $\pi : \mathfrak{A} \to \mathfrak{B}$ a *-homomorphism,*

$$\pi(f(a,a^*)) = f(\pi(a),\pi(a^*)) \quad \textit{for any } f \in C(\sigma_{\mathfrak{A}}(a)).$$

(e) *If $a = a^*$ the *-homomorphism Φ_a coincides with its analogue of Theorem 8.21.*

Proof. (a), (b) and (e) Uniqueness is evident because if two *-homomorphisms $\Phi_a : C(\sigma(a)) \to \mathfrak{A}$ and $\Phi'_a : C(\sigma(a)) \to \mathfrak{A}$ satisfy $\Phi_a(z) = \Phi'_a(z) = a$, by definition they concide on the polynomial algebra in z and $\overline{z}$, which is dense in $C(\sigma(a))$ in norm $||\,||_\infty$ by Stone–Weierstrass ($\sigma(a)$ is compact and Hausdorff). As *-homomorphisms are continuous (Theorem 8.22), $\Phi_a(f) = \Phi'_a(f)$ for any $f \in C(\sigma(a))$. The same argument proves, in the case $a = a^*$, that the *-homomorphism Φ_a coincides with its cousin in Theorem 8.21. Likewise, if Φ_a is defined, then (ii) in (b) holds, because if b commutes with a and a^* it commutes with every polynomial in a, a^*, and by continuity with any $\Phi_a(f)$.

Let us show Φ_a exists and satisfies the remaining requests in (a) and (b). Consider the commutative C^*-(sub)algebra with unit $\mathfrak{A}_a \subset \mathfrak{A}$ spanned by $\mathbb{I}$, a and a^*. It is the closure, for the norm of $\mathfrak{A}$, of the set of polynomials $p(a,a^*)$ with complex coefficients. The idea is to define $\Phi_a(a)$ by $\mathscr{G}^{-1}(f)$, because the inverse Gelfand transform $\mathscr{G}^{-1} : C(\sigma(\mathscr{A})) \to \mathfrak{A}_a$ is an isometric *-isomorphism by Theorem 8.33. The problem is that now f is defined on $\sigma(\mathfrak{A}_a)$, not on $\sigma(a)$. So let us prove $\sigma(\mathfrak{A}_a)$ and $\sigma(a)$ are homeomorphic under $F : \sigma(\mathfrak{A}_a) \ni \chi \mapsto \chi(a) \in \sigma(a)$. That $\chi(a) \in \sigma(a)$ follows from (b) in Theorem 8.30. The function is continuous because characters are continuous, by Proposition 8.29(d), and it acts between compact Hausdorff spaces. Hence it is enough to show it is bijective to have a homeomorphism (Proposition 1.23). If $F(\chi) = F(\chi')$ then $\chi(a) = \chi'(a)$, $\overline{\chi(a)} = \overline{\chi'(a)}$ (see the proof of Theorem 8.33) and $\chi(a^*) = \chi'(a^*)$. On the other hand $\chi(\mathbb{I}) = \chi'(\mathbb{I}) = 1$ by Proposition 8.29(a). Since χ preserves sums and products, by continuity $\chi(b) = \chi'(b)$ if $b \in \mathfrak{A}_a$, and F is injective.

F is onto by Theorem 8.30(b). Define

$$\Phi_a(f) := \mathscr{G}^{-1}(f \circ F)$$

for $f \in C(\sigma(a))$. By construction Φ_a is an isometric *-isomorphism from $C(\sigma(a))$ to $\mathfrak{A}_a$ such that $\Phi_a^{-1}(a) = z$, where z is $\sigma(a) \ni (z,\overline{z}) \mapsto z$. In fact, $\Phi_a^{-1}(a) = z$ means $\mathscr{G}(a) = z \circ F$ i.e. $\chi(a) = z(\chi)$ for any character $\chi \in \sigma(\mathfrak{A}_a)$. But the latter is true by definition of F. Hence (a), (b) are valid by redefining Φ_a as valued in the larger algebra $\mathfrak{A}$.

(c) By Theorem 8.23, first of all, $\sigma_{\mathfrak{A}}(f(a,a^*)) = \sigma_{\mathfrak{A}_a}(f(a,a^*))$, so we look at the spectrum of $f(a,a^*)$ in $\mathfrak{A}_a$. Then $\Phi_a : C(\sigma(a)) \to \mathfrak{A}_a$ defines an isometric *-isomorphism. The abstract function $f(a,a^*) - \lambda\mathbb{I}$ corresponds to the concrete $\sigma(s) \ni (z,\overline{z}) \mapsto f(z,\overline{z}) - \lambda$. Therefore $f(a,a^*) - \lambda\mathbb{I}$ is invertible iff $\sigma(s) \ni (z,\overline{z}) \mapsto (f(z,\overline{z}) - \lambda)^{-1}$ is in $C(\sigma(a))$. Since the range of f is compact (continuous image of a compact set), the assertion is equivalent to $\lambda \notin f(\sigma(a), \overline{\sigma(a)})$. Now (c) is immediate.

(d). We prove the equivalent $\pi(\Phi_a(f)) = \Phi_{\pi(a)}(f)$. By construction $C(\sigma(a)) \ni f \to \pi(\Phi_a(f)) \in \pi(\mathfrak{A})$ and $C(\sigma(a)) \ni f \to \Phi_{\pi(a)}(f) \in \pi(\mathfrak{A})$ are continuous *-homomorphisms. Trivially, $\pi(\Phi_a(z)) = \pi(a) = \Phi_{\pi(a)}(z)$, $\pi(\Phi_a(\overline{z})) = \pi(a)^* = \Phi_{\pi(a)}(\overline{z})$ and $\pi(\Phi_a(1)) = \mathbb{I} = \Phi_{\pi(a)}(1)$. Therefore $\pi(\Phi_a(p)) = \Phi_{\pi(a)}(p)$ on polynomials $p = p(z,\overline{z})$, and by continuity they coincide on any $f \in \sigma(a)$. □

8.2.5 *C^*-algebras of operators in $\mathfrak{B}(\mathsf{H})$: functional calculus for bounded measurable functions*

Let us return to functional calculus for operators and specialise Chapter 8.2.4 to $\mathfrak{A} = \mathfrak{B}(\mathsf{H})$, H Hilbert space, and instead of the normal $a \in \mathfrak{A}$ consider a normal operator $T \in \mathfrak{B}(\mathsf{H})$. Then the *-homomorphism Φ_T is a *representation* of $C(\sigma(T))$ on H (Definition 3.48). Here as well it is conveninet to decompose T into self-adjoint operators $X, Y \in \mathfrak{B}(\mathsf{H})$:

$$T = X + iY\,, \quad T^* = X - iY\,, \tag{8.19}$$

where

$$X := \frac{T + T^*}{2}\,, \quad Y := \frac{T - T^*}{2i}\,. \tag{8.20}$$

The operators X and Y are patently self-adjoint by construction, and commute since T is normal and commutes with T^*.

Decomposition (8.19) is akin to the real/imaginary decomposition of a complex number

$$z = x + iy\,, \quad \overline{z} = x - iy\,, \tag{8.21}$$

where

$$x := \frac{z + \overline{z}}{2}\,, \quad y := \frac{z - \overline{z}}{2i}\,. \tag{8.22}$$

As before, we may view $f(z,\overline{z})$ as a complex function in x and y. Theorem 8.36 specialises, with identical proof, as follows. We refer to Definition 3.48 for the notion of representation of a C^*-algebra.

Proposition 8.37. *Let* H *be a Hilbert space and* $T \in \mathfrak{B}(\mathsf{H})$ *a normal operator.*

(a) *There exists a unique representation of the commutative* C^**-algebra with unit* $C(\sigma(T))$ *on* H*:*

$$\Phi_T : C(\sigma(T)) \ni f \mapsto f(T,T^*) \in \mathfrak{B}(\mathsf{H}),$$

such that, if z *is the polynomial* $\sigma(T) \ni (z,\overline{z}) \mapsto z$*:*

$$\Phi_T(z) = T \tag{8.23}$$

(b) *We have:*

(i) Φ_T *is faithful, as isometric: for any* $f \in C(\sigma(T))$*,* $||\Phi_T(f)|| = ||f||_\infty$*;*
(ii) *if, for* $A \in \mathfrak{B}(\mathsf{H})$*,* $AT = TA$ *and* $AT^* = T^*A$*, then* $A\Phi_T(f) = \Phi_T(f)A$ *for any* $f \in C(\sigma(T))$*;*
(iii) Φ_T *preserves the involution:* $\Phi_T(\overline{f}) = \Phi_T(f)^*$ *for any* $f \in C(\sigma(T))$*.*

(c) $\sigma(\Phi_T(f)) = f(\sigma(T), \overline{\sigma(T)})$*, for any* $f \in C(\sigma(T))$*.*

One consequence is worth making explicit.

Corollary 8.38. *Let* H *be a Hilbert space and* $T \in \mathfrak{B}(\mathsf{H})$ *a normal operator. Consider the isometric* **-homomorphism* $\Phi_T : C(\sigma(T)) \to \mathfrak{B}(\mathsf{H})$ *defined in Proposition 8.37. Then* $\Phi_T(C(\sigma(T)))$*, set of continuous functions in the variables* T*,* T^* *(defined on* $\sigma(T)$*) is the smallest* C^**-subalgebra with unit in* $\mathfrak{B}(\mathsf{H})$ *containing* I *and* T*.*

Proof. Every C^*-subalgebra with unit $\mathfrak{A}$ of $\mathfrak{B}(\mathsf{H})$ containing I and T must contain polynomials in T, T^* (restricted to $\sigma(T)$). The construction that led to Φ_T shows $\mathfrak{A}$ contains all continuous maps in T and T^*, i.e. $\Phi_T(C(\sigma(T)))$. The latter, being the image of a C^*-algebra with unit under an injective *-homomorphism, is a C^*-subalgebra with unit of $\mathfrak{B}(\mathsf{H})$ ((b) in Theorem 8.22). □

The fact that we are now working with a concrete C^*-algebra of operators allows to make a further step forward in functional calculus. We can generalise the above theorem by defining $f(T,T^*)$ when f is a bounded measurable, not necessarily continuous, map. In order to do so, in the lack of a Stone–Weierstrass-type theorem for bounded measurable functions ($C(\mathsf{X})$ is not dense in $M_b(\mathsf{X})$ if X is compact with non-empty interior in $\mathbb{R}^n$, cf. Remark 2.26(4)), we shall use heavily Riesz's representation results (for Hilbert spaces and Borel measures).

Recall that on a topological space X, $\mathscr{B}(\mathsf{X})$ is the Borel σ-algebra on X. The C^*-algebra of bounded measurable maps $f : \mathsf{X} \to \mathbb{C}$ is indicated with $M_b(\mathsf{X})$ (Examples 2.26(3) and 3.44(1)).

Proposition 8.37 can be generalised to prove the existence and uniqueness of a *-homomorphism of C^*-algebras with unit between $M_b(\sigma(T))$ and $\mathfrak{B}(\mathsf{H})$ (the topology on $\sigma(T)$ is induced by $\mathbb{C} \supset \sigma(T)$). The next theorem has a host of consequences. It will, in particular, be a crucial ingredient to prove the existence of the spectral measure, in Theorem 8.54. Statement (iii) in (b) will be completed by Theorem 9.9.

Theorem 8.39 (Functional calculus for bounded measurable functions of normal operators). *Let* H *be a Hilbert space and* $T \in \mathfrak{B}(\mathsf{H})$ *a normal operator.*

(a) *There is a unique representation of the commutative C^*-algebra with unit $M_b(\sigma(T))$ (with respect to the norm $|| \, ||_\infty$) on H:*

$$\widehat{\Phi}_T : M_b(\sigma(T)) \ni f \mapsto f(T,T^*) \in \mathfrak{B}(\mathsf{H}),$$

such that:

(i) *if z is the polynomial $\sigma(a) \ni (z,\overline{z}) \mapsto z$,*
$$\widehat{\Phi}_T(z) = T\,; \tag{8.24}$$

(ii) *if $\{f_n\}_{n\in\mathbb{N}} \subset M_b(\sigma(T))$ is bounded and converges pointwise to $f : \sigma(T) \to \mathbb{C}$, then*
$$\widehat{\Phi}_T(f) = w\text{-}\lim_{n\to+\infty} \widehat{\Phi}_T(f_n)\,.$$

(b) *$\widehat{\Phi}_T$ enjoys these properties:*

(i) *the restriction of $\widehat{\Phi}_T$ to $C(\sigma(T))$ is the *-homomorphism Φ_T of Proposition 8.37;*
(ii) *for any $f \in M_b(\sigma(T))$, $||\widehat{\Phi}_T(f)|| \leq ||f||_\infty$;*
(iii) *with $A \in \mathfrak{B}(\mathsf{H})$, if $AT = TA$ and $AT^* = T^*A$ then $A\widehat{\Phi}_T(f) = \widehat{\Phi}_T(f)A$ for any $f \in M_b(\sigma(T))$;*
(iv) *$\widehat{\Phi}_T$ preserves the involution: $\widehat{\Phi}_T(\overline{f}) = \widehat{\Phi}_T(f)^*$ for any $f \in M_b(\sigma(T))$;*
(v) *if $\{f_n\}_{n\in\mathbb{N}} \subset M_b(\sigma(T))$ is bounded and converges pointwise to $f : \sigma(T) \to \mathbb{C}$, then*
$$\widehat{\Phi}_T(f) = s\text{-}\lim_{n\to+\infty} \widehat{\Phi}_T(f_n)\,;$$
(vi) *if $f \in M_b(\sigma(T))$ takes only real values and $f \geq 0$, then $\widehat{\Phi}_T(f) \geq 0$.*

Proof. (a) Fix $x, y \in \mathsf{H}$. The map

$$L_{x,y} : C(\sigma(T)) \ni f \mapsto (x|\Phi_T(f)y) \in \mathbb{C}$$

is linear and $||L_{x,y}||$ is given by:

$$\sup\{|L_{x,y}(f)| \mid f \in C(\sigma(T))\,,\ ||f||_\infty = 1\}$$
$$\leq ||x||\,||y|| \sup\{||\Phi_T(f)|| \mid f \in C(\sigma(T))\,,\ ||f||_\infty = 1\}$$

(Cauchy-Schwarz was used). Since Φ_T is isometric we find

$$||L_{x,y}|| \leq ||x||\,||y||\,,$$

so $L_{x,y}$ is bounded.

By Theorem 2.48 (Riesz's representation theorem for complex measures) there exists a unique complex measure $\mu_{x,y}$ (Definition 1.80) on the compact set $\sigma(T) \subset \mathbb{C}$, such that for any $f \in C(\sigma(T))$:

$$L_{x,y}(f) = (x|\Phi_T(f)y) = \int_{\sigma(T)} f(\lambda)\, d\mu_{x,y}(\lambda)\,. \tag{8.25}$$

Moreover, $|\mu_{x,y}|(\sigma(T)) = ||L_{x,y}|| \leq ||x|| \, ||y||$. Aside, note that $x = y$ forces $\mu_{x,x}$ to be a real, positive, finite measure: in fact, if $f \in C(\sigma(T))$ is real-valued $\Phi_T(f) = \Phi_T(f)^*$ by part (iii) of Proposition 8.37(b), so

$$\int_{\sigma(T)} f(\lambda)\,\overline{h(\lambda)} d|\mu_{x,x}(\lambda)| = \overline{\int_{\sigma(T)} f(\lambda)\, h(\lambda) d|\mu_{x,x}(\lambda)|} = \overline{(x|\Phi_T(f)x)}$$
$$= (\Phi_T(f)x|x) = (x|\Phi_T(f)x) = \int_{\sigma(T)} f(\lambda)\, h(\lambda) d|\mu_{x,x}(\lambda)|\,,$$

where we have decomposed $d\mu_{x,x}$ into $hd|\mu_{x,x}|$, h being a measurable map of unit norm determined, almost everywhere, by $\mu_{x,x}$ (Theorem 1.86), and $|\mu_{x,x}|$ being the positive finite measure associated to $\mu_{x,x}$ called the *total variation* (Remark 1.81(2)). By linearity

$$\int_{\sigma(T)} f(\lambda)\,\overline{h(\lambda)} d|\mu_{x,x}(\lambda)| = \int_{\sigma(T)} f(\lambda)\, h(\lambda) d|\mu_{x,x}(\lambda)|$$

must hold when $f \in C(\sigma(T))$ is complex-valued. Riesz's Theorem 2.48 on complex measures guarantees $hd|\mu_{x,x}| = \overline{h}d|\mu_{x,x}|$, so $\overline{h(\lambda)} = h(\lambda)$ almost everywhere; but $|h(\lambda)| = 1$, so $h(\lambda) = 1$ almost everywhere, and hence $\mu_{x,x}$ is a real, positive and finite measure (so is $|\mu_{x,x}|$).

Use (8.25) to generalise $L_{x,y}(f)$ to the case $f \in M_b(\sigma(T))$, since the right-hand side is well defined anyway: if $g \in M_b(\sigma(T))$,

$$L_{x,y}(g) := \int_{\sigma(T)} g(\lambda)\, d\mu_{x,y}(\lambda)\,. \tag{8.26}$$

By general properties of complex measures (cf. Example 2.45(1)):

$$|L_{x,y}(g)| \leq ||g||_\infty |\mu_{x,y}|(\sigma(T)) \leq ||g||_\infty ||x|| \, ||y||\,. \tag{8.27}$$

By construction, given $g \in C(\sigma(T))$, $(x,y) \mapsto L_{x,y}(g)$ is antilinear in x and linear in y. One can prove this is still valid for $g \in M_b(\sigma(T))$. Let us for instance show linearity in y, the other case being similar. Given $x,y,z \in \mathsf{H}$ and $g \in M_b(\sigma(T))$, if $\alpha, \beta \in \mathbb{C}$ then

$$\alpha \int_{\sigma(T)} g(\lambda)\, d\mu_{x,y}(\lambda) + \beta \int_{\sigma(T)} g(\lambda)\, d\mu_{x,z}(\lambda) = \int_{\sigma(T)} g(\lambda)\, d\nu(\lambda)\,, \tag{8.28}$$

where ν is the complex measure $\nu(E) := \alpha\mu_{x,y}(E) + \beta\mu_{x,z}(E)$ for any Borel set $E \subset \sigma(T)$. Remembering how we defined the $\mu_{x,y}$ (cf. (8.25)) and using the inner product's linearity on the right, we immediately see that for any $f \in C(\sigma(T))$ replacing g in (8.28) gives:

$$\int_{\sigma(T)} f(\lambda)\, d\mu_{x,\alpha y+\beta z}(\lambda) = \int_{\sigma(T)} f(\lambda)\, d\nu(\lambda)\,.$$

Riesz's theorem now tells $\mu_{x,\alpha y+\beta z} = \nu$. Thus (8.28) reads, *for any* $g \in M_b(\sigma(T))$:

$$\alpha \int_{\sigma(T)} g(\lambda)\, d\mu_{x,y}(\lambda) + \beta \int_{\sigma(T)} g(\lambda)\, d\mu_{x,z}(\lambda) = \int_{\sigma(T)} g(\lambda)\, \mu_{x,\alpha y+\beta z}(\lambda)\,.$$

We proved $L_{x,y}(g)$ is linear in y for any given $x \in \mathsf{H}$ and any $g \in M_b(\sigma(T))$.

Equation (8.27) implies the linear operator $y \mapsto L_{x,y}(g)$ is bounded, so by Theorem 3.16 (Riesz's once again), given $g \in M_b(\sigma(T))$ and $x \in \mathsf{H}$, there exists a unique $v_x \in \mathsf{H}$ such that $L_{x,y}(g) = (v_x|y)$ for any $y \in \mathsf{H}$. Since v_x is linear in x ($L_{x,y}(g)$ is antilinear in x and the inner product $(v_x|y)$ is antilinear in v_x), there is also a unique operator $g(T,T^*)' \in \mathfrak{L}(\mathsf{H})$ such that $v_x = g(T,T^*)'x$ for any $x \in \mathsf{H}$. Hence $L_{x,y}(g) = (g(T,T^*)'x|y)$. Condition (8.27) implies $g(T,T^*)'$ is bounded, for:

$$||g(T,T^*)'x||^2 = |(g(T,T^*)'x|g(T,T^*)'x)|$$
$$= |L_{x,g(T,T^*)'x}(g)| \le ||g||_\infty \, ||x|| \, ||g(T,T^*)'x|| \,,$$

hence

$$\frac{||g(T,T^*)'x||}{||x||} \le ||g||_\infty$$

and then $||g(T,T^*)'|| \le ||g||_\infty$.

Setting $g(T,T^*) := g(T,T^*)'^*$, we proved that for $g \in M_b(\sigma(T))$ there is a unique operator $g(T,T^*) \in \mathfrak{B}(\mathsf{H})$ such that

$$L_{x,y}(g) = (x|g(T,T^*)y)$$

for any $x,y \in \mathsf{H}$. The linear mapping

$$\widehat{\Phi}_T : M_b(\sigma(T)) \ni f \mapsto f(T,T^*) \in \mathfrak{B}(\mathsf{H}) \,,$$

where, for any $x,y \in \mathsf{H}$,

$$L_{x,y}(f) = (x|f(T,T^*)y) := \int_{\sigma(T)} f(\lambda)\, d\mu_{x,y}(\lambda) \,,$$

is, by construction, an extension of Φ_T: in particular (8.24) holds. The extension is continuous because $||\widehat{\Phi}_T(f)|| \le ||f||_\infty$ for any $f \in M_b(\sigma(T))$, in fact:

$$||\widehat{\Phi}_T(f)|| = ||f(T,T^*)|| = ||f(T,T^*)'^*|| = ||f(T,T^*)'|| \le ||f||_\infty \,.$$

As $\widehat{\Phi}_T$ extends the algebra homomorphism Φ_T, to prove $\widehat{\Phi}_T$ is an algebra homomorphism it suffices to show $\widehat{\Phi}_T(f \cdot g) = \widehat{\Phi}_T(f)\widehat{\Phi}_T(g)$ when $f,g \in M_b(\sigma(T))$. If the two maps belong in $C(\sigma(T))$, the claim is true by Proposition 8.37 above. Suppose $f,g \in C(\sigma(T))$. Then

$$\int_{\sigma(T)} f \cdot g \, d\mu_{x,y} = (x|\widehat{\Phi}_T(f \cdot g)y) = (x|\widehat{\Phi}_T(f)\widehat{\Phi}_T(g)y) = \int_{\sigma(T)} f \, d\mu_{x,\widehat{\Phi}_T(g)y} \,.$$

The mentioned theorem of Riesz on complex measures implies that $d\mu_{x,\widehat{\Phi}_T(g)y}$ coincides with $g\, d\mu_{x,y}$. Thus, if $f \in M_b(\sigma(T))$,

$$\int_{\sigma(T)} f \cdot g \, d\mu_{x,y} = \int_{\sigma(T)} f \, d\mu_{x,\widehat{\Phi}_T(g)y}.$$

From this follows, for any $x,y \in \mathsf{H}$, $f \in M_b(\sigma(T))$ and $g \in C(\sigma(T))$:

$$\int_{\sigma(T)} f \cdot g \, d\mu_{x,y} = \int_{\sigma(T)} f \, d\mu_{x,\widehat{\Phi}_T(g)y} = (x|\widehat{\Phi}_T(f)\widehat{\Phi}_T(g)y) = (\widehat{\Phi}_T(f)^* x|\widehat{\Phi}_T(g)y)$$

$$= \int_{\sigma(T)} g \, d\mu_{\widehat{\Phi}_T(f)^* x,y} \, .$$

Arguing as before and with Riesz's theorem, the equality

$$\int_{\sigma(T)} f \cdot g \, d\mu_{x,y} = \int_{\sigma(T)} g \, d\mu_{\widehat{\Phi}_T(f)^* x,y} \, , \tag{8.29}$$

valid for any $g \in C(\sigma(T))$, forces $f d\mu_{x,y} = d\mu_{\widehat{\Phi}_T(f)^* x,y}$, so (8.29) must hold for any $x,y \in \mathsf{H}$, and any $f,g \in M_b(\sigma(T))$. Therefore

$$(x|\widehat{\Phi}_T(f \cdot g)y) = \int_{\sigma(T)} f \cdot g \, d\mu_{x,y} = \int_{\sigma(T)} g \, d\mu_{\widehat{\Phi}_T(f)^* x,y}$$

$$= (\widehat{\Phi}_T(f)^* x|\widehat{\Phi}_T(g)y) = (x|\widehat{\Phi}_T(f)\widehat{\Phi}_T(g)y) \, ,$$

and consequently

$$\left(x \middle| (\widehat{\Phi}_T(f \cdot g) - \widehat{\Phi}_T(f)\widehat{\Phi}_T(g))y \right) = 0 \, .$$

Choosing x as the second argument in the inner product gives

$$\widehat{\Phi}_T(f \cdot g)y = \widehat{\Phi}_T(f)\widehat{\Phi}_T(g)y$$

for any $y \in \mathsf{H}$, $f,g \in M_b(\sigma(T))$, whence

$$\widehat{\Phi}_T(f \cdot g) = \widehat{\Phi}_T(f)\widehat{\Phi}_T(g) \, .$$

To show we have indeed a *-homomorphism we need to prove property (iv). Let $x \in \mathsf{H}$ and $g \in M_b(\sigma(T))$. Since $\mu_{x,x}$ is real, we have (beware that complex conjugation does not act on $\sigma(T)$, here thought of as subset in $\mathbb{R}^2$):

$$(x|\widehat{\Phi}_T(\overline{g})x) = \int_{\sigma(T)} \overline{g} \, d\mu_{x,x} = \overline{\int_{\sigma(T)} g \, d\mu_{x,x}} = (\widehat{\Phi}_T(g)x|x) = (x|\widehat{\Phi}_T(g)^* x) \, .$$

Hence $(x|(\widehat{\Phi}_T(\overline{g}) - \widehat{\Phi}_T(g)^*)x) = 0$ for any $x \in \mathsf{H}$. From Exercise 3.18 we have $\widehat{\Phi}_T(\overline{g}) = \widehat{\Phi}_T(g)^*$.

Property (ii) of (a) follows from (v) in (b), which we will prove below independently.

To finish (a), we show $\widehat{\Phi}_T$ is unique under (a). Let $\Psi : M_b(\sigma(T)) \to \mathfrak{B}(\mathsf{H})$ satisfy (a). It must coincide with $\widehat{\Phi}_T$ on polynomials, so by continuity (it is continuous being a *-homomorphism of C^*-algebras with unit, and Theorem 8.22 holds) it coincides with $\widehat{\Phi}_T$ on $C(\sigma(T))$. Given $x,y \in \mathsf{H}$, the map

$$\nu_{x,y} : E \mapsto (x|\Psi(\chi_E)y) \, ,$$

where E is an arbitrary Borel set in $\sigma(T)$ and χ_E its characteristic function, is a complex measure on $\sigma(T)$. In fact $\nu_{x,y}(\varnothing) = (x|\Psi(0)y) = 0$; moreover, if $\{S_k\}_{k\in\mathbb{N}}$ is a family of pairwise disjoint Borel sets,

$$\nu_{x,y}(\cup_k S_k) = (x|\Psi(\chi_{\cup_k S_k})y) = \left(x\left|\lim_{n\to+\infty}\Psi\left(\sum_{k=0}^{n}\chi_{S_k}\right)y\right.\right) = \lim_{n\to+\infty}\sum_{k=0}^{n}(x|\Psi(\chi_{S_k})y)$$

$$= \sum_{k=0}^{+\infty}\nu_{x,y}(S_k)\,,$$

where the left-hand side is always finite, we used (ii) in (a) and that, pointwise:

$$\chi_{\cup_k S_k} = \sum_{k=0}^{+\infty}\chi_{S_k}\,. \tag{8.30}$$

Observe that (8.30) does not depend on the labelling order of the S_k, for the series has positive terms. Consequently

$$\nu_{x,y}(\cup_k S_k) = \sum_{k=0}^{+\infty}\nu_{x,y}(S_k)$$

holds irrespective of the series' ordering, and the series converges absolutely (Theorem 1.82). This means $\nu_{x,y}$ is a complex measure.

Bearing in mind Ψ's and the inner product's linearity, plus the definition of integral of a simple map, we easily see

$$\int_{\sigma(T)} s\, d\nu_{x,y} = (x|\Psi(s)y)$$

for any simple map $s \in S(\sigma(T))$. If $f \in M_b(\sigma(T))$ and $\{s_n\} \subset S(\sigma(T))$ converges uniformly to f (the sequence exists by Proposition 7.49(b)), then by Ψ's continuity in norm $||\ ||_\infty$ and by dominated convergence relative to $|\nu_{x,y}S|$, we have

$$(x|\Psi(f)y) = \int_{\sigma(T)} f\, d\nu_{x,y} \tag{8.31}$$

for any $f \in M_b(\sigma(T))$. In particular, this must hold for $f \in C(\sigma(T))$, on which Ψ coincides with $\widehat{\Phi}_T$. Therefore, Riesz's Theorem 2.48 on complex measures implies that $\nu_{x,y}$ coincides with the complex measure $\mu_{x,y}$ of the beginning, using which we defined $\widehat{\Phi}_T$ by

$$(x|\widehat{\Phi}_T(f)y) = \int_{\sigma(T)} f\, d\mu_{x,y}\,,$$

for $x, y \in \mathsf{H}$ and $f \in M_b(\sigma(T))$. But then (8.31) implies $\Psi(f) = \widehat{\Phi}_T(f)$ for any $f \in M_b(\sigma(T))$, for $\nu_{x,y} = \mu_{x,y}$.

(b) We only need to prove (iii), (v) and (vi), because the rest were shown in part (a).

Property (iii) holds when $f \in C(\sigma(T))$, as we know from Proposition 8.37(b). If $AT = TA$ and $AT^* = T^*A$,

$$\int_{\sigma(T)} f\, d\mu_{x,Ay} = (x|\widehat{\Phi}_T(f)Ay) = (x|A\widehat{\Phi}_T(f)y) = (A^*x|\widehat{\Phi}_T(f)y) = \int_{\sigma(T)} f\, d\mu_{A^*x,y}\,,$$

for any vectors $x,y \in \mathsf{H}$ and any $f \in C(\sigma(T))$. Riesz's Theorem 2.48 on the representation of complex measures on Borel sets ensures $\mu_{A^*x,y} = \mu_{x,Ay}$, hence

$$(x|\widehat{\Phi}_T(f)Ay) = \int_{\sigma(T)} f\, d\mu_{x,Ay} = \int_{\sigma(T)} f\, d\mu_{A^*x,y} = (A^*x|\widehat{\Phi}_T(f)y) = (x|A\widehat{\Phi}_T(f)y)$$

for any $x,y \in \mathsf{H}$, $f \in M_b(\sigma(T))$. As the vectors x,y are arbitrary, $\widehat{\Phi}_T(f)A = A\widehat{\Phi}_T(f)$ if $f \in M_b(\sigma(T))$.

Let us prove (v), so take a sequence $\{f_n\}_{n\in\mathbb{N}} \subset M_b(\sigma(T))$, bounded (in absolute value) by $K > 0$, that converges to $f : \sigma(T) \to \mathbb{C}$. So $||f||_\infty \leq K$ and f is measurable, forcing $f \in M_b(\sigma(T))$. Given $x,y \in \mathsf{H}$ and using (iv) in (b),

$$||(\widehat{\Phi}_T(f_n) - \widehat{\Phi}_T(f))x||^2 = ((\widehat{\Phi}_T(f_n) - \widehat{\Phi}_T(f))x|(\widehat{\Phi}_T(f_n) - \widehat{\Phi}_T(f))x)$$

$$= (x|(\widehat{\Phi}_T(f_n - f)^*\widehat{\Phi}_T(f_n - f)x) = (x|\widehat{\Phi}_T(|f - f_n|^2)x)\,.$$

The last terms can be written as

$$\int_{\sigma(T)} |f - f_n|^2\, d\mu_{x,x} = \int_{\sigma(T)} |f - f_n|^2\, hd|\mu_{x,x}|\,,$$

where $|\mu_{x,x}|$ is the positive measure (the total variation of Remark 1.81(2)) associated the real (signed) measure $\mu_{x,x}$, and h is a measurable function of constant modulus 1 (Theorem 1.86). (Actually, we saw in part (a) that $\mu_{x,x}$ is a positive real measure, so $|\mu_{x,x}| = \mu_{x,x}$ and $h = 1$.) Because

$$|\mu_{x,x}|(\sigma(T)) < +\infty\,,$$

the dominated convergence theorem implies $|h||f - f_n|^2$ converges to 0 in $L^1(\sigma(T), |\mu_{x,x}|)$. Hence as $n \to +\infty$, $||(\widehat{\Phi}_T(f_n) - \widehat{\Phi}_T(f))x||^2 \to 0$ for any $x \in \mathsf{H}$.

Eventually, let us prove (vi). The proof is easy and follows from Theorem 8.25(iii), but here is an alternative argument. If $M_b(\sigma(T)) \ni f \geq 0$, then $f = g^2$ where $0 \leq g \in M_b(\sigma(T))$. By (a), $\widehat{\Phi}_T(f) = \widehat{\Phi}_T(g \cdot g) = \widehat{\Phi}_T(g)\widehat{\Phi}_T(g)$. Moreover, $\widehat{\Phi}_T(g)^* = \widehat{\Phi}_T(\overline{g}) = \widehat{\Phi}_T(g)$ (by (iv)), so $\widehat{\Phi}_T(g \cdot g) = \widehat{\Phi}_T(g)^*\widehat{\Phi}_T(g)$. The right-hand side is patently positive. □

Remarks 8.40. The spectral decomposition theorem, proved later, is in some sense a way to interpret the operator $f(T,T^*)$ in terms of an integral of f with respect to an operator-valued measure: integrating bounded measurable functions produces, instead of numbers, operators. The version of the spectral decomposition theorem presented in this chapter states that there is always such a measure, for any bounded normal operator. ■

8.3 Projector-valued measures (PVMs)

In this section we introduce *projector-valued measures* (PVM), also called *spectral measures*. They are the central tool to state spectral theorems, and represent a generalisation of the notion of measure on the Borel σ-algebra of a topological space X, where now the measure's range is not in $\mathbb{R}$, but rather a subset of orthogonal projectors $\mathfrak{P}(\mathsf{H})$ in a Hilbert space H:

$$\mathscr{B}(\mathsf{X}) \ni E \mapsto P(E) \in \mathfrak{P}(\mathsf{H}) .$$

Thereby we will be able to *integrate functions to obtain operators*. We will see, in particular, that the homomorphism $\widehat{\Phi}_T$, studied in the previous section and associated to a bounded normal operator T, is nothing else than an integral with respect to a PVM generated by T:

$$\widehat{\Phi}_T(f) = \int_{\sigma(T)} f(x) dP^{(T)}(x) .$$

Projector-valued measures made their appearance already in Chapter 7 (Definition 7.46), in the special case where the σ-algebra of the PVM was $\mathscr{B}(\mathbb{R})$. A quantum *observable*, in the sense of the previous chapter, is a special spectral measure, by virtue of Proposition 7.44. In that case the operator to which such a PVM is attached is not just normal, but self-adjoint as well.

8.3.1 Spectral measures, or PVMs

We remind that for $T, U \in \mathfrak{B}(\mathsf{H})$, H Hilbert space, $T \geq U$ means $(x|Tx) \geq (x|Ux)$ for any $x \in \mathsf{H}$ (see Definition 3.51(f) and the ensuing comments).

Definition 8.41. *If* H *is a Hilbert space,* $(\mathsf{X}, \mathscr{T})$ *a second-countable space and* $\mathscr{B}(\mathsf{X})$ *the Borel* σ*-algebra on* X*,* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ *is called* **spectral measure** *on* X*, or equivalently* **projector-valued measure on** X *(PVM), if the following requisites are satisfied.*

(a) $P(B) \geq 0$ *for any* $B \in \mathscr{B}(\mathsf{X})$.
(b) $P(B)P(B') = P(B \cap B')$ *for any* $B, B' \in \mathscr{B}(\mathsf{X})$.
(c) $P(\mathsf{X}) = I$.
(d) *if* $\{B_n\}_{n\in\mathbb{N}} \subset \mathscr{B}(\mathsf{X})$*, with* $B_n \cap B_m = \varnothing$*,* $n \neq m$*:*

$$s\text{-}\sum_{n=0}^{+\infty} P(B_n) = P(\cup_{n\in\mathbb{N}} B_n) .$$

The **support of** P *is the closed set*

$$supp(P) := \mathsf{X} \setminus \bigcup_{A\in\mathscr{T}, P(A)=0} A .$$

When $\mathsf{X} = \mathbb{R}^n$ *or* $\mathbb{C}^n$*,* P *is called* **bounded** *if* $supp(P)$ *is bounded.*

Dropping (b), what remains in Definition 8.41 gives a POVM, a *positive operator-valued measure*, that we will discuss in Chapter 13. Another related and useful definition is the following.

Definition 8.42. *If* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ *is a PVM, a measurable function* $f : \mathsf{X} \to \mathbb{C}$ *is said* **essentially bounded for** P *when*

$$P(\{x \in \mathsf{X} \mid |f(x)| \geq M\}) = 0 \quad \textit{for some } M < +\infty. \tag{8.32}$$

If f *is essentially bounded, the greatest lower bound* $||f||_\infty^{(P)}$ *on the set of* $M \geq 0$ *satisfying (8.32) is called* **essential (semi)norm** *of* f *in* P.

The next proposition treats the basic properties of PVMs. In particular, as the name PVM itself suggests, $P(E) \in \mathfrak{P}(\mathsf{H})$, where as usual $\mathfrak{P}(\mathsf{H})$ are the orthogonal projectors on the Hilbert space H.

Proposition 8.43. *Retaining Definition 8.41, the following facts hold.*

(a) $P(B) \in \mathfrak{P}(\mathsf{H})$ *for any* $B \in \mathscr{B}(\mathsf{X})$. *Conditions (a) and (b) in Definition 8.41 may be replaced by the equivalent requirement that* $P(B)$ *is an orthogonal projector if* $B \in \mathscr{B}(\mathsf{X})$.
(b) P *is* **monotone**: $P(C) \leq P(B)$ *for* $B, C \in \mathscr{B}(\mathsf{X})$, $C \subset B$.
(c) P *is* **sub-additive**: *if* $B_n \in \mathscr{B}(\mathsf{X})$, $n \in \mathbb{N}$, *then*

$$(x|P(\cup_{n\in\mathbb{N}} B_n)\, x) \leq \sum_{n\in\mathbb{N}} (x|P(B_n)x) \quad \textit{for any } x \in \mathsf{H}.$$

(d) $P(supp(P)) = I$, *so* P *is concentrated on* $supp(P)$, *i.e.*

(i) $P(B) = P(B \cap supp(P))$ *for* $B \in \mathscr{B}(\mathsf{X})$;
(ii) $P(C) = 0$ *if either* $C = \varnothing$ *or* $C \subset \mathsf{X} \setminus supp(P)$ *and* $C \in \mathscr{B}(\mathsf{X})$.

Proof. (a) The operators $P(B)$ are idempotent, $P(B)P(B) = P(B \cap B) = P(B)$, by Definition 8.41(b), and self-adjoint because bounded and positive (by Definition 8.41(a)), so they are orthogonal projectors. We claim that if (c) and (d) in Definition 8.41 hold, and every $P(B)$ is an orthogonal projector, then also (a) and (b) hold in Definition 8.41. Part (a) is trivial, for any $P(E)$ is an orthogonal projector, hence positive. By part (d), if $E_1, E_2 \in \mathscr{B}(\mathsf{X})$ and $E_1 \cap E_2 = \varnothing$ then $P(E_1 \cup E_2) = P(E_1) + P(E_2)$. Multiplying by $P(E_1 \cup E_2) = P(E_1) + P(E_2)$, and recalling we are using (idempotent) projectors, gives $P(E_1)P(E_2) + P(E_2)P(E_1) = 0$ and so $P(E_1)P(E_2) = -P(E_2)P(E_1)$; applying now $P(E_1)$ and recalling that $P(E_1)P(E_2) = -P(E_2)P(E_1)$, we find $P(E_1)P(E_2) = P(E_2)P(E_1)$. Therefore $P(E_1)P(E_2) = \frac{1}{2}(P(E_1)P(E_2) + P(E_2)P(E_1)) = 0$ if $E_1 \cap E_2 = \varnothing$. Now set $C = B \cap B'$, $E_1 = B \setminus C$, $E_2 = B' \setminus C$ for $B, B' \in \mathscr{B}(\mathsf{X})$. Remember $E_1 \cap E_2 = E_1 \cap C = E_2 \cap C = \varnothing$. Then

$$P(B)P(B') = (P(E_1) + P(C))(P(E_2) + P(C)) = P(C)P(C) = P(C) = P(B \cap B')$$

i.e. property (b) in 8.41.
(b) $B = C \cup (B \setminus C)$ and $C \cap (B \setminus C) = \varnothing$ so by Definition 8.41(d) follows $P(B) = P(C) + P(B \setminus C)$. But $P(B \setminus C) \geq 0$, so $P(C) \leq P(B)$.

(c) Define $B := \cup_{n\in\mathbb{N}} B_n$ and the sequence $\{C_n\}_{n\in\mathbb{N}}$, with $C_0 := B_0$, $C_1 := B_1 \setminus B_0$, $C_2 := B_2 \setminus (B_0 \cup B_1)$ and so on. Clearly $C_k \cap C_h = \varnothing$ if $h \neq k$ and $B = \cup_{n\in\mathbb{N}} C_n$. By Definition 8.41(d), then, $P(B) = s\text{-}\sum_{n=0}^{+\infty} P(C_k)$ and thus $(x|P(B)x) = \sum_{n=0}^{+\infty}(x|P(C_k)x)$. Since $C_k \subset B_k$ for any $k \in \mathbb{N}$, by monotonicity $(x|P(C_k)x) \leq (x|P(B_k)x)$, i.e. $(x|P(B)x) \leq \sum_{n=0}^{+\infty}(x|P(B_k)x)$.
(d) $P(supp(P)) = I$ is obviously equivalent to (Definition 8.41(d) in the finite case) to $P(A) = 0$, where $A := \mathsf{X} \setminus supp(P)$. To prove the latter, notice that by definition A is the union of open sets with null spectral measure. As X is second-countable, Lindelöf's lemma (Theorem 1.8) says we can extract a countable subcovering. Put differently, $A = \cup_{n\in\mathbb{N}} A_n$ with $P(A_n) = 0$ for any $n \in \mathbb{N}$. Using sub-additivity, for any $x \in \mathsf{H}$,

$$0 \leq ||P(A)x||^2 = (P(A)x|P(A)x) = (x|P(A)x) \leq \sum_{n\in\mathbb{N}} (x|P(A_n)x) = 0\,,$$

hence $P(A) = 0$. Property (ii) is immediate by monotonicity (note $\varnothing \subset \mathsf{X} \setminus supp(P)$ and $\varnothing \in \mathscr{B}(\mathsf{X})$). With A defined as above, property (i) is a consequence of writing $B = (B \cap supp(P)) \cup (B \cap A)$: Definition 8.41(d) in fact gives $P(B) = P(B \cap supp(P)) + P(B \cap A)$, and we can use (ii). □

Remark 8.44. **(1)** If $f : \mathsf{X} \to \mathbb{C}$ is measurable, property (ii) in Proposition 8.43(d) implies immediately the first inequality below:

$$||f||_\infty^{(P)} \leq ||f\restriction_{supp(P)}||_\infty \leq ||f||_\infty \tag{8.33}$$

(the second one is obvious). Equation (8.33) holds trivially when one among $||f||_\infty^{(P)}, ||f\restriction_{supp(P)}||_\infty, ||f||_\infty$ is $+\infty$.
(2) The set of measurable functions that are essentially bounded for P is a vector space, and $||\ ||_\infty^{(P)}$ is a *seminorm* on it.
(3) In this text we will only work with PVMs defined on Borel σ-algebras $\mathscr{B}(\mathsf{X})$ coming from second-countable spaces X. This is not strictly necessary, and almost the entire story could be developed using general σ-algebras (see for instance [Rud91]). Our choice is only motivated by simplicity. First of all, the *support* of the PVM satisfies, thus, the properties established by Proposition 8.43(d) (which justify the notion of support), showing the spectral measure is *concentrated* on it (in analogy to Proposition 1.45 for σ-additive positive measures). Secondly, the spectral decomposition theorem is stated over $\mathbb{C}$ (or $\mathbb{R}$), whose topology is second-countable. ■

8.3.2 Integrating bounded measurable functions in a PVM

We pass now to define a procedure to integrate bounded measurable functions $f : \mathsf{X} \to \mathbb{C}$ *with respect to a projector-valued measure* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$.

Recall that given a space X with a σ-algebra Σ, a (complex-valued) map $s : \mathsf{X} \to \mathbb{C}$, measurable for Σ, is called **simple** when its range is finite.

Notation 8.45. If X is a topological space, $S(\mathsf{X})$ denotes the vector space of complex-valued simple functions on X, relative to the Borel σ-algebras $\mathscr{B}(\mathsf{X})$. ■

Let a PVM be given, on a second-countable space X, with values in $\mathfrak{B}(\mathsf{H})$ for some Hilbert space H. Consider a map $s \in S(\mathsf{X})$. We can always write it, for suitable $c_i \in \mathbb{C}$ and I finite, as follows:

$$s = \sum_{i \in I} c_i \chi_{E_i} \,. \tag{8.34}$$

As, by definition, the range of a simple function consists of finitely many distinct values, the expression above is uniquely determined by s once we require the sets E_i to be measurable and pairwise disjoint, and that the complex numbers c_i are distinct. We define the **integral of s with respect to P** as the operator in $\mathfrak{B}(\mathsf{H})$:

$$\int_{\mathsf{X}} s(x)\, dP(x) := \sum_{i \in I} c_i P(E_i) \,. \tag{8.35}$$

Remarks 8.46. If we do not insist the above c_i be distinct, there are several ways to write s as a linear combination of characteristic functions of disjoint measurable sets. Using the same argument as for an ordinary measure it is easy to prove, however, that the integral of s does not depend on the particular representation of s chosen. ■

The mapping

$$\mathfrak{I} : S(\mathsf{X}) \ni s \mapsto \int_{\mathsf{X}} s(x)\, dP(x) \in \mathfrak{B}(\mathsf{X}) \,, \tag{8.36}$$

is linear, i.e. $\mathfrak{I} \in \mathfrak{L}(S(\mathsf{X}), \mathfrak{B}(\mathsf{H}))$, as the previous remark easily implies. Since $S(\mathsf{X})$ and $\mathfrak{B}(\mathsf{H})$ are normed spaces, $\mathfrak{L}(S(\mathsf{X}), \mathfrak{B}(\mathsf{H}))$ is equipped with the operator norm. $\mathfrak{I}$ turns out to be a *bounded* operator for this norm. Let us prove this fact, and consider $s \in S(\mathsf{X})$ of the form (8.34). As the E_k are pairwise disjoint, $P(E_j)P(E_i) = P(E_j \cap E_i) = 0$ if $i \neq j$ or $P(E_j)P(E_i) = P(E_i)$ if $i = j$. If $x \in \mathsf{H}$

$$\begin{aligned} ||\mathfrak{I}(s)x||^2 = (\mathfrak{I}(s)x|\mathfrak{I}(s)x) &= \left(\sum_{i \in I} c_i P(E_i)x \,\middle|\, \sum_{j \in I} c_j P(E_j)x \right) \\ &= \sum_{i,j \in I} \left(c_i P(E_j)^* P(E_i)x \middle| \, c_j x \right) = \sum_{i,j \in I} \left(c_i P(E_j) P(E_i)x \middle| \, c_j x \right) \\ &= \sum_{i \in I} |c_i|^2 (x|P(E_i)x) \leq \sup_{i \in I'} |c_i|^2 \sum_{i \in I'} (x|P(E_i)x) \,, \end{aligned}$$

where $I' \subset I$ is made by indices for which $P(E_i) \neq 0$. By additivity and monotonicity

$$\sum_{i \in I'} (x|P(E_i)x) \leq (x|P(\cup_{i \in I'} E_i)x) \leq (x|P(\mathsf{X})x) = (x|x) = ||x||^2 \,.$$

But I' is finite, so trivially $||s||_\infty^P = \sup_{i \in I'} |c_i|$, and hence $||\mathfrak{I}(s)x||^2 \leq ||x||^2 (||s||_\infty^P)^2$. Taking the least upper bound over unit vectors $x \in \mathsf{H}$:

$$||\mathfrak{I}(s)|| \leq ||s||_\infty^{(P)} \,.$$

But $||s||_\infty^{(P)}$ coincides with one of the values of $|s|$, say $|c_k|$ if we choose $x \in P(E_k)(\mathsf{H})$ ($\neq \{0\}$ by construction). Thus by $x = P(E_k)x$ we have

$$\mathfrak{I}(s)x = \sum_{i \in I'} c_i P(E_i)x = \sum_{i \in I'} c_i P(E_i)P(E_k)x = c_k P(E_k)x = c_k x \,.$$

So choosing x with $||x|| = 1$ we obtain $||\mathfrak{I}(s)x|| = ||s||_\infty^{(P)}$. Therefore $\mathfrak{I}$ is certainly continuous on $S(\mathsf{X}) \subset M_b(\mathsf{X})$ in norm $|| \; ||_\infty$, by what we have just proved and by (8.33):

$$||\mathfrak{I}(s)|| = ||s||_\infty^{(P)} \leq ||s \restriction_{supp(P)} ||_\infty \leq ||s||_\infty . \tag{8.37}$$

This settled, we can define integrals of bounded measurable functions, by prolonging $\mathfrak{I}$ by linearity and continuity to the whole Banach space $M_b(\mathsf{X})$ of bounded measurable maps $f : \mathsf{X} \to \mathbb{C}$. $M_b(\mathsf{X})$ contains $S(\mathsf{X})$ as dense subspace in norm $|| \; ||_\infty$, by Proposition 7.49(b). The operator $\mathfrak{I} : S(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ is continuous. By Proposition 2.44 there exists one and only one bounded operator from $M_b(\mathsf{X})$ to $\mathfrak{B}(\mathsf{H})$ extending $\mathfrak{I}$.

Definition 8.47. *Let X be a second-countable space, H a Hilbert space and $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ a projector-valued measure defined on the Borel σ-algebra of X.*

(a) *The unique bounded extension $\widehat{\mathfrak{I}} : M_b(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ of the operator $\mathfrak{I} : S(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ (cf. (8.35)–(8.36)) is called* **integral operator** *in P.*
(b) *For any $f \in M_b(\mathsf{X})$:*

$$\int_{\mathsf{X}} f(x)\, dP(x) := \widehat{\mathfrak{I}}(f)$$

is the **integral of** *f with respect to the projector-valued measure P.*
(c) *Let $f : \mathsf{X} \to \mathbb{C}$ be measurable, not necessarily bounded. If $f \restriction_E \in M_b(E)$ with $E \subset \mathscr{B}(\mathsf{X})$, we define:*

$$\int_E f(x)\, dP(x) := \int_{\mathsf{X}} \chi_E(x) f(x)\, dP(x) .$$

If $g \in M_b(E)$, with $E \subset \mathscr{B}(\mathsf{X})$, we set:

$$\int_E g(x)\, dP(x) := \int_{\mathsf{X}} g_0(x)\, dP(x) ,$$

where $g_0(x) := g(x)$ if $x \in E$, or $g_0(x) := 0$ if $x \notin E$.

Remarks 8.48. If P is a spectral measure on X and $supp(P) \neq \mathsf{X}$, we can *restrict* P to a spectral measure $P\restriction_{supp(P)}$ on $supp(P)$ (with induced topology), by defining $P\restriction_{supp(P)} (E) := P(E)$ for any Borel set $E \subset \mathscr{B}(supp(P))$. The fact that $P\restriction_{supp(P)}$ is a PVM is immediate using Proposition 8.43, especially part (d). From (i) in (d) we have, for any $s \in S(\mathsf{X})$,

$$\int_{\mathsf{X}} s dP = \int_{supp(P)} s dP = \int_{supp(P)} s\restriction_{supp(P)} dP\restriction_{supp(P)} ,$$

where the second integral is understood in the sense of Definition 8.47(c). If $S(\mathsf{X}) \ni s_n \to f$ in norm $|| \; ||_\infty$, then $S(\mathsf{X}) \ni s_n\restriction_{supp(P)} \to f\restriction_{supp(P)}$ in the same norm. Therefore the definition of integral of some $f \in M_b(\mathsf{X})$ with respect to P tells that

$$\int_{\mathsf{X}} f dP = \int_{supp(P)} f dP = \int_{supp(P)} f\restriction_{supp(P)} dP\restriction_{supp(P)} \quad \text{for any } f \in M_b(\mathsf{X}). \tag{8.38}$$

■

Examples 8.49. (1) Let us see a concrete example lest the procedure seem too abstract. The (generalisation of this) example actually covers all possibilities, as we shall explain.

Consider the Hilbert space $\mathsf{H} = L^2(\mathsf{X}, \mu)$, where X is second countable and μ a positive σ-additive measure on the Borel σ-algebra of X. A spectral measure on H arises by defining, for any $\psi \in L^2(\mathsf{X}, \mu)$ and $E \in \mathscr{B}(\mathsf{X})$,

$$(P(E)\psi)(x) := \chi_E(x)\psi(x)\,, \quad \text{for almost every } x \in \mathsf{X}\,. \tag{8.39}$$

The map $\mathscr{B}(\mathsf{X}) \ni E \mapsto P(E)$ easily defines a spectral measure on $L^2(\mathsf{X}, \mu)$. We want to understand what the operators $\int_{\mathsf{X}} f(x)\,dP(x)$ look like, for any map of $M_b(\mathsf{X})$.

If $\psi \in L^2(\mathsf{X}, \mu)$ and $f \in M_b(\mathsf{X})$, then $f \cdot \psi \in L^2(\mathsf{X}, \mu)$, where $\cdot$ is the pointwise product of maps, for:

$$\int_{\mathsf{X}} |f(x)\psi(x)|^2\,d\mu(x) \leq ||f||_\infty^2 \int_{\mathsf{X}} |\psi(x)|^2\,d\mu(x) < +\infty\,.$$

In particular, we proved

$$||f \cdot \psi|| \leq ||f||_\infty ||\psi||$$

if $f \in M_b(\mathsf{X})$ and $\psi \in L^2(\mathsf{X}, \mu)$. Consequently:
if $\{f_n\}_{n\in\mathbb{N}} \subset M_b(\mathsf{X})$ and $f_n \to f \in M_b(\mathsf{X})$ in norm $||\ ||_\infty$, as $n \to +\infty$, then also $f_n \cdot \psi \to f \cdot \psi$ in $L^2(\mathsf{X}, \mu)$.

Moreover, if $s \in S(\mathsf{X})$, the operator $\int_{\mathsf{X}} s(x)\,dP(x)$ can be made explicit using (8.39) and (8.35): for any $\psi \in L^2(\mathsf{X})$, in fact,

$$\left(\int_{\mathsf{X}} s(y)\,dP(y)\psi\right)(x) = s(x)\psi(x)\,.$$

Hence if $\{s_n\} \subset S(\mathsf{X})$ converges uniformly to $f \in M_b(\mathsf{X})$ (by Proposition 7.49(b) such a sequence exists for any $f \in M_b(\mathsf{X})$), we have

$$s_n \cdot \psi = \int_{\mathsf{X}} s_n(x)\,dP(x)\psi \to \int_{\mathsf{X}} f(x)\,dP(x)\psi$$

as $n \to +\infty$, by the definition of integral via the continuous prolongation $\widehat{\mathfrak{I}}$ of $\mathfrak{I}$. On the other hand we saw at the beginning that under our assumptions (with $f_n := s_n$) we have $s_n \cdot \psi \to f \cdot \psi$ in $L^2(\mathsf{X})$, as $n \to +\infty$, so

$$\left(\int_{\mathsf{X}} f(y)\,dP(y)\psi\right)(x) = f(x)\psi(x) \quad \text{for almost every } x \in \mathsf{X}, \tag{8.40}$$

for any $f \in M_b(\mathsf{X})$, $\psi \in L^2(\mathsf{X}, \mu)$. Equation (8.40) gives the explicit form of the integral operator of f with respect to the PVM of (8.39).

(2) For the second example consider a basis N of a separable Hilbert space H. Endow N with the discrete topology of the power set of N, for which singlets are open and the associated Borel σ-algebra is the topology itself, and hence is the power set. Note N is second countable as a topological space. If $E \subset N$ is a Borel subset, consider the closed subspace $\mathsf{H}_E := \overline{< \{z\}_{z\in E} >}$. The orthogonal projector onto such subspace is (cf. Proposition 3.58(d))

$$P(E) := \text{s-}\sum_{z\in E} (z|\ )z\,,$$

E being a basis of H_E. It is easy to check $P : \mathscr{B}(N) \ni E \mapsto P(E)$ is indeed a projector-valued measure. One can also prove, for any $f : N \to \mathbb{C}$ bounded,

$$\int_N f(z)\,dP(z) = \text{s-}\sum_{z\in N} f(z)\,(z|\ )\,z\,. \tag{8.41}$$

The proof can be obtained using example (1), because (Theorem 3.28) H and $L^2(N,\mu)$ are isomorphic Hilbert spaces under the surjective isometry $U : \mathsf{H} \to L^2(N,\mu)$ sending $x \in \mathsf{H}$ to the map $z \mapsto \psi_x(z) := (z|x)$, where μ is N's counting measure. $Q(E) := UP(E)U^{-1}$ is indeed the operator in $L^2(N,\mu)$ that multiplies by the characteristic function of E: we obtain thus a spectral measure $Q : \mathscr{B}(N) \ni E \mapsto Q(E)$ of the kind of example (1). Using the integral of a map $f \in M_b(\mathsf{X})$ defined by simple integrals, for which

$$\int_N s(z)\,dQ(z) = \sum_i c_i Q(E_i) = U\sum_i c_i P(E_i)U^{-1} = U\int_N s(z)\,dP(z)U^{-1}\,,$$

we obtain

$$\int_N f(z)\,dQ(z) = U\int_N f(z)\,dP(z)U^{-1}\,, \tag{8.42}$$

by continuity of the composite in $\mathfrak{B}(\mathsf{H})$. Equation (8.40) implies

$$\int_N f(z)\,dQ(z)\psi = f\cdot\psi\,. \tag{8.43}$$

From (8.42) and (8.43), then

$$\int_N f(z)\,dP(z)\phi = U^{-1}f\cdot U\phi = \sum_{z\in N} f(z)\,(z|\phi)\,z\,,$$

where we used the definition of U (cf. Theorem 3.28):

$$U : \mathsf{H} \ni \phi \mapsto \{(z|\phi)\}_{z\in N} \in L^2(N,\mu)$$

and the inverse:

$$U^{-1} : L^2(N,\mu) \ni \{\alpha_z\}_{z\in N} \mapsto \sum_{z\in N} \alpha_z z \in \mathsf{H}\,.$$

Altogether, we proved that

$$\int_N f(z)\,dP(z) = \text{s-}\sum_{z\in N} f(z)\,(z|\ )\,z$$

as required.

(3) The third example generalises the previous one. Consider a set X equipped with a second-countable topology in which every singlet $\{x\}$, $x \in \mathsf{X}$, belongs to the associated Borel σ-algebra $\mathscr{B}(\mathsf{X})$. For instance, take X countable with the discrete topology, or, more trivially, $\mathsf{X} = \mathbb{R}$ with standard topology, or $\mathsf{X} := \{0\} \cup \{\pm 1/n \mid n = 1, 2, \ldots\} \subset \mathbb{R}$ with topology induced by $\mathbb{R}$. Let us define a family of orthogonal projectors $P_\lambda : \mathsf{H} \to \mathsf{H}$ on the Hilbert space H, for any $\lambda \in \mathsf{X}$. In order to have a PVM we impose three conditions:

(a) $P_\lambda P_\mu = 0$, for $\lambda, \mu \in \mathsf{X}$, $\lambda \neq \mu$;
(b) $\sum_{\lambda \in \mathsf{X}} ||P_\lambda \psi||^2 < +\infty$, for any $\psi \in \mathsf{H}$;
(c) $\sum_{\lambda \in \mathsf{X}} P_\lambda \psi = \psi$, for any $\psi \in \mathsf{H}$.

Condition (b) implies that only countably many (at most, see Proposition 3.21) elements $P_\lambda \psi$ are non-zero, even if X is not countable; by (a), the vectors $P_\lambda \psi$ and $P_\mu \psi$ are orthogonal if $\lambda \neq \mu$, so Lemma 3.25 guarantees that the sum of (c) is well defined and may be rearranged at will.

That (a), (b), (c) hold is proved by exhibiting a family that satisfies them. The simplest case is given by the projectors $P(\{z\})$, $z \in N$, of example (2) when X is a basis. An instance where X is not a basis will be described in Example 8.58(1). Take a self-adjoint compact operator T, set $\mathsf{X} = \sigma_p(T)$ and define P_λ, $\lambda \in \sigma_p(T)$, to be the orthogonal projector onto the λ-eigenspace. X has the topology induced by $\mathbb{R}$. By Theorems 4.17 and 4.18 conditions (a), (b) and (c) follow.
We shall rephrase the latter two combined as:

$$\text{s-}\sum_{\lambda \in \mathsf{X}} P_\lambda = I\,. \tag{8.44}$$

With these assumptions, $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$, defined so that

$$P(E) = \text{s-}\sum_{\lambda \in E} P_\lambda\,, \tag{8.45}$$

for any $E \subset \mathscr{B}(\mathsf{X})$, is a projector-valued measure on H. The sum $\sum_{\lambda \in E} P_\lambda \psi$ always exists in H, for any $\psi \in \mathsf{H}$, and does not depend on the ordering: this fact is a consequence of condition (b), because of Lemma 3.25. Now we wish to prove

$$\int_{\mathsf{X}} f(x)\, dP(x) = \text{s-}\sum_{x \in \mathsf{X}} f(x) P_x \tag{8.46}$$

for any $f \in M_b(\mathsf{X})$. The right-hand side is well defined and can be re-ordered by Lemma 3.25, because for any $\psi \in \mathsf{H}$:

$$\sum_{x \in \mathsf{X}} ||f(x)P_x\psi||^2 \leq ||f||_\infty^2 \sum_{x \in \mathsf{X}} ||P_x\psi||^2 = ||f||_\infty^2 \sum_{x \in \mathsf{X}} (P_x\psi|P_x\psi) = ||f||_\infty^2 \sum_{x \in \mathsf{X}} (\psi|P_x^2\psi)$$
$$= ||f||_\infty^2 \sum_{x \in \mathsf{X}} (\psi|P_x\psi) = ||f||_\infty^2 \left(\psi \middle| \sum_{x \in \mathsf{X}} P_x\psi\right) = ||f||_\infty^2 (\psi|\psi) = ||f||_\infty^2 ||\psi||^2\,,$$

the last equality coming from (8.44). If $s \in S(\mathsf{X})$ is simple, using (8.45) and the definition of integral, we have

$$\int_{\mathsf{X}} s(x)\, dP(x)\psi = \sum_i c_i P(E_i)\psi = \sum_i \sum_{x\in E_i} s(x)P_x\psi = \sum_{x\in\mathsf{X}} s(x)P_x\psi\,, \tag{8.47}$$

for any $\psi \in \mathsf{H}$. Note that in the second equality we used that $s(x) = \sum_i c_i \chi_{E_i}$ implies $c_i = s(x)$ for all $x \in E_i$.

If $\{s_n\} \subset S(\mathsf{X})$ and $s_n \to f \in M_b(\mathsf{X})$ uniformly, then for any $\psi \in \mathsf{H}$:

$$\int_{\mathsf{X}} f(x)\, dP(x)\psi - \int_{\mathsf{X}} s_n(x)\, dP(x)\psi \to 0\,, \tag{8.48}$$

as $n \to +\infty$, by definition of integral of bounded measurable maps. At the same time, (8.47) and condition (a) give

$$\left|\left|\sum_{x\in\mathsf{X}} f(x)\, P_x\psi - \int_{\mathsf{X}} s_n(x)\, dP(x)\psi\right|\right|^2 = \sum_{x\in\mathsf{X}} |f(x)-s_n(x)|^2||P_x\psi||^2 \le ||f-s_n||_\infty^2||\psi||^2\,.$$

The last term goes to zero as $n \to +\infty$. By (8.48) and uniqueness of limits in H,

$$\sum_{x\in\mathsf{X}} f(x)\, P_x\psi = \int_{\mathsf{X}} f(x)\, dP(x)\psi\,,$$

for any $\psi \in \mathsf{H}$, so (8.46) holds. ∎

8.3.3 Properties of operators obtained integrating bounded maps with respect to PVMs

In this section we examine the properties of the integral operator, separating them in two groups.

The first theorem establishes basic features of the integral operator. Concerning item (v) in (c), we remind that if $\mathscr{B}(\mathsf{X})$ denotes the Borel σ-algebra of a topological space X, the **support** of a measure μ on $\mathscr{B}(\mathsf{X})$, either positive, σ-additive or complex (Definitions 1.44 and 1.83), is the closed set $supp(\mu) \subset \mathsf{X}$ given by the complement of the union of all open sets $A \subset \mathsf{X}$ with $\mu(A) = 0$.

Theorem 8.50. *Let X be a second-countable space, $(\mathsf{H}, (\ |\ ))$ a Hilbert space and $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ a projector-valued measure.*

(a) *For any $f \in M_b(\mathsf{X})$,*

$$\left|\left|\int_{\mathsf{X}} f(x)\, dP(x)\right|\right| = ||f||_\infty^{(P)} \le ||f\restriction_{supp(P)}||_\infty\,. \tag{8.49}$$

(b) *The integral operator with respect to P is* **positive***:*

$$\int_{\mathsf{X}} f(x)\, dP(x) \ge 0 \ \text{ if } 0 \le f \in M_b(\mathsf{X})\,.$$

(c) *For any* $\psi, \phi \in \mathsf{H}$*, the map*

$$\mu_{\psi,\phi} : \mathscr{B}(\mathsf{X}) \ni E \mapsto \left(\psi \middle| \int_{\mathsf{X}} \chi_E \, dP(x)\phi\right)$$

satisfies the following properties:

(i) $\mu_{\psi,\phi}$ *is a complex measure on* X*, called* **complex spectral measure associated to** ψ *and* ϕ*;*
(ii) *if* $\psi = \phi$*, then* $\mu_\psi := \mu_{\psi,\psi}$ *is a finite positive measure on* X*, called* **(positive) spectral measure associated to** ψ*;*
(iii) $\mu_{\psi,\phi}(\mathsf{X}) = (\psi|\phi)$*, and in particular* $\mu_\psi(\mathsf{X}) = ||\psi||^2$*;*
(iv) *for any* $f \in M_b(\mathsf{X})$*:*

$$\left(\psi \middle| \int_{\mathsf{X}} f(x)\, dP(x)\phi\right) = \int_{\mathsf{X}} f(x)\, d\mu_{\psi,\phi}(x)\,; \tag{8.50}$$

(v) $supp(\mu_{\psi,\phi}) \subset supp(|\mu_{\psi,\phi}|) \subset supp(P)$ *and* $supp(\mu_\psi) \subset supp(P)$*.*

(d) *If* $f \in M_b(\mathsf{X})$*,* $\int_{\mathsf{X}} f(x)dP(x)$ *commutes with every operator* $B \in \mathfrak{B}(\mathsf{H})$ *such that* $P(E)B = BP(E)$ *for any* $E \in \mathscr{B}(\mathsf{X})$*.*

Proof. (a) Consider a sequence of simple functions $s_n \to f$ in norm $||\ ||_\infty$. Then $||s_n - f||_\infty^{(P)} \le ||s_n - f||_\infty \to 0$, so $|\,||s_n||_\infty^{(P)} - ||f||_\infty^{(P)}\,| \le ||s_n - f||_\infty^{(P)}$ implies $||s_n||_\infty^{(P)} \to ||f||_\infty^{(P)}$. We also know $||\int_{\mathsf{X}} s_n dP|| = ||s_n||_\infty^{(P)}$ by (8.37). From the definition of integral of bounded maps $||\int_{\mathsf{X}} s_n dP|| \to ||\int_{\mathsf{X}} f dP||$, hence $||s_n||_\infty^{(P)} \to ||f||_\infty^{(P)} = ||\int_{\mathsf{X}} f dP||$, proving the first equality in (8.49). The inequality follows from the property of $supp(P)$ discussed in Remark 8.44.
(b) Using Proposition 7.49(c) if $0 \le f \in M_b(\mathsf{X})$ there is a sequence of simple functions $\{s_n\}_{n\in\mathbb{N}}$, $0 \le s_n \le s_{n+1} \le f$ for any n, that converges uniformly to f. Keeping in mind the definition of integral with respect to P, and that uniform convergence implies weak convergence, we have $(\psi|\int_{\mathsf{X}} s_n dP\psi) \to (\psi|\int_{\mathsf{X}} f dP\psi)$, as $n \to +\infty$, for any $\psi \in \mathsf{H}$. For the positivity of $\int_{\mathsf{X}} f dP$ it suffices to show $(\psi|\int_{\mathsf{X}} s_n dP\psi) \ge 0$ for any n. Directly from (8.35) we find

$$\left(\psi \middle| \int_{\mathsf{X}} s_n dP\ \psi\right) = \sum_{i\in I_n} c_i^{(n)} \left(\psi \middle| P(E_i^{(n)})\ \psi\right) \ge 0\,,$$

because every orthogonal projector is positive and the numbers $c_i^{(n)}$ are non-negative for $s_n \ge 0$.
(c) By (8.35),

$$\mu_{\psi,\phi}(E) = \left(\psi \middle| \int_{\mathsf{X}} \chi_E(x)\, dP(x)\phi\right) = (\psi|1 \cdot P(E)\phi) = (\psi|P(E)\phi)\,, \tag{8.51}$$

and $(\psi|P(E)\psi) \ge 0$. Then Definition 8.41(d) and the inner product's continuity imply $\mu_{\psi,\phi}$ is a complex measure on the Borel σ-algebra $\mathscr{B}(\mathsf{X})$; moreover, parts (d) and

(a) in Definition 8.41 say that if $\psi = \phi$, μ_ψ is a positive, σ-additive, finite measure on the Borel σ-algebra $\mathscr{B}(\mathsf{X})$. At last Definition 8.41(c) forces $\mu_{\psi,\phi}(\mathsf{X}) = (\psi|\phi)$, in particular $\mu_\psi(\mathsf{X}) = (\psi|\psi) = ||\psi||^2$. As μ_ψ and $|\mu_{\psi,\phi}|$ are finite measures, their integral is continuous in norm $|| \, ||_\infty$ on $M_b(\mathsf{X})$. (In fact, for any $f \in M_b(\mathsf{X})$,

$$\left| \int_{\mathsf{X}} f(x)\, d\mu_{\psi,\phi}(x) \right| \leq \int_{\mathsf{X}} |f(x)|\, d|\mu_{\psi,\phi}(x)| \leq ||f||_\infty |\mu_{\psi,\phi}|(\mathsf{X}),$$

whence the integral's continuity in sup norm.)

If $s_n \in S(\mathsf{X})$, using (8.51) and (8.35) we immediately see

$$\left(\psi \middle| \int_{\mathsf{X}} s_n(x)\, dP(x)\phi \right) = \int_{\mathsf{X}} s_n(x)\, d\mu_{\psi,\phi}(x).$$

If now $f \in M_b(\mathsf{X})$ and $\{s_n\}_{n\in\mathbb{N}} \subset S(\mathsf{X})$ converges to f, as $n \to +\infty$, in uniform sense (cf. Proposition 7.49(b)), we can use the continuity of the inner product and of the integral associated to $\mu_{\psi,\phi}$ with respect to the uniform convergence to obtain

$$\left(\psi \middle| \int_{\mathsf{X}} f(x)\, dP(x)\phi \right) = \left(\psi \middle| \lim_{n\to+\infty} \int_{\mathsf{X}} s_n(x)\, dP(x)\phi \right)$$

$$= \lim_{n\to+\infty} \left(\psi \middle| \int_{\mathsf{X}} s_n(x)\, dP(x)\phi \right) = \lim_{n\to+\infty} \int_{\mathsf{X}} s_n(x)\, d\mu_{\psi,\phi}(x) = \int_{\mathsf{X}} f(x)\, d\mu_{\psi,\phi}(x).$$

Let us prove (v), or equivalently, $\mathsf{X} \setminus supp(|\mu_{\psi,\phi}|) \supset \mathsf{X} \setminus supp(P)$. Take $x \in \mathsf{X} \setminus supp(P)$: then there is an open set $A \subset \mathsf{X}$ with $x \in A$ and $P(A) = 0$. By monotonicity $P(B) = 0$ if $\mathscr{B}(\mathsf{X}) \ni B \subset A$, and therefore

$$\mu_{\psi,\phi}(B) = \int_{\mathsf{X}} \chi_B(x)\, d\mu_{\psi,\phi}(x) = \left(\psi \middle| \int_{\mathsf{X}} \chi_B(x)\, dP(x)\phi \right) = (\psi|P(B)\phi) = 0.$$

By the definition of total variation (Remark 1.81(2)) $|\mu_{\psi,\phi}|(A) = 0$, so $x \in \mathsf{X} \setminus supp(|\mu_{\psi,\phi}|)$. But $|\mu_{\psi,\phi}|(A) \geq |\mu_{\psi,\phi}(A)|$ for any $A \in \mathscr{B}(\mathsf{X})$, in particular with A open, hence we have: $supp(\mu_{\psi,\phi}) \subset supp(|\mu_{\psi,\phi}|)$. The case μ_ψ is analogous.
(d) The claim is obvious when f is simple, and extends by continuity to any f. □

Remark 8.51. (1) It must be said that if we want the positive measures μ_ψ, defined on the Borel σ-algebra of X, to be proper *Borel measures*, then we should also demand X be Hausdorff and locally compact. In concrete situations, like when using PVMs that define the spectral expansion of an operator, X is always (a subset of) $\mathbb{R}$ or $\mathbb{R}^2$, so the extra assumptions hold. In such case the measures are also regular, see the remark preceding Theorem 2.48.
(2) A useful remark is that the complex measure $\mu_{\psi,\phi}$ decomposes as complex linear combination of 4 positive finite measures μ_χ. Since $\mu_{\psi,\phi}(E) = (\psi|P(E)\phi) = (P(E)\psi|P(E)\phi)$, by the polarisation formula (3.4) we obtain:

$$\mu_{\psi,\phi}(E) = \mu_{\psi+\phi}(E) - \mu_{\psi-\phi}(E) - i\mu_{\psi+i\phi}(E) + i\mu_{\psi-i\phi}(E) \quad \text{for any } E \in \mathscr{B}(\mathsf{X}). \ ■$$

The next theorem establishes the primary feature of projector-valued measures: they give rise to *-homomorphisms of C^*-algebras on $\mathfrak{B}(\mathsf{H})$, i.e. representations of C^*-algebras on H.

This will be a crucial point in proving the spectral theorem, proven immediately after.

Theorem 8.52. *Let H be a Hilbert space, X a second-countable space and $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ a projector-valued measure.*

(a) *The integral operator:*

$$\widehat{\mathfrak{I}} : M_b(\mathsf{X}) \ni f \mapsto \int_{\mathsf{X}} f(x) dP(x) \in \mathfrak{B}(\mathsf{H})$$

is a (continuous) representation on H of the C^*-algebra with unit $M_b(\mathsf{X})$. Equivalently: beside (8.49) the following hold:

(i) *if 1 is the constant map on X,*

$$\int_{\mathsf{X}} 1\, dP(x) = I\,;$$

(ii) *for any $f, g \in M_b(\mathsf{X})$, $\alpha, \beta \in \mathbb{C}$,*

$$\int_{\mathsf{X}} (\alpha f(x) + \beta g(x))\, dP(x) = \alpha \int_{\mathsf{X}} f(x) dP(x) + \beta \int_{\mathsf{X}} g(x) dP(x)\,;$$

(iii) *for any $f, g \in M_b(\mathsf{X})$,*

$$\int_{\mathsf{X}} f(x)\, dP(x) \int_{\mathsf{X}} g(x)\, dP(x) = \int_{\mathsf{X}} f(x) g(x)\, dP(x)\,;$$

(iv) *for any $f \in M_b(\mathsf{X})$,*

$$\int_{\mathsf{X}} \overline{f(x)}\, dP(x) = \left(\int_{\mathsf{X}} f(x) dP(x) \right)^* .$$

(b) *If $\psi \in \mathsf{H}$ and $f \in M_b(\mathsf{X})$ then*

$$\left|\left| \int_{\mathsf{X}} f(x)\, dP(x) \psi \right|\right|^2 = \int_{\mathsf{X}} |f(x)|^2\, d\mu_\psi(x)\,.$$

(c) *If $\{f_n\}_{n \in \mathbb{N}} \subset M_b(\mathsf{X})$ is bounded and converges to $f : \mathsf{X} \to \mathbb{C}$ pointwise, the integral of f with respect to the spectral measure P exists and equals:*

$$\int_{\mathsf{X}} f(x)\, dP(x) = s\text{-}\lim_{n \to +\infty} \int_{\mathsf{X}} f_n(x)\, dP(x)\,.$$

(d) *If $\mathsf{X} = \mathbb{R}^2$ with Euclidean topology, and $supp(P)$ is bounded, then*

$$supp(P) = \sigma(T)\,,$$

where $\sigma(T)$ is viewed as subset in $\mathbb{R}^2$, and we defined the normal operator:

$$T := \int_{supp(P)} z\, dP(x,y)\,,$$

with z denoting the map $\mathbb{R}^2 \ni (x,y) \mapsto z := x+iy$.
In this case let us identify $M_b(\sigma(T))$ with the closed subspace of $M_b(\mathbb{R}^2)$ of maps vanishing outside the compact set $\sigma(T)$. Then the restriction

$$\widehat{\mathfrak{I}}\restriction_{M_b(\sigma(T))} : M_b(\sigma(T)) \to \mathfrak{B}(\mathsf{H})$$

coincides with representation $\widehat{\Phi}_T$ of Theorem 8.39 for the normal operator T, and we can write

$$f(T,T^*) = \int_{\sigma(T)} f(x,y) dP(x,y)\,, \quad f \in M_b(\sigma(T)),$$

where $f(T,T^*) := \widehat{\Phi}_T(f)$.

Remark 8.53. (1) Part (iii) of (a) implies, in particular, the following commutation relation:

$$\int_{\mathsf{X}} f(x)\, dP(x) \int_{\mathsf{X}} g(x)\, dP(x) = \int_{\mathsf{X}} g(x)\, dP(x) \int_{\mathsf{X}} f(x)\, dP(x)\,,$$

for any $f, g \in M_b(\mathsf{X})$.
(2) From (iv) and (iii) of (a) the operator $\int_{\mathsf{X}} f(x)\, dP(x)$ is *normal*, for any $f \in M_b(\mathsf{X})$. ∎

Proof of Theorem 8.52. (a) The only facts not entirely trivial are (iii) and (iv), so let us begin with the former. Choose two sequences of simple functions $\{s_n\}$ and $\{t_m\}$ that converge uniformly to f and g respectively. A direct computations shows

$$\int_{\mathsf{X}} s_n(x)\, dP(x) \int_{\mathsf{X}} t_m(x)\, dP(x) = \int_{\mathsf{X}} s_n(x) t_m(x)\, dP(x)\,.$$

Given m, $s_n \cdot t_m$ tends to $f \cdot t_m$ uniformly, as $n \to +\infty$, because t_m is bounded. By continuity (in the sense of Theorem 8.50(a)) and linearity of the integral and taking the limit $n \to +\infty$ above, we obtain

$$\int_{\mathsf{X}} f(x)\, dP(x) \int_{\mathsf{X}} t_m(x)\, dP(x) = \int_{\mathsf{X}} f(x) t_m(x)\, dP(x)\,,$$

where we used the fact that the composite of bounded operators is continuous in its arguments. Similarly, since $f \cdot t_m$ tends to $f \cdot g$ uniformly as $m \to +\infty$, we obtain (iii). Property (iv) is proven by choosing a sequence of simple functions $\{s_n\}$ uniformly converging to f. Take $\psi, \phi \in \mathsf{H}$. Directly by definition of integral of a simple function (NB: orthogonal projectors are self-adjoint), we have

$$\left(\int_{\mathsf{X}} \overline{s_n(x)}\, dP(x) \psi \middle| \phi \right) = \left(\psi \middle| \int_{\mathsf{X}} s_n(x)\, dP(x) \phi \right)\,.$$

Notice $\overline{s_n} \to \overline{f}$ uniformly, as $n \to +\infty$. Hence by continuity and linearity of the integral (in the sense of Theorem 8.50(a)), plus the continuity of the inner product, when we take the limit as $n \to +\infty$, the above identity gives

$$\left(\int_{\mathsf{X}} \overline{f(x)}\, dP(x)\psi \middle| \phi \right) = \left(\psi \middle| \int_{\mathsf{X}} f(x)\, dP(x)\phi \right),$$

so:

$$\left(\left[\int_{\mathsf{X}} \overline{f(x)}\, dP(x) - \left(\int_{\mathsf{X}} f(x)\, dP(x) \right)^* \right] \psi \middle| \phi \right) = 0\,.$$

As $\psi, \phi \in \mathsf{H}$ are arbitrary, (iv) holds.
(b) If $\psi \in \mathsf{H}$, using (iii) and (iv) of (a), we find

$$\left|\left| \int_{\mathsf{X}} f(x)\, dP(x)\psi \right|\right|^2 = \left(\psi \middle| \int_{\mathsf{X}} |f(x)|^2\, dP(x)\psi \right) = \int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x)\,,$$

where the last equality used Theorem 8.50(c).
(c) As first thing let us observe $f \in M_b(\mathsf{X})$, because f is measurable, as limit of measurable functions, and bounded by the constant that bounds the sequence f_n. If $\psi \in \mathsf{H}$ the integral's linearity and (b) imply

$$\left|\left| \left(\int_{\mathsf{X}} f(x)\, dP(x) - \int_{\mathsf{X}} f_n(x)\, dP(x) \right) \psi \right|\right|^2 = \int_{\mathsf{X}} |f(x) - f_n(x)|^2\, d\mu_\psi(x)\,.$$

The measure μ_ψ is finite, so constant maps are integrable. By assumption $|f_n| < K < +\infty$ for any $n \in \mathbb{N}$, so $|f| \leq K$ and then $|f_n - f|^2 \leq (|f_n| + |f|)^2 < 4K^2$. Since $|f_n - f|^2 \to 0$ pointwise, we can invoke the dominated convergence theorem to obtain, as $n \to +\infty$,

$$\left|\left| \int_{\mathsf{X}} f(x)\, dP(x)\psi - \int_{\mathsf{X}} f_n(x)\, dP(x)\psi \right|\right| = \sqrt{\int_{\mathsf{X}} |f(x) - f_n(x)|^2\, d\mu_\psi(x)} \to 0\,.$$

Given the freedom in $\psi \in \mathsf{H}$, (c) is proved.
(d) If $supp(P)$ is bounded, it is compact (as closed by definition) and every continuous map on it is bounded. The mapping $\mathbb{R}^2 \ni (x,y) \mapsto z\chi_{supp(P)}(x,y) \in \mathbb{C}$ is thus bounded, so $T := \int_{supp(P)} z\, dP(x,y) := \int_{\mathbb{R}^2} z\chi_{supp(P)}(x,y)\, dP(x,y)$ is well defined and a normal operator ((ii) and (iv) in (a)) in $\mathfrak{B}(\mathsf{H})$. Its residual spectrum is in particular empty, by Proposition 8.7(c).

By definition of resolvent set, the claim is the same as asking:

$$\mathbb{C} \ni \lambda \notin \; supp(P) \; \textit{iff}\; \lambda \in \rho(T)\,.$$

Let us prove $\lambda \notin supp(P)$ implies $\lambda \in \rho(T)$. Since $\mathbb{R}^2 \ni (x,y) \mapsto z = x+iy$ is bounded on $supp(P)$, suppose $\lambda \notin supp(P)$. If so, there is an open subset in $\mathbb{R}^2$, $A \ni (x_0, y_0)$ with $x_0 + iy_0 = \lambda$, such that $P(A) = 0$. It follows that $(x,y) \mapsto (z-\lambda)^{-1}$ is bounded on the closed set $supp(P)$. Then we have the operator

$$\int_{supp(P)} \frac{1}{z-\lambda}\, dP(x,y)$$

of $\mathfrak{B}(\mathsf{H})$. By virtue of (iii) and (i) in (a),

$$\int_{supp(P)} \frac{1}{z-\lambda}\, dP(x,y) \int_{supp(P)} (z-\lambda)\, dP(x,y)$$

$$= \int_{supp(P)} (z-\lambda)\, dP(x,y) \int_{supp(P)} \frac{1}{z-\lambda}\, dP(x,y)$$

$$= \int_{supp(P)} 1\, dP(x,y) = \int_{\mathbb{R}^2} 1\, dP(x,y) = I\,,$$

which we may write, by (i) and (ii) of (a),

$$\int_{supp(P)} \frac{1}{z-\lambda}\, dP(x,y)(T-\lambda I) = (T-\lambda I) \int_{supp(P)} \frac{1}{z-\lambda}\, dP(x,y) = I\,.$$

Put differently, $T-\lambda I$ is a bijection of H so that $\lambda \in \rho(T)$. By Theorem 8.4(a) $\lambda \in \rho(T)$.

Now we show $\lambda \in \rho(T)$ implies $\lambda \notin supp(P)$, and actually we prove the equivalent statement: $\lambda \in supp(P)$ *implies* $\lambda \in \sigma(T) = \sigma_p(T) \cup \sigma_c(T)$.

If $\lambda \in supp(P)$, is may happen that $T-\lambda I : \mathsf{H} \to \mathsf{H}$ is not one-to-one, in which case $\lambda \in \sigma_p(T)$ and the proof ends. If, on the contrary, $T-\lambda I : \mathsf{H} \to \mathsf{H}$ is injective we can prove the inverse $(T-\lambda I)^{-1} : Ran(T-\lambda I) \to \mathsf{H}$ cannot be bounded, so $\lambda \in \sigma_c(T)$. For that it is enough to show, for $\lambda \in supp(P)$, $n = 1,2,\ldots$, that there exists $\psi_n \in \mathsf{H}$, $\psi_n \neq 0$, such that

$$||(T-\lambda I)\psi_n||/||\psi_n|| \leq 1/n\,.$$

(Under our assumptions $\psi_n = (T-\lambda I)^{-1}\phi_n$, for any $n = 1,2,\ldots$, with $\phi_n \neq 0$ so that $\psi_n \neq 0$. Then

$$1/n \geq ||(T-\lambda I)\psi_n||/||\psi_n|| = ||(T-\lambda I)(T-\lambda I)^{-1}\phi_n||/||(T-\lambda I)^{-1}\phi_n||\,.$$

In other terms, for $n = 1,2,\ldots$, there is $\phi_n \in \mathsf{H}$, $\phi_n \neq 0$, such that

$$\frac{||(T-\lambda I)^{-1}\phi_n||}{||\phi_n||} \geq n\,.$$

Then $(T-\lambda I)^{-1}$ cannot be bounded, and hence $\lambda \in \sigma_c(T)$.)

If $\lambda \in supp(P)$, any open set $A \ni \lambda$ must satisfy $P(A) \neq 0$. Set $x_0 + iy_0 := \lambda$ and consider the family of open discs $D_n \subset \mathbb{R}^2$, centred at (x_0, y_0) and of radii $1/n$, $n = 1,2,\ldots$. As $P(D_n) \neq 0$, there exists $\psi_n \neq 0$ with $\psi_n \in P(D_n)(\mathsf{H})$. In such a case

$$(T-\lambda I)\psi_n = \int_{supp(P)} (z-\lambda)\, dP(x,y)\psi_n$$

$$= \int_{supp(P)} (z-\lambda)\, dP(x,y) \int_{supp(P)} \chi_{D_n}(z)\, dP(x,y)\psi_n\,,$$

where we used $P(D_n) = \int_{\mathbb{R}^2} \chi_{D_n}(z)\, dP(x,y)$ and $P(D_n)\psi_n = \psi_n$. By part (iii) in (a) we find

$$(T - \lambda I)\psi_n = \int_{\mathbb{R}^2} \chi_{D_n}(z)(z-\lambda)dP(x,y)\,.$$

Hence property (b) yields

$$||(T-\lambda I)\psi_n||^2 = \int_{\mathbb{R}^2} |\chi_{D_n}(z)|^2 |z-\lambda|^2 d\mu_{\psi_n}(x,y) \leq \int_{\mathbb{R}^2} 1\cdot n^{-2}\, d\mu_{\psi_n}(x,y)$$

$$= n^{-2}\int_{\mathbb{R}^2} 1\, d\mu_{\psi_n}(x,y) = n^{-2}||\psi_n||^2\,,$$

where the last equality made use of $\mu_{\psi_n}(\mathbb{R}^2) = ||\psi_n||^2$, by (iii) in Theorem 8.50(c). Taking the square root of both sides produces

$$\frac{||(T-\lambda I)\psi_n||}{||\psi_n||} \leq 1/n\,,$$

and concludes the proof.

The final statement is an easy consequence of $\widehat{\Phi}_T$'s uniqueness, because $\widehat{\mathfrak{I}}$ restricted to $M_b(\sigma(T)) = M_b(supp(P))$ trivially satisfies all the conditions (see Theorem 8.39). □

8.4 Spectral theorem for normal operators in $\mathfrak{B}(\mathsf{H})$

At this juncture enough material has been gathered to allow to state the first important *spectral decomposition theorem* for normal operators in $\mathfrak{B}(\mathsf{H})$. Later in this section we will prove another version of the theorem that concerns a useful *spectral representation* for bounded normal operators T, in terms of multiplicative operators on certain L^2 spaces built on the spectrum of T.

8.4.1 Spectral decomposition of normal operators in $\mathfrak{B}(\mathsf{H})$

The spectral decomposition theorem establishes how any normal operator of $\mathfrak{B}(\mathsf{H})$ can be constructed integrating a certain PVM, whose support is the operator's spectrum and which is completely determined by the operator. In view of the applications it is important to point out that the theorem holds in particular for self-adjoint operators in $\mathfrak{B}(\mathsf{H})$ and unitary operators, both subcases of normal operators.

Theorem 8.54 (Spectral decomposition of normal operators in $\mathfrak{B}(\mathsf{H})$). *Let* H *be a Hilbert space and* $T \in \mathfrak{B}(\mathsf{H})$ *a normal operator.*

(a) *There exists a unique, and bounded, projector-valued measure* $P^{(T)}$ *on* $\mathbb{R}^2$ *(standard) such that:*

$$T = \int_{supp(P^{(T)})} z\, dP^{(T)}(x,y)\,, \tag{8.52}$$

where z is the map $\mathbb{R}^2 \ni (x,y) \mapsto z := x+iy \in \mathbb{C}$.
(a)' *If T is self-adjoint, or unitary, statement (a) can be refined replacing $\mathbb{R}^2$ with:*

$$\mathbb{R} \quad or \quad \mathbb{S}^1 := \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\}, respectively.$$

(b) *We have*

$$supp(P^{(T)}) = \sigma(T) .$$

In particular, for $\lambda = x+iy \in \mathbb{C}$ ($\lambda = x \in \mathbb{R}$, or $\lambda = x+iy \in \mathbb{S}^1$ respectively):

(i) $\lambda \in \sigma_p(T) \Leftrightarrow P^{(T)}(\{\lambda\}) \neq 0$;
(ii) $\lambda \in \sigma_c(T) \Leftrightarrow P^{(T)}(\{\lambda\}) = 0$, *and* $P^{(T)}(A_\lambda) \neq 0$ *for any open set* $A_\lambda \subset \mathbb{R}^2$ *($\mathbb{R}$ or $\mathbb{S}^1$ respectively), $A_\lambda \ni \lambda$;*
(iii) *if $\lambda \in \sigma(T)$ is isolated, $\lambda \in \sigma_p(T)$;*
(iv) *if $\lambda \in \sigma_c(T)$, then for any $\varepsilon > 0$ there exists $\phi_\varepsilon \in \mathsf{H}$ with $||\phi_\varepsilon|| = 1$ and*

$$0 < ||T\phi_\varepsilon - \lambda\phi_\varepsilon|| \leq \varepsilon .$$

(c) *If $f \in M_b(\sigma(T))$, the operator $\int_{\sigma(T)} f(x,y)\, dP^{(T)}(x,y)$ commutes with every operator in $\mathfrak{B}(\mathsf{H})$ that commutes with T and T^*.*

Remark 8.55. (1) In practice, property (iv) of (b) says that when $\lambda \in \sigma_c(T)$, despite there are no eigenvectors of T with eigenvalue λ (the continuous and discrete spectra are disjoint), we can still construct vectors that solve the characteristic equation with arbitrary approximation.
(2) Bearing in mind Theorem 8.52(d) we may rephrase part (c), that restates (iii) of Theorem 8.39(b), as follows. By definition of *von Neumann algebra* (see Definition 3.47 and ensuing remarks): *the *-subalgebra of $\mathfrak{B}(\mathsf{H})$ of operators $f(T,T^*)$, for $f \in M_b(\sigma(T))$, is contained in the von Neumann algebra generated by T, T^* in $\mathfrak{B}(\mathsf{H})$.*

In Theorem 9.9 we will prove this inclusion is actually an equality, provided H is separable. ■

Proof of Theorem 8.54. (a), (a)' and (c) *Uniqueness*. We begin with the spectral measure's uniqueness. First, note that if a spectral measure P satisfies (8.52) it must have bounded support, since the map z is not bounded on unbounded sets and the right-hand side in (8.52) is understood as in Definition 8.47(c). So let P, P' be projector-valued measures with bounded support (so compact, for $supp(P)$ is closed in $\mathbb{R}^2$ by definition) and such that:

$$T = \int_{supp(P)} z\, dP(x,y) = \int_{supp(P')} z\, dP'(x,y) . \tag{8.53}$$

Using (i)–(iv) in Theorem 8.52(a), this gives, for any polynomial $p = p(z,\bar{z})$,

$$p(T,T^*) = \int_{supp(P)} p(x+iy, x-iy)\, dP(x,y) = \int_{supp(P')} p(x+iy, x-iy)\, dP'(x,y) ,$$

where the polynomial $p(T,T^*)$ is defined in the most obvious manner, i.e. reading multiplication as composition of operators and setting $T^0 := I$, $(T^*)^0 := I$. If $u,v \in \mathsf{H}$ are arbitrary, for any complex polynomial $p = p(z,\bar{z})$ on $\mathbb{R}^2$,

$$\int_{supp(\mu_{u,v})} p(z,\bar{z})\, d\mu_{u,v}(x,y) = \left(u \middle| \int_{supp(P)} p(z,\bar{z})\, dP(x,y) v \right)$$

$$= \left(u \middle| \int_{supp(P')} p(z,\bar{z})\, dP'(x,y) v \right) = \int_{supp(\mu'_{u,v})} p(z,\bar{z})\, d\mu'_{u,v}(x,y)\,.$$

The two complex measures $\mu_{u,v}$ and $\mu'_{u,v}$ are those of Theorem 8.50(c) (where u, v were ψ, ϕ). Moreover $supp(\mu_{u,v})$, $supp(\mu'_{u,v})$ are compact subsets of $\mathbb{R}^2$ (by (v) Theorem 8.50(c)), so there exists a compact set $K \subset \mathbb{R}^2$ containing both. Let us extend in the obvious way the measures to K, maintaing the supports intact, by defining the measure of a Borel set E in K by $\mu_{u,v}(E \cap supp(\mu_{u,v}))$, and similarly for $\mu'_{u,v}$.

Since polynomials in $z, \bar{z}$ with complex coefficients are bijectively identified with complex polynomials $q(x,y)$ in the real x,y (under the usual identification $z := x+iy$ and $\bar{z} := x-iy$, so $p(x+iy, x-iy) = q(x,y)$), we can rewrite the above identities in terms of polynomials with complex coefficients in $(x,y) \in K$:

$$\int_K p(x+iy, x-iy)\, d\mu_{u,v}(x,y) = \int_K p(x+iy, x-iy)\, d\mu'_{u,v}(x,y)\,.$$

By the Stone–Weierstrass theorem (2.27), the algebra of complex polynomials $q(x,y)$ is dense inside $C(K)$ for the sup norm. Therefore the algebra of complex polynomials $p(x+iy, x-iy)$ restricted to K is dense in $C(K)$ for the sup norm. Since integrals of complex measures are continuous functionals in sup norm,

$$\int_K f(x,y)\, d\mu_{u,v}(x,y) = \int_K f(x,y)\, d\mu'_{u,v}(x,y) \quad \text{for any } f \in C(K)\,.$$

Riesz's Theorem 2.48 for complex measures ensures the two extended measures must concide. Consequently the yet-to-be-extended measures must have the same support and coincide. Now by (iv) in Theorem 8.50(c), for any pair of vectors $u,v \in \mathsf{H}$ and any bounded measurable g on $\mathbb{R}^2$ we have

$$\left(u \middle| \int_{supp(P)} g(x,y)\, dP(x,y) v \right) = \left(u \middle| \int_{supp(P')} g(x,y)\, dP'(x,y) v \right),$$

i.e.

$$\left(u \middle| \int_{\mathbb{R}^2} g(x,y)\, dP(x,y) v \right) = \left(u \middle| \int_{\mathbb{R}^2} g(x,y)\, dP'(x,y) v \right).$$

Thus

$$\int_{\mathbb{R}^2} g(x,y)\, dP(x,y) = \int_{\mathbb{R}^2} g(x,y)\, dP'(x,y)$$

because u and v are arbitrary. If E is an arbitrary Borel set of $\mathbb{R}^2$ and we pick $g = \chi_E$, the above equation implies

$$P(E) = \int_{\mathbb{R}^2} \chi_E(x,y)\, dP(x,y) = \int_{\mathbb{R}^2} \chi_E(x,y)\, dP'(x,y) = P'(E)\,,$$

proving $P = P'$.

Observe, furthermore, that (8.53) and Theorem 8.52(d) give $supp(P^{(T)}) = \sigma(T)$.

Uniqueness for the cases of (a)' follows by what we have just proved, by $supp(P^{(T)}) = \sigma(T)$ and by (a)-(b) (i) in Proposition 8.7.

Existence. Let us see to the existence of the spectral measure $P^{(T)}$. Consider the *-homomorphism $\widehat{\Phi}_T$ associated to the normal operator $T \in \mathfrak{B}(\mathsf{H})$, as of Theorem 8.39. If E is a Borel set in $\mathbb{R}^2$, define $E' := E \cap \sigma(T)$ whence $P^{(T)}(E) := \widehat{\Phi}_T(\chi_{E'})$. The operator $P^{(T)}(E)$ is patently idempotent, for $\widehat{\Phi}_T$ is a homomorphism and $\chi_{E'} \cdot \chi_{E'} = \chi_{E'}$. By property (vi) of Theorem 8.39(b) and the positivity of characteristic functions, $P^{(T)}(E) \geq 0$, so $P^{(T)}(E)$ is self-adjoint. Therefore every $P^{(T)}(E)$ is an orthogonal projector. It is easy to check $\mathscr{B}(\mathbb{R}) \ni E \mapsto P^{(T)}(E)$ is a PVM: $P^{(T)}(E) \geq 0$ we saw above. Concerning Definition 8.41: (b) follows from $\chi_{E'} \cdot \chi_{F'} = \chi_{E' \cap F'}$ and because $\widehat{\Phi}_T$ is a homomorphism; (c) descends from $\widehat{\Phi}_T(\chi_{\sigma(T)}) = I$, which holds by definition of algebra homomorphism; eventually, (d) follows from (v) in Theorem 8.41(b), because, pointwise, $\lim_{N\to+\infty} \sum_{n=0}^{N} \chi_{E'_n} = \chi_{\cup_{n\in\mathbb{N}} E'_n}$ when the E'_n are pairwise disjoint. By construction, $supp(P^{(T)})$ is bounded because $supp(P^{(T)}) \subset \sigma(T)$, the latter being compact by Theorem 8.4(c).

To continue with the proof, let us remark the following fact. By the above argument, and because both the integral operator associated to $P^{(T)}$ and $\widehat{\Phi}_T$ are linear,

$$\widehat{\Phi}_T(s \restriction_{\sigma(T)}) = \int_{supp(P^{(T)})} s(x,y)\, dP^{(T)}(x,y)\,,$$

for any simple function $s : \mathbb{R}^2 \to \mathbb{C}$. Either functional is continuous for the sup topology ((ii) in Theorem 8.39(b) and (a)), so Proposition 7.49 gives

$$\widehat{\Phi}_T(f \restriction_{\sigma(T)}) = \int_{supp(P^{(T)})} f(x,y)\, dP^{(T)}(x,y)\,, \tag{8.54}$$

for any $f : \mathbb{R}^2 \to \mathbb{C}$ measurable and bounded. In particular, by (i) Theorem 8.39(a)

$$T = \int_{supp(P^{(T)})} z\, dP^{(T)}(x,y)\,.$$

As far as the proof of (c) is concerned, notice that (8.54) implies $A \in \mathfrak{B}(\mathsf{H})$ commutes with $\int_{supp(P^{(T)})} f(x,y)\, dP^{(T)}(x,y)$, when A commutes with T, T^*; that is because A commutes with $\widehat{\Phi}_T(f \restriction_{\sigma(T)})$ in consequence of (iii) in Theorem 8.39(b).

(b) Let us prove the claim for the general case where T is not necessarily self-adjoint nor unitary; these special cases are easily proved with this argument. The fact that $supp(P^{(T)}) = \sigma(T)$ is a straightforward consequence of Theorem 8.52(d). Let us prove (i). We shall write P instead of $P^{(T)}$ to simplify the notation. Let $\lambda := x_0 + iy_0$ be a complex number. By (iii) of Theorem 8.52(a),

$$\begin{aligned} TP(\{(x_0,y_0)\}) &= \int_{\sigma(T)} (x+iy)\chi_{\{(x_0,y_0)\}}(x,y)\, dP(x,y) \\ &= \int_{\sigma(T)} (x_0+iy_0)\chi_{\{(x_0,y_0)\}}(x,y)\, dP(x,y) = \lambda \int_{\sigma(T)} \chi_{\{(x_0,y_0)\}}(x,y)\, dP(x,y)\,. \end{aligned}$$

Hence $TP(\{(x_0,y_0)\}) = \lambda P(\{(x_0,y_0)\})$. We conclude that $P(\{(x_0,y_0)\}) \neq 0$ implies $\lambda := x_0 + iy_0$ is an eigenvalue of T, because any vector $u \neq 0$ in the target subspace of $P(\{(x_0,y_0)\})$ is an eigenvector relative to λ.

Suppose conversely $Tu = \lambda u$, $u \neq 0$ and $\lambda := x_0 + iy_0$. Then (cf. (b) of (i) in Proposition 3.54) $T^*u = \overline{\lambda} u$, $T^n (T^*)^m u = \lambda^n \overline{\lambda}^m u$, and by linearity

$$p(T,T^*)u = \int_{supp(P)} p(x+iy, x-iy)\, dP(x,y)u = p(\lambda,\overline{\lambda})u \tag{8.55}$$

for any polynomial $p = p(x+iy, x-iy)$, because the integral defines a *-homomorphism. Every polynomial $p = p(x+iy, x-iy)$ is also a complex polynomial $q = q(x,y)$ in the real variables x, y by setting $q(x,y) := p(x+iy, x-iy)$ pointwise; the correspondence is bijective. Since the $q(x,y)$ approximate continuous maps $f(x,y)$ in sup norm, the second equality of (8.55) holds when $p(x+iy, x-iy) = q(x,y)$ is replaced by the continuous $f = f(x,y)$. If $\lambda = x_0 + iy_0$, it is not hard to see $\chi_{\{(x_0,y_0)\}}$ is the pointwise limit of a bounded sequence of continuous maps f_n. Using Theorem 8.50(c) and dominated convergence (μ_u is finite), we have:

$$\begin{aligned}
(u|P_{\{(x_0,y_0)\}}u) &= \left(u \middle| \int_{supp(P)} \chi_{\{(x_0,y_0)\}}(x,y) dP(x,y)\, u \right) \\
&= \int_{supp(P)} \chi_{\{(x_0,y_0)\}}(x,y) d\mu_u(x,y) = \lim_{n\to+\infty} \int_{supp(P)} f_n(x,y) d\mu_u(x,y) \\
&= \lim_{n\to+\infty} \left(u \middle| \int_{supp(P)} f_n(x,y) dP(x,y)\, u \right) = \lim_{n\to+\infty} (u|f_n(x_0,y_0)u) \\
&= \chi_{\{(x_0,y_0)\}}(x_0,y_0)(u|u)\,.
\end{aligned}$$

Since orthogonal projectors are idempotent and self-adjoint, and since $\chi_{\{(x_0,y_0)\}}(x_0,y_0) = 1$ by definition,

$$(P_{\{(x_0,y_0)\}}u|P_{\{(x_0,y_0)\}}u) = (u|u) \neq 0\,.$$

Hence $P_{\{(x_0,y_0)\}} \neq 0$.

Let us pass to (ii). As $\sigma_c(T) \cup \sigma_p(T) = \sigma(T)$ (by (i) Proposition 8.7(c)) and $\sigma_c(T) \cap \sigma_p(T) = \varnothing$ by definition, we must have $\lambda \in \sigma_c(T)$ if and only if $\lambda \in \sigma(T)$ and $\lambda \notin \sigma_p(T)$. But $supp(P^{(T)}) = \sigma(T)$, so $\lambda \in \sigma(T)$ is the same as saying, for any open set A in $\mathbb{R}^2$ containing (x_0,y_0), $x_0 + iy_0 = \lambda$, that $P(A) \neq 0$. On the other hand, by (i), $\lambda \notin \sigma_p(T)$ means $P^{(T)}(\{(x_0,y_0)\}) = 0$.

Now (iii). If $\lambda = x_0 + iy_0 \in \mathbb{C}$ is an isolated point in $\sigma(T)$, by definition there is an open set $A \ni \{(x_0,y_0)\}$ disjoint from the remaining part of $\sigma(T)$. If $P(\{(x_0,y_0)\})$ were 0, then λ would belong to $supp(P^{(T)})$, for in that case $P(A) = 0$. Necessarily, then, $P^{(T)}(\{(x_0,y_0)\}) \neq 0$. By (i) we have $\lambda \in \sigma_p(T)$.

The proof of (iv) is contained in Theorem 8.52(d), where we proved, among other facts, that if $\lambda \in \sigma_c(T)$, for any natural number $n > 0$ there exists $\psi_n \in \mathsf{H}$ with $||\psi_n|| \neq 0$ and $0 < ||T\psi_n - \lambda\psi_n||/||\psi_n|| \leq 1/n$. To have (iv) it suffices to define $\phi_n := \psi_n/||\psi_n||$ with n such that $1 \leq \varepsilon n$ for the given ε. □

8.4.2 Spectral representation of normal operators in $\mathfrak{B}(\mathsf{H})$

The next important result provides a *spectral representation* for any normal operator $T \in \mathfrak{B}(\mathsf{H})$; it is shown that every bounded normal operator can be viewed as a multiplicative operator, on a suitable space L^2, basically built on the spectrum of T. In the sequel we shall refer to Definition 7.34 and the ensuing remark.

Theorem 8.56 (Spectral representation of normal operators in $\mathfrak{B}(\mathsf{H})$). *Let H be a Hilbert space, $T \in \mathfrak{B}(\mathsf{H})$ a normal operator, $P^{(T)}$ the spectral measure associated to T as of Theorem 8.54(a) (or (a)').*

(a) H *splits as a Hilbert sum $\mathsf{H} = \oplus_{\alpha\in A}\mathsf{H}_\alpha$ (A at most countable if H is separable), where the subspaces H_α are closed and mutually orthogonal, such that:*

(i) *for any $\alpha \in A$, $T\mathsf{H}_\alpha \subset \mathsf{H}_\alpha$ and $T^*\mathsf{H}_\alpha \subset \mathsf{H}_\alpha$;*

(ii) *for any $\alpha \in A$ there exist a positive, finite Borel measure μ_α on the spectrum $\sigma(T) \subset \mathbb{R}^2$, and a surjective isometry $U_\alpha : \mathsf{H}_\alpha \to L^2(\sigma(T), \mu_\alpha)$, such that*

$$U_\alpha \left(\int_{\sigma(T)} f(x,y)dP^{(T)}(x,y) \right) \restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = f\cdot\,,$$

for $f \in M_b(\sigma(T))$. In particular:

$$U_\alpha T \restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = (x+iy)\cdot\,, \quad U_\alpha T^* \restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = (x-iy)\cdot$$

where $f\cdot$ is the multiplication by f in $L^2(\sigma(T), \mu_\alpha)$:

$$(f\cdot g)(x,y) = f(x,y)g(x,y)$$

almost everywhere on $\sigma(T)$ if $g \in L^2(\sigma(T), \mu_\alpha)$;

(ii)' *if T is self-adjoint or unitary, there exist, for any $\alpha \in A$, a positive finite Borel measure, on Borel sets of $\sigma(T) \subset \mathbb{R}$ or $\sigma(T) \subset \mathbb{S}^1$ (respectively), and a surjective isometry $U_\alpha : \mathsf{H}_\alpha \to L^2(\sigma(T), \mu_\alpha)$, such that*

$$U_\alpha \left(\int_{\sigma(T)} f(x)dP^{(T)}(x) \right) \restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = f\cdot\,,$$

for $f \in M_b(\sigma(T))$. In particular,

$$U_\alpha T \restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = x\cdot\,,$$

where $f\cdot$ is the multiplication by f on $L^2(\sigma(T), \mu_\alpha)$:

$$(f\cdot g)(x) = f(x)g(x) \text{ almost everywhere on } \sigma(T)$$

for any $g \in L^2(\sigma(T), \mu_\alpha)$.

(b)

$$\sigma(T) = supp\{\mu_\alpha\}_{\alpha\in A}\,,$$

where $supp\{\mu_\alpha\}_{\alpha\in A}$ *is the complement to the set of* $\lambda \in \mathbb{C}$ *(respectively* $\mathbb{R}$*, or* $\mathbb{S}^1$*) for which there is an open set* $A_\lambda \subset \mathbb{C}$ *(respectively* $\mathbb{R}$*, or* $\mathbb{S}^1$*) such that* $A_\lambda \ni \lambda$ *and* $\mu_\alpha(A_\lambda) = 0$ *for any* $\alpha \in A$*.*
(c) *If* H *is separable, there exist a measure space* $(\mathsf{M}_T, \Sigma_T, \mu_T)$*, with* $\mu_T(\mathsf{M}_T) < +\infty$*, a bounded map* $F_T : M_T \to \mathbb{C}$ *(respectively* $\mathbb{R}$*, or* $\mathbb{S}^1$ *if* T *is self-adjoint, or unitary), and a unitary operator* $U_T : \mathsf{H} \to L^2(\mathsf{M}_T, \mu_T)$*, satisfying*

$$\left(U_T T U_T^{-1} f\right)(m) = F_T(m) f(m)\,,\ \left(U_T T^* U_T^{-1} f\right)(m) = \overline{F_T(m)} f(m) \quad \textit{for any } f \in \mathsf{H}. \tag{8.56}$$

Proof. (a) We prove (i), (ii) and (iii). The proof of (ii)' is similar to (ii). To begin with, assume there is a vector $\psi \in \mathsf{H}$ whose subspace H_ψ, containing vectors of type $\int_{\sigma(T)} g(x,y)\, dP^{(T)}(x,y)\psi$, $g \in M_b(\sigma(T))$, is dense in H. If μ_ψ is the spectral measure associated to ψ (finite since

$$\int_{supp(P^{(T)})} 1\, d\mu_\psi = ||\psi||^2\,)$$

we have $supp(\mu_\psi) \subset supp(P^{(T)})$ by (iv) in Theorem 8.50(c). Consider the Hilbert space $L^2(\sigma(T), \mu_\psi)$ and the linear surjective operator

$$V_\psi : M_b(\sigma(T)) \ni g \mapsto \int_{\sigma(T)} g(x,y)\, dP^{(T)}(x,y)\psi \in H_\psi\,.$$

As μ_ψ is finite, $M_b(\sigma(T)) \subset L^2(\sigma(T), \mu_\psi)$ as subspace. Hence for any $g_1, g_2 \in M_b(\sigma(T))$,

$$\int_{\sigma(T)} \overline{g_1(x,y)} g_2(x,y)\, d\mu_\psi(x,y)$$
$$= \left(\int_{\sigma(T)} g_1(x,y)\, dP^{(T)}(x,y)\psi \middle| \int_{\sigma(T)} g_2(x,y)\, dP^{(T)}(x,y)\psi \right), \tag{8.57}$$

or equivalently,

$$\int_{\sigma(T)} \overline{g_1(x,y)} g_2(x,y)\, d\mu_\psi(x,y) = \left(V_\psi g_1 | V_\psi g_2\right)\,. \tag{8.58}$$

The proof of (8.57) descends by the following observation. If $E, E' \subset \sigma(T)$ are Borel sets, using (iv) in Theorem 8.50(c), (iii) in Theorem 8.52(a) and (iv) in Theorem 8.52(a),

$$\int_{\sigma(T)} \overline{\chi_E} \chi_{E'}\, d\mu_\psi = \int_{\sigma(T)} \chi_{E\cap E'}\, d\mu_\psi = \left(\psi \middle| \int_{\sigma(T)} \chi_{E\cap E'}\, dP^{(T)}\ \psi\right) =$$
$$\left(\psi \middle| \int_{\sigma(T)} \chi_E \chi_{E'}\, dP^{(T)}\ \psi\right) = \left(\psi \middle| \int_{\sigma(T)} \chi_E\, dP^{(T)} \int_{\sigma(T)} \chi_{E'}\, dP^{(T)}\ \psi\right)$$
$$= \left(\int_{\sigma(T)} \chi_E\, dP^{(T)}\psi \middle| \int_{\sigma(T)} \chi_{E'}\, dP^{(T)}\ \psi\right);$$

by linearity of the inner product, if s and t are simple,

$$\int_{\sigma(T)} \overline{s} t \, d\mu_\psi = \left(\int_{\sigma(T)} s \, dP^{(T)} \psi \middle| \int_{\sigma(T)} t \, dP^{(T)} \psi \right) .$$

By Proposition 7.49, using the definition of integral of a measurable bounded map for a spectral measure, by dominated convergence with respect to the finite measure μ_ψ and by the inner product's continuity, the above identity implies (8.57). Thus we have proved V_ψ is a surjective isometry mapping $M_b(\sigma(T))$ to H_ψ. Notice that $M_b(\sigma(T))$ is dense in $L^2(\sigma(T), \mu_\psi)$, because if $g \in L^2(\sigma(T), \mu_\psi)$, the maps $g_n := \chi_{E_n} \cdot g$,

$$E_n := \{(x,y) \in \sigma(T) \mid |g(x,y)| < n\} ,$$

belong in $M_b(\sigma(T))$, and $g_n \to g$ in $L^2(\sigma(T), \mu_\psi)$ by dominated convergence (pointwise $|g_n - g|^2 \to 0$, as $n \to +\infty$, and $|g_n - g|^2 \leq 2|g|^2 \in L^1(\sigma(T), \mu_\psi)$). So we can extend V_ψ to a unique surjective isometry $\widehat{V}_\psi : L^2(\sigma(T), \mu_\psi) \to \overline{H_\psi}$, whose inverse will be denoted U_ψ. Then $\overline{H_\psi} = \mathsf{H}$.

If $f \in M_b(\sigma(T))$, from (8.57) and using (iii) in Theorem 8.52(a) we see that

$$\int_{\sigma(T)} \overline{g_1(x,y)} f(x,y) g_2(x,y) \, d\mu_\psi(x,y)$$

$$= \left(\int_{\sigma(T)} g_1(x,y) \, dP^{(T)}(x,y) \psi \middle| \int_{\sigma(T)} f(x,y) g_2(x,y) \, dP^{(T)}(x,y) \psi \right)$$

$$= \left(\int_{\sigma(T)} g_1(x,y) \, dP^{(T)}(x,y) \psi \middle| \int_{\sigma(T)} f(x,y) \, dP^{(T)}(x,y) \int_{\sigma(T)} g_2(x,y) \, dP^{(T)}(x,y) \psi \right)$$

$$= \left(V_\psi g_1 \middle| \int_{\sigma(T)} f(x,y) \, dP^{(T)}(x,y) V_\psi g_2 \right) .$$

We have proved that for any triple $g_1, g_2, f \in M_b(\sigma(T))$

$$\int_{\sigma(T)} \overline{g_1(x,y)} f(x,y) g_2(x,y) \, d\mu_\psi(x,y) = \left(V_\psi g_1 \middle| \int_{\sigma(T)} f(x,y) \, dP^{(T)}(x,y) V_\psi g_2 \right) .$$

The operator $f \cdot : L^2(\sigma(T), \mu_\psi) \to L^2(\sigma(T), \mu_\psi)$, i.e. the multiplication by $f \in M_b(\sigma(T))$, is easily bounded; since $M_b(\sigma(T))$ is dense in $L^2(\sigma(T), \mu_\psi)$, by definition of U_ψ, because

$$\int_{\sigma(T)} f(x,y) \, dP^{(T)}(x,y)$$

is bounded and, eventually, by continuity of the inner product, we have

$$\int_{\sigma(T)} \overline{g_1(x,y)} f(x,y) g_2(x,y) \, d\mu_\psi(x,y) = \left(U_\psi^{-1} g_1 \middle| \int_{\sigma(T)} f(x,y) \, dP^{(T)}(x,y) U_\psi^{-1} g_2 \right) ,$$

for any $g_1, g_2 \in L^2(\sigma(T), \mu_\psi)$. That is to say

$$U_\psi \int_{\sigma(T)} f(x,y) \, dP^{(T)}(x,y) U_\psi^{-1} = f \cdot \, . \tag{8.59}$$

Now we pass to the case in which there is no ψ with $\overline{H_\psi} = \mathsf{H}$.

If so, let ψ be an arbitrary vector in H, and indicate by H_ψ the space of vectors $\int_{\sigma(T)} f(x,y)\,dP^{(T)}(x,y)\psi$, $f \in M_b(\sigma(T))$. We have $T(H_\psi) \subset H_\psi$ and $T^*(H_\psi) \subset H_\psi$, because for any $f \in M_b(\sigma(T))$

$$T\int_{\sigma(T)} f(x,y)\,dP^{(T)}\psi = \int_{\sigma(T)} (x+iy)\,dP^{(T)} \int_{\sigma(T)} f(x,y)\,dP^{(T)}\psi$$

$$= \int_{\sigma(T)} (x+iy)f(x,y)\,dP^{(T)}\psi$$

(Theorem 8.54(a) and (iii) in Theorem 8.52(a)). Hence $T\int_{\sigma(T)} f(x,y)\,dP^{(T)}\psi \in H_\psi$, for $(x,y) \mapsto (x+iy)f(x,y)$ is an element of $M_b(\sigma(T))$). The proof for T^* is analogous, using

$$T^* = \int_{\sigma(T)} (x-iy)\,dP^{(T)}\,.$$

By continuity $T(\overline{\mathsf{H}_\psi}) \subset \overline{\mathsf{H}_\psi}$ and $T^*(\overline{H_\psi}) \subset \overline{\mathsf{H}_\psi}$. Defining U_ψ as before we have (8.59).

Now let us show how to build another closed subspace $\overline{H_{\psi'}}$, orthogonal to $\overline{H_\psi}$, invariant under T, T^* and satisfying (8.59) for a corresponding surjective isometry $U_{\psi'} : \overline{H_{\psi'}} \to L^2(\sigma(T), \mu_{\psi'})$. If $0 \neq \psi' \perp \mathsf{H}_\psi$ then

$$\left(\psi' \middle| \int_{\sigma(T)} f(x,y)\,dP^{(T)}(x,y)\psi\right) = 0\,,$$

for any $f \in M_b(\sigma(T))$. But then the properties of the integral with respect to spectral measures ((iii)-(iv) in Theorem 8.52(a)) imply, for any $g, f \in M_b(\sigma(T))$:

$$\left(\int_{\sigma(T)} g\,dP^{(T)}\psi' \middle| \int_{\sigma(T)} f\,dP^{(T)}\psi\right) = \left(\psi' \middle| \int_{\sigma(T)} \overline{g}\,dP^{(T)} \int_{\sigma(T)} f\,dP^{(T)}\psi\right)$$

$$= \left(\psi' \middle| \overline{g(x,y)} f(x,y)\,dP^{(T)}(x,y)\psi\right) = 0\,,$$

where we used $\overline{g} \cdot f \in M_b(\sigma(T))$. Overall, if $\psi' \perp H_\psi$ then $H_{\psi'}$ is orthogonal to H_ψ, so the same holds for the closures by continuity of the scalar product. The space $\overline{H_{\psi'}}$ is invariant under T and T^* (the proof is the same as for $\overline{\mathsf{H}_\psi}$), and (8.59) holds for the surjective isometry $U_{\psi'} : \overline{H_{\psi'}} \to L^2(\sigma(T), \mu_{\psi'})$ (the proof was given at the beginning). Thus, choosing $\{\psi_\alpha\}$ suitably, we can construct closed subspaces $\mathsf{H}_\alpha = \overline{H_{\psi_\alpha}}$, each with a surjective isometry $U_\alpha : \overline{H_\alpha} \to L^2(\sigma(T), \mu_{\psi'})$, so that: (a) the spaces are pairwise orthogonal; (b) each one is T-invariant and T^*-invariant; (c) they satisfy

$$U_\alpha \int_{\sigma(T)} f(x,y)\,dP^{(T)}(x,y) \upharpoonright_{\mathsf{H}_\alpha} U_\alpha^{-1} = f\cdot \tag{8.60}$$

for any $f \in M_b(\sigma(T))$. Call $\mathfrak{C}$ the collection of these subspaces. We can partially order $\mathfrak{C}$ with the inclusion. Then every ordered subset in $\mathfrak{C}$ is upper bounded, and Zorn's lemma gives us a maximal element $\{\mathsf{H}_\alpha\}_{\alpha\in A}$ in $\mathfrak{C}$. We claim $\mathsf{H} = \oplus_{\alpha\in A}\mathsf{H}_\alpha$. It suffices to show that if ψ belongs to the orthogonal complement of every H_a, then $\psi = 0$. The trivial subspace $\mathsf{H}_0 := \{0\}$ (where μ_0 is the null measure – for which $L^2(\sigma(T),\mu_0)$ is trivial – and U_0 maps the null vector of $\{0\}$ to the null vector of $L^2(\sigma(T),\mu_0)$) is contained in $\{\mathsf{H}_\alpha\}_{\alpha\in A}$: if not, $\{\mathsf{H}_\alpha\}_{\alpha\in A}\cup\{\mathsf{H}_0\}$ would bound $\{\mathsf{H}_\alpha\}_{\alpha\in A}$ from above, a contradiction. If there existed $\psi\in\mathsf{H}$ with $\psi\perp\mathsf{H}_\alpha$ for any $\alpha\in A$ and $\psi\neq 0$, we would be able to construct $\overline{H_\psi}$, distinct from every H_α but satisfying (a), (b), (c). Consequently $\{\mathsf{H}_\alpha\}_{\alpha\in A}\cup\{\overline{H_\psi}\}$ would bound $\{\mathsf{H}_\alpha\}_{\alpha\in A}$ from above, a contradiction. We conclude that if ψ is orthogonal to every H_α, it must vanish $\psi=0$. Put equivalently, $\overline{<\{\mathsf{H}_\alpha\}_{\alpha\in A}>}=\mathsf{H}$, so $\mathsf{H}=\oplus_{\alpha\in A}\mathsf{H}_\alpha$ for the spaces are mutually orthogonal.

Now we prove (b) when T is normal (T self-adjoint or unitary is obtained specialising the same proof). We shall prove the logical equivalence: $\lambda\notin supp\{\mu_\alpha\}_{\alpha\in A}\Leftrightarrow\lambda\in\rho(T)$, equivalent to the claim.

$\Rightarrow$. If $\lambda\notin supp\{\mu_\alpha\}_{\alpha\in A}$, take D_R an open disc of radius $R>0$ centred at λ, with $\mu_\alpha(D_R)=0$ for any $\alpha\in A$; such a disc exists by the assumptions. On every space $L^2(\sigma(T),\mu_\alpha)$ the multiplication by $(x,y)\mapsto(x+iy-\lambda)^{-1}$ is bounded, with norm not exceeding $1/R$ (independent from α), and inverts (on the left and the right) the multiplication by $(x+iy-\lambda)$. Let $R_\lambda(\alpha):\mathsf{H}_\alpha\to\mathsf{H}_\alpha$ be the operator $U_\alpha^{-1}(x+iy-\lambda)^{-1}\cdot U_\alpha$. $R_\lambda(\alpha)$ has the same norm of the operator $(x+iy-\lambda)^{-1}\cdot$, since U_a is a surjective isometry, so $||R_\lambda(\alpha)||\leq 1/R$. Define $R_\lambda:\mathsf{H}\to\mathsf{H}$ so that

$$R_\lambda:\sum_{\alpha\in A}P_\alpha\psi\mapsto\sum_{\alpha\in A}R_\lambda(\alpha)P_\alpha\psi,$$

for any $\psi\in\mathsf{H}$. Remembering the H_α are invariant under T and R_λ (i.e. $R_\lambda(\alpha)$ on each one), we easily see that $||R_\lambda||\leq 1/R$ and $R_\lambda(T-\lambda I)=(T-\lambda I)R_\lambda=I$. In fact, $RanR_\lambda(\alpha)=\mathsf{H}_\alpha$ implies

$$||R_\lambda\psi||^2=||\sum_{\alpha\in A}R_\lambda(\alpha)P_\alpha\psi||^2=||\sum_{\alpha\in A}P_aR_\lambda(\alpha)P_\alpha\psi||^2=\sum_{\alpha\in A}||P_aR_\lambda(\alpha)P_\alpha\psi||^2$$

$$=\sum_{\alpha\in A}||R_\lambda(\alpha)P_\alpha\psi||^2\leq R^{-2}\sum_{\alpha\in A}||P_\alpha\psi||^2=R^{-2}||\psi||^2.$$

Moreover

$$(T-\lambda I)R_\lambda\psi=(T-\lambda I)R_\lambda\sum_{\alpha\in A}P_\alpha\psi$$

$$\sum_{\alpha\in A}(T-\lambda I)R_\lambda P_\alpha\psi=\sum_{\alpha\in A}(T-\lambda I)\restriction_{H_\alpha}R_\lambda(a)P_\alpha\psi=\sum_{\alpha\in A}IP_\alpha\psi=\psi,$$

hence $(T-\lambda I)R_\lambda=I$. Similarly

$$R_\lambda(T-\lambda I)\psi=R_\lambda(T-\lambda I)\sum_{\alpha\in A}P_\alpha\psi$$

$$\sum_{\alpha\in A}R_\lambda(T-\lambda I)P_\alpha\psi=\sum_{\alpha\in A}R_\lambda(a)(T-\lambda I)\restriction_{H_\alpha}P_\alpha\psi=\sum_{\alpha\in A}IP_\alpha\psi=\psi,$$

so $R_\lambda(T-\lambda I)=I$. By Theorem 8.4(a) $\lambda\in\rho(T)$.

$\Leftarrow$. Suppose now $\lambda\in\rho(T)$, so $(T-\lambda I)^{-1}:\mathsf{H}\to\mathsf{H}$ is the closed inverse to $T-\lambda I$. Pick $\varepsilon>0$ so that $||(T-\lambda I)^{-1}||=:1/\varepsilon$. We claim $\mu_\alpha(D_\varepsilon)=0$ for any $\alpha\in A$, D_ε

being the open disc of radius ε centred at λ. We proceed by contradiction. Suppose the last assertion is false, so there exists $\beta \in A$ such that $\mu_\beta(D_\varepsilon) > 0$. Let $D'_\delta \subset D_\varepsilon$ be another open disc, at a point of D_ε with radius δ, such that $0 < \delta < \varepsilon$ and $\mu_\beta(D_\delta) > 0$; if such a disc D_δ did not exist we would have $\mu_\beta(D_\varepsilon) = 0$ [2]. Consider a $\psi \in \mathsf{H} \setminus \{0\}$ defined by $P_\alpha \psi = 0$ if $\alpha \neq \beta$ and $U_\beta \psi = f$, such that $supp f \subset D_\delta$. We can always redefine ψ so that $||\psi|| = 1$ by multiplying it by the suitable factor. If $|x+iy-\lambda| < \varepsilon$ ($x+iy \in D_\delta$) then,

$$||(T-\lambda I)\psi||^2 = \int_{D_\delta} |(x+iy)-\lambda|^2 |f(x,y)|^2 \, d\mu_\beta(x,y) < \varepsilon^2 \int_{D_\delta} |f(x,y)|^2 \, d\mu_\beta(x,y) = \varepsilon^2 .$$

Therefore

$$||(T-\lambda I)\psi|| < \varepsilon .$$

On the other hand, by definition of norm,

$$||(T-\lambda I)^{-1}|| \geq \frac{||(T-\lambda I)^{-1}\phi||}{||\phi||}$$

for any $\phi \in \mathsf{H} \setminus \{0\}$. Setting $(T-\lambda I)^{-1}\phi = \psi$, we have

$$||(T-\lambda I)^{-1}|| \geq \frac{||\psi||}{||(T-\lambda I)\psi||} ,$$

so

$$1/\varepsilon \geq \frac{1}{||(T-\lambda I)\psi||} > 1/\varepsilon ,$$

because ψ is unit. But that is a contradiction.

We finish the proof by showing (c). If H is separable, consider the collection of orthogonal vectors $\{\psi_n\}_{n\in\mathbb{N}}$ built as the $\{\psi_\alpha\}_{\alpha\in A}$ above, where now the index α is called $n \in \mathbb{N}$. We are free to choose them so that $||\psi_n||^2 = 2^{-n}$. Define $M_T := \bigsqcup_{n=1}^{+\infty} \sigma(T)$ to be the disjoint union of infinitely many copies of $\sigma(T)$. Now call μ_T the measure that restricts to μ_n on the nth factor $\sigma(T)$. It is clear that, in this way, $\mu_T(M_T) = \sum_{n=0}^{+\infty} ||\psi_n||^2 < +\infty$. The map F_T clearly restricts to $(x+iy)\cdot$ on each component $\sigma(T)$. Hence F_T is bounded, because every copy of $\sigma(T)$ is bounded. The operator U_T is built in the obvious manner using the U_n. □

One can rearrange canonically the decomposition of H into spaces isomorphic to L^2. In particular ([DS88] vol.II) the following fact holds. In the statement we use the symbol $A\Delta B := (A\cup B)\setminus(A\cap B)$ for the symmetric difference of two sets.

[2] For any $z \in D_\varepsilon$ we can choose an open disc, centred at z of positive radius $\delta < \varepsilon$, so that $D_\delta \subset D_\varepsilon$. This gives an open covering of D_ε. Lindelöf's lemma (Theorem 1.8) guarantees we can extract a countable subcovering $\{D^{(i)}_{\delta_i}\}_{i\in\mathbb{N}}$. Since $D_\varepsilon = \cup_{i\in\mathbb{N}} D^{(i)}_{\delta_i}$ then, $\mu_\beta(E_\varepsilon) \leq \sum_{i\in\mathbb{N}} \mu_\beta(D^{(i)}_{\delta_i})$. If we had $\mu_\beta(D^{(i)}_{\delta_i}) = 0$ for any i, we would obtain $\mu_\beta(D_\varepsilon) = 0$.

Theorem 8.57. *Let $T \in \mathfrak{B}(\mathsf{H})$ be a normal operator on the separable Hilbert space* H.
(a) *There exists a pair $(\mu_T, \{E_{Tn}\}_{n\in\mathbb{N}})$, where μ_T is a positive, finite Borel measure on $\sigma(T)$ and $\{E_{Tn}\}_{n\in\mathbb{N}} \subset \mathscr{B}(\sigma(T))$ satisfies $\sigma(T) = E_{T1} \supset E_{T2} \supset \cdots$, so that Theorem 8.56(a) holds with $A = \mathbb{N}$, $\mu_\alpha(F) := \mu_T(F \cap E_{T\alpha})$ for any $\alpha \in \mathbb{N}$, $F \in \mathscr{B}(\sigma(T))$.*
(b) *If $(\mu'_T, \{E'_{Tn}\}_{n\in\mathbb{N}})$ satisfies part (a) and $\mu_T \prec \mu'_T \prec \mu_T$, then $\mu_T(E_{Tn}\Delta E'_{Tn}) = \mu'(E_{Tn}\Delta E'_{Tn}) = 0$ for any $n \in \mathbb{N}$.*
(c) *Let $S \in \mathfrak{B}(\mathsf{H})$ be a normal operator and suppose $(\{E_{Sn}\}_{n\in\mathbb{N}}, \mu_S)$ satisfies (a) together with S. Then there exists a unitary operator $U : \mathsf{H} \to \mathsf{H}$ with $USU^{-1} = T$ iff $\mu_T \prec \mu_S \prec \mu_T$ and $\mu_T(E_{Tn}\Delta E_{Sn}) = \mu_S(E_{Tn}\Delta E_{Sn}) = 0$ for any $n \in \mathbb{N}$.*

In case $\mu(E_{Tn-1}) \neq 0$ but $\mu(E_{Tn}) = 0$, one says T has **spectral multiplicity** n (including $n = +\infty$ if $\mu(E_{Tn-1}) \neq 0$ for any n). The definition is clearly independent of the pair $(\mu_T, \{E_{Tn}\}_{n\in\mathbb{N}})$ of (a). If $\mu_T(E_{Tk}) = 0$ for some k then $\mu'_T(E_{Tk}) = 0$, for $\mu_T \prec \mu'_T \prec \mu_T$. Since $\mu'_T(E_{Tk}\Delta E'_{Tk}) = 0$ we have $\mu'_T(E'_{Tk}) = 0$. By symmetry $\mu'_T(E'_{Tk}) = 0$ implies $\mu_T(E_{Tk}) = 0$. The unabridged theory of spectral multiplicity can be found in [Hal51].

Examples 8.58. (1) Consider a compact self-adjoint operator $T \in \mathfrak{B}(\mathsf{H})$ on the Hilbert space H. By Theorem 4.17, $\sigma_p(T)$ is discrete in $\mathbb{R}$, with possible unique limit point 0. Consequently $\sigma(T) = \sigma_p(T)$, except in case $\sigma_p(T)$ accumulates in 0, but $0 \notin \sigma_p(T)$. In that case ($\sigma(T)$ is closed by Theorem 8.4) $\sigma(T) = \sigma_p(T) \cup \{0\}$ and 0 is the *only point in* $\sigma_c(T)$ (for $\sigma_r(T) = \varnothing$ by Proposition 8.7). Following Example 8.49(3), we can define a PVM on $\mathbb{R}$ that vanishes outside $\sigma(T)$:

$$P_E := \text{s-}\sum_{\lambda\in E} P_\lambda$$

where $E \subset \sigma(T)$, while P_λ is either the null projector $P_\lambda = 0$, or an orthogonal projector on the eigenspace relative to λ. The former case can occur only when $\lambda = 0$ is no eigenvalue. Mimicking Example 8.49(3), we see

$$\int_{\sigma(T)} \lambda P(\lambda)\psi = \sum_{\lambda\in\sigma(T)} \lambda P_\lambda \psi\,,$$

for any $\psi \in \mathsf{H}$. On the other hand Theorem 4.18 gives

$$\sum_{\lambda\in\sigma(T)} \lambda P_\lambda = T\,,$$

where $P_0 = 0$ if $0 \in \sigma_c(T)$.

The statement of Theorem 4.18 explains that the decomposition is valid in the uniform topology provided we label eigenvalues properly. Using such an ordering, for any $\psi \in \mathsf{H}$

$$\sum_{\lambda\in\sigma(T)} \lambda P_\lambda \psi = T\psi\,.$$

We may interpret the sum as an integral in the projector-valued measure on $\sigma(T)$ defined above.This also proves that the series on the left can be rearranged (when projectors are applied to some $\psi \in \mathsf{H}$). By construction, $supp(P) = \sigma(T)$. In conclusion: the above measure on $\sigma(T)$ is the spectral measure of T, uniquely associated to T by the spectral theorem. Moreover, the spectral decomposition of T is precisely the eigenspace decomposition with respect to the strong topology:

$$T = \text{s-}\sum_{\lambda \in \sigma_p(T)} \lambda P_\lambda \,.$$

The point $0 \in \sigma_c(T)$, if present, brings no contribution to the integral.

(2) Consider the operator T on $\mathsf{H} := L^2([0,1] \times [0,1], dx \otimes dy)$ defined by

$$(Tf)(x,y) = xf(x,y)$$

almost everywhere on $\mathsf{X} := [0,1] \times [0,1]$, for any $f \in \mathsf{H}$. It is not hard to show T is bounded, self-adjoint and its spectrum is $\sigma(T) = \sigma_c(T) = [0,1]$.

A spectral measure on $\mathbb{R}$, with bounded support, that reproduces T as integral operator is given by orthogonal projectors $P_E^{(T)}$ that multiply by characteristic functions $\chi_{E'}$, $E' := (E \cap [0,1]) \times [0,1]$, for any Borel set $E \subset \mathbb{R}$. Proceeding as in Example 8.49(1), and choosing appropriate domains, allows to see that

$$\left(\int_{[0,1]} g(\lambda)\, P(\lambda) f \right)(x,y) = g(x) f(x,y) \,, \quad \text{almost everywhere on } \mathsf{X}$$

for any $g \in M_b(\mathsf{X})$. In particular

$$\left(\int_{[0,1]} \lambda\, P(\lambda) f \right)(x,y) = xf(x,y) \,, \quad \text{almost everywhere on } \mathsf{X} \,,$$

so

$$T = \int_{[0,1]} \lambda \, dP(\lambda) \,,$$

as required. This spectral measure is therefore the unique measure on $\mathbb{R}$ satisfying condition (a) in the spectral representation theorem.

We concern ourselves with part (c) in the spectral representation theorem. A decomposition of H of the kind prescribed in (c) can be obtained as follows. Let $\{u_n\}_{n\in\mathbb{N}}$ be a basis of $L^2([0,1], dy)$. Consider subspaces of $\mathsf{H} := L^2([0,1] \times [0,1], dx \otimes dy)$ given, for any $n \in \mathbb{N}$, by

$$\mathsf{H}_n := \{ f \cdot u_n \mid f \in L^2([0,1], dx) \} \,.$$

It is easy to see that these subspaces, with respect to T, fulfill every request of the theorem's item (c). In particular, H_n is by construction isomorphic to $L^2([0,1], dx)$ under the surjective isometry $f \cdot u_n \mapsto f$, so $\mu_n = dx$. ∎

8.5 Fuglede's theorem and consequences

In the final section we state and prove a well-known result, called *Fuglede's theorem*: it establishes that an operator $B \in \mathfrak{B}(\mathsf{H})$ commutes with A^*, for $A \in \mathfrak{B}(\mathsf{H})$

normal, provided it commutes with A. The result is far from obvious, and given the aforementioned theorems, it has immediate consequences. Due to the spectral decomposition 8.54(c), for example, one corollary is that if B commutes with A then it commutes with every operator $\int_{\sigma(T)} f(x,y)dP^{(A)}(x,y)$, for any measurable bounded map $f : \sigma(T) \to \mathbb{C}$.

8.5.1 Fuglede's theorem

Theorem 8.59 (Fuglede). *Let* H *be a Hilbert space. If* $A \in \mathfrak{B}(\mathsf{H})$ *is normal and* $B \in \mathfrak{B}(\mathsf{H})$ *commutes with* A*, then* B *commutes with* A^* *as well.*

Proof. For $s \in \mathbb{C}$ consider the function $Z(s) = e^{-sA^*}Be^{sA^*}$, where the exponentials are spectrally defined by integrals of $\mathbb{C} \ni x+iy \mapsto e^{\mp s(x-iy)}$ with respect to the spectral measure $P^{(A)}$ of A. As usual $z = x+iy$ and $\overline{z} = x-iy$. Now observe $e^{\mp s(x-iy)} = \sum_{n=0}^{+\infty} \frac{(\mp s(x-iy))^n}{n!}$, and for given s, the convergence is uniform in (x,y) on every compact set, like $\sigma(A)$. In particular this means the sequence of partial sums is bounded in norm $|| \ ||_\infty$. Using again the PVM spectrally associated to A, by Theorem 8.52(c) we have

$$e^{\mp sA^*} = \text{s-}\sum_{n=0}^{+\infty} \frac{(\mp sA^*)^n}{n!} . \tag{8.61}$$

Expanding $Z(s)A\psi$ and $AZ(s)\psi$ as above, and recalling A^* and B commute with A, we see $A^nZ(s)\psi = Z(s)A^n\psi$ for any $\psi \in \mathsf{H}$. Hence the exponential expansion gives

$$e^{\mp\overline{s}A}Z(s)\psi = Z(s)e^{\mp\overline{s}A}\psi \quad \text{for any } \psi \in \mathsf{H}.$$

Therefore

$$Z(s) = Z(s)e^{+\overline{s}A}e^{-\overline{s}A} = e^{+\overline{s}A}Z(s)e^{-\overline{s}A} = e^{-sA^*}e^{+\overline{s}A}Be^{sA^*}e^{-\overline{s}A} = e^{-sA^*+\overline{s}A}Be^{sA^*-\overline{s}A} .$$

To obtain this we need the identities $e^{-sA^*}e^{+\overline{s}A} = e^{-sA^*+\overline{s}A}$ and $e^{+\overline{s}A}e^{-\overline{s}A} = I$, which are proved exactly as in $\mathbb{C}$, i.e. using the expansion (8.61) and the commutation of A, A^*. With the same technique one proves $U_s := e^{-sA^*+\overline{s}A} = \left(e^{sA^*-\overline{s}A}\right)^*$ and $U_s^* = U_s^{-1}$. Therefore U_s is unitary and $||Z(s)|| = ||U_sBU_s^*|| \leq ||U_s||\,||B||\,||U_s^*|| = 1||B||1 = ||B||$. The map $\mathbb{C} \ni s \mapsto (\psi|Z(s)\phi)$ is then bounded on the entire complex plane. If this function were entire (i.e. analytic on $\mathbb{C}$), Liouville's theorem would force it to be constant. So let us assume the map is entire, hence constant. Consequently, since ψ, ϕ are arbitrary, $Z(s) = Z(0)$ for any $s \in \mathbb{C}$. This identity reads $e^{-sA^*}Be^{sA^*} = B$, i.e. $Be^{sA^*} = e^{sA^*}B$. Therefore $(\psi|Be^{sA^*}\phi) = (\psi|e^{sA^*}B\phi)$ for any $\psi, \phi \in \mathsf{H}$. This equation can be written $(B^*\psi|e^{sA^*}\phi) = (\psi|e^{sA^*}B\phi)$, and by the properties of PVMs and the spectral theorem:

$$\int_K e^{s\overline{z}}d\mu_{B^*\psi,\phi} = \int_K e^{s\overline{z}}d\mu_{\psi,B\phi} ,$$

where $K \subset \mathbb{R}^2 \equiv \mathbb{C}$ is a compact set large enough to contain the supports of the measures of the integrals. Let us differentiate in s, and evaluate at $s = 0$, by swapping derivative and integral (the derivatives of the integrands are continuous in $(s,(x,y))$, hence bounded on the compact set $C \times K$, with C a compact subset containing $s = 0$; thus Theorem 1.87 applies). The outcome is

$$\int_K \bar{z} d\mu_{B^*\psi,\phi} = \int_K \bar{z} d\mu_{\psi,B\phi} \,,$$

which we can write $(\psi|BA^*\phi) = (\psi|A^*B\phi)$. Varying ψ and ϕ, we obtain B commutes with A^*: $BA^* = A^*B$.

There remains to prove $\mathbb{C} \ni s \mapsto (\psi|Z(s)\phi)$ is an analytic function. The expansion (8.61) and the inner product's continuity imply

$$(\psi|Z(s)\phi) = \sum_{n=0}^{+\infty}\sum_{m=0}^{+\infty} \frac{(-s)^{n+m}}{n!m!}(\psi|(A^*)^n B(A^*)^m\phi)\,. \tag{8.62}$$

We may interpret the double series as an iterated integral *for* $\mathbb{N}$*'s counting measure*; we shall denote the latter by $d\mu(n)$. By Schwarz's inequality and known norm properties:

$$\left|\frac{(-s)^{n+m}}{n!m!}(\psi|(A^*)^n B(A^*)^m\phi)\right| \leq \frac{(|s|\,||A||)^n}{n!}\frac{(|s|\,||A||)^m}{m!}||B||\,||\psi||\,||\phi||\,.$$

The positive function on $\mathbb{N}\times\mathbb{N}$ of the right-hand side is integrable in the product measure (the integral is clearly $e^{|s|\,||A||}e^{|s|\,||A||}||B||\,||\psi||\,||\phi||$), by Fubini–Tonelli, so $(n,m) \mapsto \frac{(-s)^{n+m}}{n!m!}(\psi|(A^*)^n B(A^*)^m\phi) =: f_s(n,m)$ is L^1 for the product masure, and (8.62) reads:

$$(\psi|Z(s)\phi) = \int_{\mathbb{N}\times\mathbb{N}} f_s(n,m)d\mu(n)\otimes d\mu(m)\,. \tag{8.63}$$

By dominated convergence we have

$$\int_{\mathbb{N}\times\mathbb{N}} f_s(n,m)d\mu(n)\otimes d\mu(m) = \lim_{N\to+\infty}\int_{\mathbb{N}\times\mathbb{N}} \chi_{\{(n,m)\in\mathbb{N}\times\mathbb{N}\,|\,n+m\leq N\}} f_s(n,m)d\mu(n)\otimes d\mu(m)\,.$$

Writing the right side using sums:

$$(\psi|Z(s)\phi) = \lim_{N\to+\infty}\sum_{M=0}^{N}\sum_{n+m=M}\frac{(-s)^{n+m}}{n!m!}(\psi|(A^*)^n B(A^*)^m\phi)\,,$$

i.e.

$$(\psi|Z(s)\phi) = \sum_{N=0}^{+\infty} C_N s^N \quad \forall s\in\mathbb{C}\,, \tag{8.64}$$

where

$$C_N = (-1)^N \sum_{n+m=N}\frac{(\psi|(A^*)^n B(A^*)^m\phi)}{n!m!}\,.$$

The series (8.64) says we may express $(\psi|Z(s)\phi)$ as a power series in s, with s roaming the whole complex plane. Hence $\mathbb{C} \ni s \mapsto (\psi|Z(s)\phi)$ is an entire function, as claimed. □

The theorem generalises by dropping the boundedness of A (but keeping B's). This was Fuglede's original statement [Fug50], whose proof requires the spectral theory of unbounded normal operators that we will not develop.

8.5.2 Consequences to Fuglede's theorem

Corollary 8.60. *If $M,N \in \mathfrak{B}(\mathsf{H})$, H Hilbert space, are normal and satisfy $NM = MN$, then NM is normal.*

Proof. $MN(MN)^* = MNM^*N^*$. By Fuglede's theorem the right-hand side is $MM^*NN^* = M^*MN^*N = M^*N^*MN = (NM)^*MN$. But N,M commute, so $(NM)^*MN = (MN)^*MN$. Hence we have proved $MN(MN)^* = (MN)^*MN$, i.e. the claim. □

Corollary 8.61 (Fuglede–Putnam–Rosenblum). *Let $T,M,N \in \mathfrak{B}(\mathsf{H})$, H Hilbert space. If M,N are normal and $MT = TN$ then $M^*T = TN^*$.*

Proof. Consider the Hilbert space $\mathsf{H} \oplus \mathsf{H}$ with standard inner product $((u,v)|(u',v')) := (u|u')(v|v')$, and hence the operators of $\mathfrak{B}(\mathsf{H}\oplus\mathsf{H})$:

$$T' : (u,v) \mapsto (0,Tu)\,, \quad N' : (u,v) \mapsto (Nu,Mv)\,.$$

By direct computation $N'N'^* = N'N'^*$, i.e. N' is normal, and $N'T' = T'N'$ by the fact that $MT = TN$. We can apply Fuglede's theorem to get $N'^*T' = T'N'^*$. Since $N'^* : (u,v) \mapsto (N^*u, M^*v)$, taking the components of the identity $N'^*T'(u,v) = T'N'^*(u,v)$ gives $M^*Tu = TN^*u$ for any $u \in \mathsf{H}$, i.e. $M^*T = TN^*$. □

Corollary 8.62. *Let $M,N \in \mathfrak{B}(\mathsf{H})$ be normal operators on the Hilbert space H. If there is a bijection $S \in \mathfrak{B}(\mathsf{H})$ such that*

$$MS = SN\,,$$

then there is also a operator unitary $U \in \mathfrak{S}(\mathsf{H})$ such that

$$UMU^{-1} = N\,.$$

Proof. Observe preliminarly $S^{-1} \in \mathfrak{B}(\mathsf{H})$ by Theorem 2.92. By polar decomposition $S = U|S|$, with U unitary; therefore $MU|S| = U|S|N$. In our case $|S|^{-1}$ exists and equals $|S^{-1}|$, as is easy to see. The proof finishes if we can show that $|S|N = N|S|$, for then we can left-apply $|S|^{-1}$ on $MU|S| = UN|S|$. Let us prove that. By Putnam's theorem $MS = SN$ implies $M^*S = SN^*$; taking adjoints, $S^*M = NS^*$. Using $MS = SN$ again, we get $S^*MS = S^*SN = NS^*S$, i.e. $|S|^2N = N|S|^2$. By Theorem 3.66(b), $|S|N = N|S|$. □

Exercises

8.1. Let $\mathsf{H} = \ell^2(\mathbb{N})$ and consider the operator

$$T : (x_0, x_1, x_2, \ldots) \mapsto (0, x_0, x_1, \ldots) .$$

Determine the spectrum of T.

8.2. Let H be a Hilbert space and $T = T^* \in \mathfrak{B}(\mathsf{H})$ be compact. Show that if $dim(RanT)$ is not finite, then $\sigma_c(T) \neq \varnothing$ and consists of one point.

Hint. Decompose T as in Theorem 4.18, use Theorem 4.17 and the fact that $\sigma(T)$ is closed by Theorem 8.4.

8.3. If T is self-adjoint on the Hilbert space H and $\lambda \in \rho(T)$, show $R_\lambda(T)$ is a normal operator of $\mathfrak{B}(\mathsf{H})$ such that

$$R_\lambda(T)^* = R_{\overline{\lambda}}(T) .$$

8.4. Prove that the residual spectrum of a unitary operator is empty, without using the fact that unitary implies normal.

Solution. If $\lambda \in \sigma_r(U)$, $Ran(U - \lambda I)$ is not dense, so there exists $f \neq 0$ orthogonal to $Ran(U - \lambda I)$. For any $g \in \mathsf{H}$, $(f|\lambda g) = (f|Ug)$, so $(\overline{\lambda} f|g) = (U^* f|g)$ for any $g \in \mathsf{H}$. Hence $U^* f = \overline{\lambda} f$. Letting U act on either side gives $f = \overline{\lambda} U f$, and then $Uf = \lambda f$, because $1/\lambda = \overline{\lambda}$ by $|\lambda| = 1$. Consequently $\lambda \in \sigma_p(U)$; but this is absurd, for point spectrum and residual spectrum are disjoint; hence $\sigma_r(U) = \varnothing$.

8.5. If $U : \mathsf{H} \to \mathsf{H}$ is an isometry on a Hilbert space H that is *not surjective*, then $\sigma_r(U) \neq \varnothing$.

Solution. $0 \in \sigma_r(U)$. $U - 0I$ is one-to-one, but $Ran(U - 0I) = RanU$ is not dense. Let us prove that by contradiction. If it were dense, for any $f \in \mathsf{H}$ there would be $\{f_n\}_{n\in\mathbb{N}} \subset \mathsf{H}$ with $Uf_n \to f$. Since $||f_n - f_m|| = Uf_n - Uf_m$, $\{f_n\}$ would be a Cauchy sequence, $f_n \to g \in \mathsf{H}$. Hence $Ug = f$ for any $f \in \mathsf{H}$, which cannot be, for U is not surjective.

8.6. Build a self-adjoint operator with point spectrum dense in, but not coinciding with, $[0, 1]$.

Hint. Take the Hilbert space $\mathsf{H} = \ell^2(\mathbb{N})$, and label rationals in $[0, 1]$ arbitrarily: $r_0, r_1, \ldots$ Define

$$T : (x_0, x_1, x_2, \ldots) \mapsto (r_0 x_0, r_1 x_1, r_2 x_2, \ldots)$$

with domain $D(T)$ given by sequences $(x_0, x_1, x_2, \ldots) \in \ell^2(\mathbb{N})$ such that

$$\sum_{n=0}^{+\infty} |r_n x_n|^2 < +\infty .$$

8.7. Define a bounded normal operator $T : \mathsf{H} \to \mathsf{H}$, for some Hilbert space H, such that $\sigma(T) = \sigma_p(T) = \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$. Can H be separable?

Hint. Define $\mathsf{H} := L^2(D, \mu)$, where μ is the counting measure and $D := \{\lambda \in \mathbb{C} \mid |\lambda| \leq 1\}$. Then set $(Tf)(z) := zf(z)$, $f \in \mathsf{H}$.

8.8. If $P : \mathsf{X} \to \mathfrak{B}(\mathsf{H})$ is a PVM, prove: (1) the set of P-essentially bounded, measurable maps $f : \mathsf{X} \to \mathbb{C}$ is a vector space, and (2) $|| \; ||_\infty^{(P)}$ is a seminorm on that space.

8.9. Let A be an operator on the Hilbert space H with domain $D(A)$, and let $U : \mathsf{H}' \to \mathsf{H}$ be a Hilbert space isometry onto H. If $A' := U^{-1}AU : D(A') \to \mathsf{H}'$, $D(A') = U^{-1}D(A)$, prove $\sigma_c(A) = \sigma_c(A')$, $\sigma_p(A) = \sigma_p(A')$, $\sigma_r(A) = \sigma_r(A')$.

8.10. Conside the position operator X_i introduced in Definition 5.22. Show $\sigma(X_i) = \sigma_c(X_i) = \mathbb{R}$.

8.11. Consider the momentum operator P_i introduced in Definition 5.27. Show $\sigma(P_i) = \sigma_c(P_i) = \mathbb{R}$.

Hint. Use Proposition 5.31.

8.12. Find two operators A and B on a Hilbert space such that $\sigma(A) = \sigma(B)$, but $\sigma_p(A) \neq \sigma_p(B)$.

Hint. Consider the operator of Exercise 8.6 and the operator that multiplies by the coordinate x in $L^2([0,1], dx)$, where dx is Lebesgue's measure on $\mathbb{R}$.

8.13. Take Volterra's operator $A : L^2([0,1], dx) \to L^2([0,1], dx)$:

$$(Af)(x) = \int_0^x f(t)dt \, .$$

Study its spectrum and prove $\sigma(A) = \sigma_c(A) = \{0\}$. Conclude, without computations, that A cannot be normal.

Outline of solution. Since $[0,1]$ has finite Lebesgue measure, then $L^2([0,1], dx) \subset L^1([0,1], dx)$, and we can view Lebesgue's integral as a function of the upper limit of integration (in particular, Theorem 1.75 holds). Notice the spectrum of A cannot be empty by Theorem 8.4, since A is bounded and hence closed. If $\lambda \neq 0$, then $(\lambda^{-1}A)^n$ is a contraction operator for n large enough, as we saw in Exercise 4.19. By the fixed point theorem $\lambda \neq 0$ cannot be an eigenvalue, since the unique solution ψ to the characteristic equation $\lambda^{-1}A\psi = \psi$ is $\psi = 0$, not an eigenvector. As A is compact, Lemma 4.49 guarantees that if $0 \neq \lambda$ (hence $\lambda \notin \sigma_p(A)$), then $Ran(A - \lambda I) = \mathsf{H}$ (i.e. the Hilbert space $L^2([0,1], dx)$); moreover, since $A - \lambda I$ is bijective, $(A - \lambda I)^{-1} : \mathsf{H} \to \mathsf{H}$ is bounded by the inverse-operator theorem. Therefore $\lambda \notin \sigma(A)$ if $\lambda \neq 0$. So the unique point in the spectrum is $\lambda = 0$. By Theorem 1.75(b) there are no non-zero solutions to $A\psi = 0$, and we conclude $0 \in \sigma_r(A) \cup \sigma_c(A)$. If 0 were in $\sigma_r(A)$, $Ran(A)$ would not be dense in $L^2([0,1],dx)$, i.e. $Ker(A^*) \neq \{0\}$ because $\mathsf{H} = \overline{Ran(A)} \oplus Ker(A^*)$. This is not possible, because $(A^* f)(x) = \int_x^1 f(t)dt$ (see Exercise 3.26), so applying Theorem 1.75(b) would give a contradiction.

If A were normal, as bounded we would get $||A|| = r(A)$. But $r(A) = 0$, for $\sigma(A) = \{0\}$. Therefore A would be forced to be null.

8.14. Consider the bounded, self-adjoint operator T on $\mathsf{H} := L^2([0,1],dx)$ that multiplies functions by x^2:

$$(Tf)(x) := x^2 f(x) .$$

Find its spectral measure.

Hint. Find a unitary transformation from H to $L^2([0,1],dy)$ that maps the multiplication by x^2 to the multiplication by y.

8.15. Consider the bounded, self-adjoint operator T on $\mathsf{H} := L^2([-1,1],dx)$ that multiplies by x^2:

$$(Tf)(x) := x^2 f(x) .$$

Determine its spectral measure.

Hint. Argue as in Exercise 8.14, after having decomposed

$$L^2([-1,1],dx) = L^2([-1,0],dx) \oplus L^2([0,1],dx) .$$

8.16. Let $T \in \mathfrak{B}(\mathsf{H})$, H a Hilbert space, be a normal operator. Prove, for any $\alpha \in \mathbb{C}$, that

$$e^{\alpha T} = \int_{\sigma(T)} e^{\alpha(x+iy)}\, dP^{(T)}(x,y) ,$$

where the term on the left is defined, in *uniform topology*, as

$$e^{\alpha T} := \sum_{n=0}^{+\infty} \frac{\alpha^n T^n}{n!} .$$

Hint. The series $\sum_{n=0}^{+\infty} \frac{\alpha^n z^n}{n!}$ converges absolutely and uniformly on any closed disc of finite radius and centred at the origin of $\mathbb{C}$. Moreover, for any polynomial $p(z)$ ($z = x + iy$),

$$p(T) = \int_{\sigma(T)} p(x+iy)\, dP^{(T)}(x,y) .$$

Now use the first part of Theorem 8.50.

8.17. For any given Hilbert space H, build a compact self-adjoint operator $T : \mathsf{H} \to \mathsf{H}$ such that $T \notin \mathfrak{B}_1(\mathsf{H})$, $T \notin \mathfrak{B}_2(\mathsf{H})$.

Hint. It suffices to show $\sum_{\lambda \in \sigma_p(T)} |\lambda| = +\infty$ and $\sum_{\lambda \in \sigma_p(T)} |\lambda|^2 = +\infty$, see Exercise 4.4.

8.18. Take $T \in \mathfrak{B}(\mathsf{H})$ with $T \geq 0$ and H Hilbert space. Prove that if T is compact then

$$T^\alpha := \int_{\sigma(T)} \lambda^\alpha dP^{(T)}(\lambda)$$

is compact for any real $\alpha > 0$.

Outline of solution. If $\sigma(T)$ is finite the claim is obvious by the spectral theorem and because operators with finite-dimensional range are compact. Consider the other case. Expand T spectrally: $T = \sum_j \lambda_j (\psi_j | \ )\psi_j$, where $||T|| \geq \lambda_j \geq \lambda_{j+1} \to 0_+$ by T's compactness, and for any given j, $\lambda_{j+k} = \lambda_j$ only for a finite number of k (if the eigenvalue is non-null). Recall that for compact operators the expansion converges in uniform topology too. If $\alpha \geq 1$ and m,n are large enough, then $\sum_{j=n}^m \lambda_j^\alpha (\psi_j | \ )\psi_j \leq \sum_{j=n}^m \lambda_j (\psi_j | \ )\psi_j$; hence (positive operators are self-adjoint) $||\sum_{j=n}^m \lambda_j^\alpha (\psi_j | \ )\psi_j|| \leq ||\sum_{j=n}^m \lambda_j (\psi_j | \ )\psi_j||$. Bearing in mind the spectral decomposition theorem for the self-adjoint T^α, conclude that T^α, $\alpha > 1$, is compact as uniform limit of compact operators (their range is finite-dimensional). When $\alpha < 1$, observe $||T^{1/2}x - T^{1/2}y||^2 = ((x-y)|T(x-y)) \leq ||x-y||||Tx - Ty||$; conclude that $T^{1/2}$ is compact if T is by using the definition of compact operator. When $\alpha \in [1/2, 1)$, $T^\alpha = (T^{1/2})^\beta$ for some $\beta \in [1,2)$. Relying on the previous proof recover that T^α is compact if T is when $\alpha \in [1/2, 1)$. Iterating the procedure obtain that $T^{1/4}$ and T^α, $\alpha \in [1/4, 1/2)$, are compact if T is, and so on; hence reach any T^α, with $\alpha \in (0,1)$, because in that case $\alpha \in [1/2^{k+1}, 1/2^k)$ for some $k = 0, 1, 2, \ldots$.

9

Spectral theory II: unbounded operators on Hilbert spaces

The language of mathematics turns out to be unreasonably effective in natural sciences, a wounderful gift that we don't understand nor deserve.

Eugene Paul Wigner

With this chapter we shall examine a number of mathematical issues concerning the spectral theory of self-adjoint operators, in general unbounded.

The first section is devoted to extending the spectral decomposition theorem of the previous chapter to unbounded self-adjoint operators. To do so we will generalise the integration procedure for spectral measures to *unbounded* functions. Using this and the Cayley transform we will prove the spectral decomposition theorem for unbounded self-adjoint operators. The resulting technique will also enable us to prove, in passing, and important characterisation of the von Neumann algebra generated by a bounded normal operator and its adjoint. Then we will describe two physically-relevant examples of unbounded self-adjoint operators and their spectral decomposition: the Hamiltonian of the harmonic oscillator, and the position and momentum operators. Finally we will state a spectral representation theorem for unbounded self-adjoint operators and introduce the notion of *joint spectral measure*.

The second, very short, section is dedicated to exponentiating unbounded operators, in relationship to earlier-defined analytic vectors.

In section three we will focus on the theory of strongly continuous one-parameter groups of unitary operators. First we will establish that the various notions of continuity are equivalent. Next we will show *von Neumann's theorem* on the continuity of measurable one-parameter groups of unitary operators, and then go on to prove *Stone's theorem* and a few important corollaries. In particular, we will discuss a useful criterion, based on one-parameter unitary groups generated by self-adjoint operators, to establish when the spectral measures of two self-adjoint operators commute.

9.1 Spectral theorem for unbounded self-adjoint operators

We now generalise some of the material of Chapter 8. In particular we want to prove the spectral decomposition theorem in the case of unbounded self-adjoint operators. The physical relevance lies in that most self-adjoint operators representing interesting observables in Quantum Mechanics are *unbounded*. The paradigmatic case is the position operator of Chapter 5.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_9,

9.1.1 Integrating unbounded functions with respect to spectral measures

We will often use the following natural definition.

Definition 9.1. *Let* X *be a complex vector space,* T *an operator on* X *with domain* $D(T)$ *and* $p(x) = \sum_{k=0}^{m} a_k x^k$ *a polynomial of degree m with complex coefficients.*

(a) *The operator* $p(T)$ *on* X *is defined by writing* T *in place of the variable x, with* $T^0 := I$*,* $T^1 := T$*,* $T^2 := TT$*, and so on.*

(b) *The domain of* $p(T)$ *is*

$$D(p(T)) := \bigcap_{k=0}^{m} D(a_k T^k)\,, \tag{9.1}$$

with $D(a_k T \cdots T)$ *given in Definition 5.1.*

Extending the previous chapter's results to unbounded operators requires first a definition for the integral of *unbounded* functions with respect to a PVM. If P is a spectral measure on the second-countable space X, in the sense of Definition 8.41, and if $f : \mathsf{X} \to \mathbb{C}$ is a measurable function (for the Borel algebra of X), but not necessarily bounded, then $\int_{\mathsf{X}} f(x)dP(x)$ is as-of-yet meaningless. The whole point is to make sense of this integral.

Consider a vector ψ, in a Hilbert space H, of the PVM P such that

$$\int_{\mathsf{X}} |f(x)|^2 \, d\mu_\psi(x) < +\infty\,, \tag{9.2}$$

where the spectral measure for ψ, μ_ψ, was defined in Theorem 8.50(c). We can find a sequence of bounded measurable maps f_n such that $f_n \to f$ as $n \to +\infty$ in $L^2(\mathsf{X}, \mu_\psi)$. For example, using Lebesgue's dominated convergence it suffices to consider $f_n := \chi_{F_n} \cdot f$, where χ_{F_n} is the characteristic function of $F_n := \{x \in \mathsf{X} \mid |f(x)| < n\}$. Using (iii) in (a) and (b) of Theorem 8.52 we immediately find

$$\left|\left| \int_{\mathsf{X}} f_n(x)dP(x)\psi - \int_{\mathsf{X}} f_m(x)dP(x)\psi \right|\right|^2 = \int_{\mathsf{X}} |f_n(x) - f_m(x)|^2 d\mu_\psi(x)\,. \tag{9.3}$$

Thus the sequence of vectors $\int_{\mathsf{X}} f_n(x)dP(x)\psi$ converges to some $\int_{\mathsf{X}} f(x)dP(x)\psi$:

$$\int_{\mathsf{X}} f(x)dP(x)\psi := \lim_{n\to+\infty} \int_{\mathsf{X}} f_n(x)dP(x)\psi\,. \tag{9.4}$$

We may use (9.4) as *definition* of the integral in P of the unbounded measurable function f. This is *well defined* since $\int_{\mathsf{X}} f(x)dP(x)\psi$ *does not* depend of the sequence $\{f_n\}_{n\in\mathbb{N}}$. In fact if $\{g_n\}_{n\in\mathbb{N}}$ is another sequence of bounded measurable maps converging to f in $L^2(\mathsf{X}, \mu_\psi)$, proceeding as before we obtain

$$\left|\left| \int_{\mathsf{X}} f_n(x)dP(x)\psi - \int_{\mathsf{X}} g_n(x)dP(x)\psi \right|\right|^2 = \int_{\mathsf{X}} |f_n(x) - g_n(x)|^2 d\mu_\psi(x)\,,$$

so

$$\lim_{n\to+\infty}\int_{\mathsf{X}} f_n(x)dP(x)\psi = \lim_{n\to+\infty}\int_{\mathsf{X}} g_n(x)dP(x)\psi\,.$$

If we use (9.4) to define the integral of an unbounded function we have to recall this operator is not defined on the whole Hilbert space, but only on vectors satisfying (9.2). Consequently we have to check these vectors form a subspace in H. To show this, and much more, we need to a lemma that relates the spectral measure μ_ψ to $\mu_{\phi,\psi}$ via (9.2), for $\psi \in \mathsf{H}$.

Lemma 9.2. *Let* X *be a second-countable space,* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ *a PVM,* H *a Hilbert space, and* $f : \mathsf{X} \to \mathbb{C}$ *a measurable function.*
Given $\phi, \psi \in \mathsf{H}$*, if the measures* μ_ψ *and* $\mu_{\phi,\psi}$ *are defined as in Theorem 8.50 and*

$$\int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x) < +\infty\,,$$

then $f \in L^1(\mathsf{X}, |\mu_{\phi,\psi}|)$ *and*

$$\int_{\mathsf{X}} |f(x)|d|\mu_{\phi,\psi}|(x) \leq ||\phi||\,\sqrt{\int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x)}\,. \tag{9.5}$$

Proof. If f is bounded, by (iv) in Theorem 8.50(c):

$$\left(\phi\left|\int_{\mathsf{X}} |f(x)|dP(x)\psi\right.\right) = \int_{\mathsf{X}} |f(x)|d\mu_{\phi,\psi}(x)\,.$$

From Theorem 1.86 there exists a map $h : \mathsf{X} \to \mathbb{C}$, $|h(x)| = 1$, such that $d\mu_{\phi,\psi} = hd|\mu_{\phi,\psi}|$, and so

$$\int_{\mathsf{X}} |f(x)|d|\mu_{\phi,\psi}|(x) = \int_{\mathsf{X}} |f(x)|h^{-1}(x)d\mu_{\phi,\psi}(x) = \left(\phi\left|\int_{\mathsf{X}} |f(x)|h^{-1}(x)dP(x)\psi\right.\right)\,.$$

Using Theorem 8.52(b) and noting $||f(x)|h^{-1}(x)|^2 = |f(x)|^2$, we have

$$\int_{\mathsf{X}} |f(x)|d|\mu_{\phi,\psi}|(x) \leq ||\phi||\left|\left|\int_{\mathsf{X}} |f(x)|h^{-1}(x)dP(x)\psi\right|\right| = ||\phi||\,\sqrt{\int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x)}\,.$$

Let now f be unbounded. Define bounded maps $f_n := \chi_{F_n} \cdot f$ as above, so that $0 \leq |f_n(x)| \leq |f_{n+1}(x)| \to |f(x)|$ as $n \to +\infty$. By monotone convergence, since $f \in L^2(\mathsf{X}, d\mu_\psi)$, we obtain

$$\int_{\mathsf{X}} |f(x)|d|\mu_{\phi,\psi}|(x) = \lim_{n\to+\infty}\int_{\mathsf{X}} |f_n(x)|d|\mu_{\phi,\psi}|(x) \leq ||\phi||\lim_{n\to+\infty}\sqrt{\int_{\mathsf{X}} |f_n(x)|^2 d\mu_\psi(x)}$$

$$= ||\phi||\,\sqrt{\int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x)} < +\infty\,.$$

This proves the general case. □

The next theorem gathers several technical facts seen above, and establishes the first general properties of integrals of unbounded maps with respect to a spectral measure.

Theorem 9.3. *Let* X *be second-countable with Borel* σ*-algebra* $\mathscr{B}(\mathsf{X})$, H *a Hilbert space and* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ *a PVM.*
For any measurable $f : \mathsf{X} \to \mathbb{C}$ *define*

$$\Delta_f := \left\{ \psi \in \mathsf{H} \;\middle|\; \int_{\mathsf{X}} |f(x)|^2 \, d\mu_\psi(x) < +\infty \right\} . \tag{9.6}$$

(a) Δ_f *is a dense subspace in* H.
(b) *Given a sequence of bounded measurable maps* $\{f_n\}_{n\in\mathbb{N}}$ *converging to* f *in* $L^2(\mathsf{X}, \mu_\psi)$*, the mapping*

$$\int_{\mathsf{X}} f(x) dP(x) : \Delta_f \ni \psi \mapsto \int_{\mathsf{X}} f(x) dP(x) \psi \,, \tag{9.7}$$

with right-hand-side term as in (9.4), is a linear operator.
(c) *If* $f\restriction_{supp(P)}$ *is bounded:*

$$\Delta_f = \mathsf{H} \quad and \quad \int_{\mathsf{X}} f(x) dP(x) = \int_{supp(P)} f(x) dP(x) \in \mathfrak{B}(\mathsf{H}) \,,$$

where the right side is the operator of Definition 8.47(c).

Proof. (a) and (b) As first thing we have to prove, for any given measurable $f : \mathsf{X} \to \mathbb{C}$, that $\phi + \psi \in \Delta_f$ and $c\phi \in \Delta_f$ for any $c \in \mathbb{C}$ if $\phi, \psi \in \Delta_f$. Note Δ_f contains the null vector of H, so it is non-empty.
If $\phi, \psi \in \Delta_f$, $E \in \mathscr{B}(\mathsf{X})$:

$$||P_E(\phi + \psi)||^2 \leq (||P_E\phi|| + ||P_E\psi||)^2 \leq 2||P_E\phi||^2 + 2||P_E\psi||^2 \,;$$

since $\mu_\chi(E) = (\chi|P_E\chi) = (\chi|P_E P_E\chi) = (P_E\chi|P_E\chi) = ||P_E\chi||^2$:

$$\mu_{\phi+\psi}(E)^2 \leq 2(\mu_\phi(E) + \mu_\psi(E)) \,.$$

This implies, for $L^2(\mathsf{X}, \mu_\phi) \ni f$ and $L^2(\mathsf{X}, \mu_\psi) \ni f$, that $L^2(\mathsf{X}, \mu_{\phi+\psi}) \ni f$. I.e. $\phi, \psi \in \Delta_f$ forces $\phi + \psi \in \Delta_f$. On the other hand $\mu_{c\phi}(E) = |c|^2 (P_E\phi|\phi) = |c|^2 \mu_\phi(E)$, so $f \in L^2(\mathsf{X}, \mu_{c\phi})$ for $f \in L^2(\mathsf{X}, \mu_\phi)$ and $c \in \mathbb{C}$. That is to say $\phi \in \Delta_f$ implies $c\phi \in \Delta_f$, and Δ_f is a subspace. That $\int_{\mathsf{X}} f(x) dP(x) : \Delta_f \ni \psi \mapsto \int_{\mathsf{X}} f(x) dP(x) \psi$ is linear is a consequence of the definition of $\int_{\mathsf{X}} f(x) dP(x) \psi$ and of the linearity of the integral of a bounded map in a PVM.

Now we show Δ_f is dense in H. Given f as in the statement, let:

$$E_n := \{x \in \mathsf{X} \mid n - 1 \leq |f(x)| < n\} \,, \quad \text{for any } n \in \mathbb{N}, n \geq 1.$$

Note $E_n \cap E_m = \varnothing$ if $n \neq m$ and $\cup_n E_n = \mathsf{X}$. By Definition 8.41 the closed subspaces $\mathsf{H}_n := P(E_n)\mathsf{H}$ are mutually orthogonal, and by propriety (d) of the same definition finite combinations over the H_n form a dense space inside H. We claim Δ_f contains this subspace. By monotone convergence, if $\psi \in \mathsf{H}$:

$$\int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x) = \lim_{k\to+\infty} \sum_{n=1}^{k} \int_{\mathsf{X}} |\chi_{E_n}(x) f(x)|^2 d\mu_\psi(x) \leq +\infty \,. \tag{9.8}$$

The integral inside the sum can be written as follows, using Theorem 8.52(b):

$$\left(\int_{\mathsf{X}} \chi_{E_n}(x) f(x) dP(x) \psi \middle| \int_{\mathsf{X}} \chi_{E_n}(x) f(x) dP(x) \psi \right) .$$

But since $x \mapsto \chi_{E_n}(x) f(x)$ is bounded and $\chi_{E_n} = \chi_{E_n} \cdot \chi_{E_n}$, using (iii) in Theorem 8.52(a) gives

$$\int_{\mathsf{X}} \chi_{E_n}(x) f(x) dP(x) \psi = \int_{\mathsf{X}} \chi_{E_n}(x) f(x) dP(x) \int_{\mathsf{X}} \chi_{E_n}(x) dP(x) \psi$$

$$= \int_{\mathsf{X}} \chi_{E_n}(x) f(x) dP(x) \circ P(E_n) \psi \,.$$

If $\psi \in \mathsf{H}_n$, then, as projectors $P(E_m)$ are orthogonal,

$$\int_{\mathsf{X}} \chi_{E_k}(x) f(x) dP(x) \psi = 0, \quad \text{for } k \neq n \,.$$

Under the assumptions on ψ, therefore, the series of (9.8) becomes

$$\int_{\mathsf{X}} |f(x)|^2 d\mu_\psi(x) = \int_{\mathsf{X}} |\chi_{E_n}(x) f(x)|^2 d\mu_\psi(x) \leq \int_{\mathsf{X}} n^2 d\mu_\psi(x) = n^2 ||\psi||^2 < +\infty \,.$$

We conclude $\mathsf{H}_n \subset \Delta_f$, for any $n = 1, 2, \ldots$. But Δ_f is a subspace so it contains also the dense space of finite combinations of the H_n. Hence Δ_f is dense.
(c) With $f : \mathsf{X} \to \mathbb{C}$ let

$$F_k := \{x \in \mathsf{X} \mid |f(x)| < k\} \,. \tag{9.9}$$

Suppose $f\restriction_{supp(P)}$ is bounded. Define *bounded* measurable maps $f_n := \chi_{supp(P)} \cdot f + g_n$ where $g_n = \chi_{F_n} \cdot \chi_{\mathsf{X} \setminus supp(P)} \cdot f$. Since $supp(\mu_\psi) \subset supp(P)$ by Theorem 8.50(v), for any $\psi \in \mathsf{H}$ we have $f \in L^2(\mathsf{X}, \mu_\psi)$, hence $\Delta_f = \mathsf{H}$ because μ_f is finite, and:

$$\int_{\mathsf{X}} |f_n(x) - f(x)|^2 d\mu_\psi(x) = \int_{supp(P)} |f(x) - f(x)|^2 d\mu_\psi(x) = 0 \,.$$

Consequently $f_n \to f$ in $L^2(\mathsf{X}, \mu_\psi)$ for any $\psi \in \mathsf{H}$, so:

$$\int_{\mathsf{X}} f(x) dP(x) \psi := \lim_{n\to+\infty} \int_{\mathsf{X}} f_n(x) dP(x) \psi = \lim_{n\to+\infty} \int_{\mathsf{X}} \chi_{supp(P)} f_n(x) dP(x) \psi$$

$$= \lim_{n\to+\infty} \int_{\mathsf{X}} \chi_{supp(P)} f(x) dP(x) \psi = \int_{\mathsf{X}} \chi_{supp(P)} f(x) dP(x) \psi =: \int_{supp(P)} f(x) dP(x) \psi \,,$$

where the last integral is meant as in Definition 8.47(c), so the operator $\int_{supp(P)} f(x) dP(x)$ belongs in $\mathfrak{B}(\mathsf{H})$. $\square$

Now a result that deals, in particular, with composites of integrals of spectral measures, and characterises in a *very precise* way the corresponding domains.

Given a pair of operators A, B, we remind that $A \subset B$ means that B extends A (Definition 5.3).

Theorem 9.4. *Let X be a second-countable space, $\mathscr{B}(\mathsf{X})$ the Borel σ-algebra of X, H a Hilbert space and $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ a PVM. For any measurable $f : \mathsf{X} \to \mathbb{C}$, in the same notation of Theorem 9.3:*

(a) $\int_{\mathsf{X}} f(x)dP(x) : \Delta_f \to \mathsf{H}$ *is a closed operator.*
(b) $\int_{\mathsf{X}} f(x)dP(x)$ *is self-adjoint if f is real, and more generally:*

$$\left(\int_{\mathsf{X}} f(x)dP(x)\right)^* = \int_{\mathsf{X}} \overline{f(x)}dP(x) \quad \text{and} \quad \Delta_{\overline{f}} = \Delta_f \,. \tag{9.10}$$

(c) *If $f : \mathsf{X} \to \mathbb{C}$, $g : \mathsf{X} \to \mathbb{C}$ are measurable, D the standard domain (Definition 5.1) and $f \cdot g$ the pointwise product,*

$$\int_{\mathsf{X}} f(x)dP(x) + \int_{\mathsf{X}} g(x)dP(x) \subset \int_{\mathsf{X}} (f+g)(x)dP(x) \tag{9.11}$$

$$D\left(\int_{\mathsf{X}} f(x)dP(x) + \int_{\mathsf{X}} g(x)dP(x)\right) = \Delta_f \cap \Delta_g \tag{9.12}$$

with equality in (9.11) $\Leftrightarrow$ $\Delta_{f+g} = \Delta_f \cap \Delta_g$.

$$\int_{\mathsf{X}} f(x)dP(x) \int_{\mathsf{X}} g(x)dP(x) \subset \int_{\mathsf{X}} (f \cdot g)(x)dP(x) \tag{9.13}$$

$$D\left(\int_{\mathsf{X}} f(x)dP(x) \int_{\mathsf{X}} g(x)dP(x)\right) = \Delta_{f\cdot g} \cap \Delta_g \tag{9.14}$$

with $=$ in (9.13) $\Leftrightarrow$ $\Delta_{f\cdot g} \subset \Delta_g$. *In particular:*

$$\int_{\mathsf{X}} \overline{f(x)}dP(x) \int_{\mathsf{X}} f(x)dP(x) = \int_{\mathsf{X}} |f(x)|^2 dP(x) \tag{9.15}$$

$$D\left(\int_{\mathsf{X}} \overline{f(x)}dP(x) \int_{\mathsf{X}} f(x)dP(x)\right) = \Delta_{|f|^2} \,. \tag{9.16}$$

Moreover

$$\left(\int_{\mathsf{X}} f(x)dP(x) \int_{\mathsf{X}} g(x)dP(x)\right)\restriction_{\Delta_f \cap \Delta_g \cap \Delta_{f\cdot g}} = \left(\int_{\mathsf{X}} g(x)dP(x) \int_{\mathsf{X}} f(x)dP(x)\right)\restriction_{\Delta_f \cap \Delta_g \cap \Delta_{f\cdot g}} . \tag{9.17}$$

Eventually, if f is bounded on the Borel set $E \subset \mathsf{X}$, then $\Delta_{\chi_E \cdot f} = \mathsf{H}$ and

$$\int_{\mathsf{X}} \chi_E(x)dP(x) \int_{\mathsf{X}} f(x)dP(x) \subset \int_{\mathsf{X}} f(x)dP(x) \int_{\mathsf{X}} \chi_E(x)dP(x) = \int_{\mathsf{X}} (\chi_E \cdot f)(x)dP(x) \in \mathfrak{B}(\mathsf{H}). \tag{9.18}$$

(d) *If* $\mathsf{X} = \mathbb{R}$, $p : \mathbb{R} \to \mathbb{C}$ *is a polynomial of degree* $m \in \mathbb{N}$, *and* $T := \int_{\mathsf{X}} x dP(x)$, *then*

$$p(T) = \int_{\mathbb{R}} p(x) dP(x) \quad \text{and} \quad D(p(T)) = D(T^m) = \Delta_p \,. \tag{9.19}$$

(e) *Defining* $\mu_{\phi,\psi}$ *as in Theorem 8.50(c),* $\int_{\mathsf{X}} f(x) dP(x)$ *is the unique operator in* H *with domain* Δ_f *such that, for any* $\psi \in \Delta_f$, $\phi \in \mathsf{H}$*:*

$$\left(\phi \middle| \int_{\mathsf{X}} f(x) dP(x) \psi \right) = \int_{\mathsf{X}} f(x) d\mu_{\phi,\psi}(x) \,. \tag{9.20}$$

(f) *For any* $\psi \in \Delta_f$*:*

$$\left|\left| \int_{\mathsf{X}} f(x) dP(x) \psi \right|\right|^2 = \int_{\mathsf{X}} |f(x)|^2 d\mu_{\psi}(x) \,. \tag{9.21}$$

(g) *Every operator* $\int_{\mathsf{X}} f(x) dP(x)$ *is positive when* f *is positive, i.e.:*

$$\left(\psi \middle| \int_{\mathsf{X}} f(x) dP(x) \psi \right) \geq 0 \ \text{ for any } \psi \in \Delta_f, \text{ if } f(x) \geq 0,\ x \in \mathsf{X}. \tag{9.22}$$

(h) *If* X' *is second countable and* $\phi : \mathsf{X} \to \mathsf{X}'$ *is measurable (i.e.* $\phi^{-1}(E') \in \mathscr{B}(\mathsf{X})$ *for* $E' \in \mathscr{B}(\mathsf{X}')$*), then*

$$\mathscr{B}(\mathsf{X}') \ni E' \mapsto P'(E') := P(\phi^{-1}(E'))$$

is a PVM on X', *and for any measurable* $f : \mathsf{X}' \to \mathbb{C}$*:*

$$\int_{\mathsf{X}'} f(x') dP'(x') = \int_{\mathsf{X}} (f \circ \phi)(x) dP(x) \quad \text{and} \quad \Delta'_f = \Delta_{f \circ \phi} \,, \tag{9.23}$$

where Δ'_f *is the domain of the operator on the left.*

Proof. First of all notice that part (f) follows, by continuity, from the similar property of bounded maps, seen in Theorem 8.52(b), when we use our definition of integral of unbounded maps. Likewise, Theorem 8.50(b) implies (g). In fact if $f \geq 0$, $\psi \in \Delta_f$, the sequence of $\chi_{F_n} \cdot f_n \geq 0$ tends to f in $L^2(\mathsf{X}, \mu_\psi)$, so

$$0 \leq \left(\psi \middle| \int_{\mathsf{X}} (\chi_{F_n} \cdot f(x)) dP(x) \psi \right) \to \left(\psi \middle| \int_{\mathsf{X}} f(x) dP(x) \psi \right), \quad \text{as } n \to +\infty,$$

and $(\psi | \int_{\mathsf{X}} f(x) dP(x) \psi) \geq 0$. Let us see to the rest.
(a) We claim $T := \int_{\mathsf{X}} f(x) dP(x)$, defined on Δ_f, is closed. Notice, first, that the bounded operators (F_k as in (9.9))

$$T_k := \int_{\mathsf{X}} \chi_{F_k}(x) f(x) dP(x) \,, \tag{9.24}$$

are such, for $\psi \in \Delta_f$, that: (1) $TP_{F_k}\psi = P_{F_k}T\psi = T_k\psi$ and (2) $T_k\psi \to T\psi$, $k \to +\infty$. The proof of (1) is similar to Theorem 9.3(c), whilst (2) follows from the argument preceding Lemma 9.2. So let $\{\psi_n\}_{n\in\mathbb{N}} \subset \Delta_f$ be such that $\psi_n \to \psi \in \mathsf{H}$ and $T\psi_n \to \phi$, $n \to +\infty$. We claim $\psi \in \Delta_f$ and $T\psi = \phi$, implying T's closure. By (1) and because $P_{F_k} \to I$ strongly as $k \to +\infty$:

$$T_k\psi = \lim_{n\to+\infty} T_k\psi_n = \lim_{n\to+\infty} P_{F_k}T\psi_n = P_{F_k}\phi \to \phi \quad \text{in H as } k \to +\infty.$$

Define $\phi_k := T_k\psi$; then

$$\int_{\mathsf{X}} \chi_{F_k}(x)f(x)dP(x)\psi = \phi_k \to \phi \quad \text{in H as } k \to +\infty. \tag{9.25}$$

By Theorem 8.52(b):

$$\int_{\mathsf{X}} \chi_{F_k}(x)|f(x)|^2 d\mu_\psi(x) = ||\phi_k||^2 \to ||\phi||^2 < +\infty \quad \text{as } n \to +\infty.$$

Monotone convergence ensures $f \in L^2(\mathsf{X},\mu_\psi)$, i.e. $\psi \in \Delta_f$. Rewriting (9.25) as $T_k\psi = \phi_k$, and taking the limit for $k \to +\infty$ using (2), gives $T\psi = \phi$, as required.
(b) $\Delta_f = \Delta_{\overline{f}}$ is an obvious consequence of the definition of Δ_f and $|f| = |\overline{f}|$. We will show $\int_{\mathsf{X}} \overline{f(x)}dP(x) \subset (\int_{\mathsf{X}} f(x)dP(x))^*$. If $\psi \in \Delta_{\overline{f}}$, $\phi \in \Delta_f$ and $f_n \to f$ in $L^2(\mathsf{X},\mu_\phi)$ so $\overline{f_n} \to \overline{f}$ in $L^2(\mathsf{X},\mu_\psi)$, where f_n are bounded, we have:

$$\left(\psi \middle| \int_{\mathsf{X}} f(x)dP(x)\phi\right) = \lim_{n\to+\infty}\left(\psi \middle| \int_{\mathsf{X}} f_n(x)dP(x)\phi\right)$$
$$= \lim_{n\to+\infty}\left(\int_{\mathsf{X}} \overline{f_n(x)}dP(x)\psi \middle| \phi\right) = \left(\int_{\mathsf{X}} \overline{f(x)}dP(x)\psi \middle| \phi\right)$$

where we used the definition of integral of f and $\overline{f}$ in P, plus property (iv) in Theorem 8.52(a). This means $\int_{\mathsf{X}} \overline{f(x)}dP(x) \subset (\int_{\mathsf{X}} f(x)dP(x))^*$. We will prove $\int_{\mathsf{X}} \overline{f(x)}dP(x) \supset (\int_{\mathsf{X}} f(x)dP(x))^*$ by showing $D\left((\int_{\mathsf{X}} f(x)dP(x))^*\right) \subset \Delta_{\overline{f}}$. Let $T := \int_{\mathsf{X}} \overline{f(x)}dP(x)$ and take the bounded operators T_k of (9.24). Fix $\psi \in D(T^*)$. Then there is $h \in \mathsf{H}$ such that, for any $\phi \in \Delta_f$, $(\psi|T\phi) = (h|\phi)$. Choosing $\phi = T_k^*\psi$ we obtain $(\psi|T_kT_k^*\psi) = (h|T_k^*\psi)$, where we used $TT_k^* = T_kT_k^*$ because $T_k^* = P_{F_k}T_k^*$ from $TP_{F_k} = T_k$. Therefore $||T_k^*\psi||^2 = (h|T_k^*\psi)$, so $||T_k^*\psi||^2 \leq ||T_k^*\psi||\,||h||$, i.e. $||T_k^*\psi|| \leq ||h||$. Consequently, by Theorem 8.52(b):

$$\int_{\mathsf{X}} \chi_{F_k}(x)|\overline{f(x)}|^2 d\mu_\psi(x) \leq ||h||^2 \quad \text{for any } k \in \mathbb{N},$$

which implies $\psi \in \Delta_{\overline{f}}$ by monotone convergence. So we have $D(T^*) \subset \Delta_{\overline{f}}$.
(c) Formulas (9.11), (9.12) and the ensuing remark are trivial consequences of the given definitions and of standard domains. Let us prove (9.13), (9.14). Assume f is bounded so that $\Delta_{f\cdot g} \subset \Delta_g$, and $\psi \in \Delta_g$. Take $\{g_n\}_{n\in\mathbb{N}}$ a sequence of bounded measurable maps converging to g in $L^2(\mathsf{X},d\mu_g)$. Then $f\cdot g_n \to f\cdot g$ in $L^2(\mathsf{X},d\mu_g)$, and because the integrals of f, g_n, $f\cdot g_n$ are as in Definition 8.47 plus (iii) in Theorem 8.52(a), we immediately have, for $n \to +\infty$:

$$\int_{\mathsf{X}} f(x)dP(x)\int_{\mathsf{X}} g_n(x)dP(x)\psi = \int_{\mathsf{X}}(f\cdot g_n)(x)dP(x)\psi \to \int_{\mathsf{X}}(f\cdot g)(x)dP(x)\psi.$$

As $\int_X f dP$ is continuous, we will prove that f bounded and $\psi \in \Delta_g$ imply

$$\int_X f(x)dP(x) \int_X g(x)dP(x)\psi = \int_X (f \cdot g)(x)dP(x)\psi . \tag{9.26}$$

Let now $\phi := \int_X g dP\psi$, by (f) the identity shows

$$\int_X |f(x)|^2 d\mu_\phi(x) = \int_X |(f \cdot g)(x)|^2 d\mu_\psi(x) \quad \text{if } f \text{ is bounded and } \psi \in \Delta_g . \tag{9.27}$$

Take now f just measurable, possibly unbounded. As (9.27) holds for bounded maps, it holds for unbounded ones too. Since

$$D\left(\int_X f(x)dP(x) \int_X g(x)dP(x)\right)$$

contains all $\psi \in \Delta_g$ such that $\phi \in \Delta_f$, which happens by (9.27) precisely when $\psi \in \Delta_{f \cdot g}$, we conclude:

$$D\left(\int_X f(x)dP(x) \int_X g dP(x)\right) = \Delta_g \cap \Delta_{f \cdot g} .$$

If now $\phi \in \Delta_g \cap \Delta_{f \cdot g}$, if $\psi = \int_X g(x)dP(x)\phi$ and $f_n := \chi_{F_n} \cdot f$ (F_n as previously), then $f_n \to f$ in $L^2(X, \mu_\psi)$, $f_n \cdot g \to f \cdot g$ in $L^2(X, \mu_\phi)$ and (9.26), (f) (f_n replacing f) imply:

$$\int_X f(x)\, dP(x) \int_X g(x)\, dP(x)\phi = \int_X f(x)\, dP(x)\psi = \lim_{n\to+\infty} \int_X f_n(x)\, dP(x)\psi =$$

$$= \lim_{n\to+\infty} \int_X (f_n \cdot g)(x)\, dP(x)\phi = \int_X (f \cdot g)(x)\, dP(x)\phi .$$

This ends the proof of (9.13) and (9.14).

Inclusion (9.13) plus the equality in case $\Delta_g \supset \Delta_{f \cdot g}$ easily imply (9.17) and (9.18). Concerning (9.17), we have $\Delta_f \supset \Delta_{\overline{f} \cdot f} = \Delta_{|f|^2}$ for the following reason: as μ_ψ is finite, if $\psi \in \Delta_{|f(x)|^2}$

$$\int_X |f(x)|^2 d\mu_\psi(x) = \int_X |f(x)|^2 \cdot 1 d\mu_\psi(x) \le \sqrt{\int_X |f(x)|^4 d\mu_\psi(x)} \sqrt{\int_X 1^2 d\mu_\psi(x)} < +\infty .$$

(d) By (9.13) and (9.11) we have

$$p(T) \subset \int_X p(x)\, dP(x) .$$

Hence the claim is true when $D(p(T)) = \Delta_p$. Let us prove this, starting from showing by induction

$$D(T^n) = \Delta_{x^n} \quad \text{for } n \in \mathbb{N}. \tag{9.28}$$

When $n = 0, 1$, the identity is true: $D(T^0) = \Delta_1 = H$, $D(T) = \Delta_x$. Assume it true for a given n and let us prove it for $n+1$: $D(T^{n+1}) = \Delta_{x^{n+1}}$. We have to show

$D(TT^n) = \Delta_{x\circ x^n}$. Using the special property stated after (9.14), we know that is equivalent to $\Delta_{x\circ x^n} \subset \Delta_{x^n}$. The latter holds because μ_ψ is positive and finite, and $|x^{n+1}| > |x^n|$ outside a compact set $J \subset \mathbb{R}$, so $\psi \in \Delta_{x^{n+1}}$ implies

$$\int_{\mathbb{R}} |x|^{2n} d\mu_\psi(x) = \int_{\mathbb{R}\setminus J} |x|^{2n} d\mu_\psi(x) + \int_J |x|^{2n} d\mu_\psi(x)$$
$$\leq \int_{\mathbb{R}\setminus J} |x|^{2n+2} d\mu_\psi(x) + \sup_J |x|^{2n} \int_J 1 d\mu_\psi(x)$$
$$\leq \int_{\mathbb{R}} |x|^{2n+2} d\mu_\psi(x) + \sup_J |x|^{2n} \int_{\mathbb{R}} 1 d\mu_\psi(x) < +\infty .$$

We remark for later that we have also obtained

$$D(T^{n+1}) = \Delta_{x^{n+1}} \subset \Delta_{x^n} = D(T^n) .$$

To finish the proof of $D(p(T)) = \Delta_p$ we compute separately the two sides. Take the leading coefficient $a_m \neq 0$ in p. As $D(T^{n+1}) \subset D(T^n)$, and in general $D(A+B) = D(A) \cap D(B)$, we have

$$D(p(T)) = D(T^m) . \tag{9.29}$$

Let us compute Δ_p. Since $\Delta_{x^{n+1}} \subset \Delta_{x^n}$, we find $\Delta_{x^m} \subset \Delta_p$. Let us prove the opposite inclusion. From $|p(x)|/|x|^m \to |a_m|$, $|x| \to +\infty$, follows $|p(x)|/|x|^m \leq |a_m| + \varepsilon > 0$ for any $\varepsilon > 0$, provided x does not belong to a too big compact set $J_\varepsilon \subset \mathbb{R}$. Hence if $\psi \in \Delta_p$:

$$\int_{\mathbb{R}} |x|^{2m} d\mu_\psi$$
$$\leq \int_{\mathbb{R}\setminus J_\varepsilon} |x|^{2m} d\mu_\psi + \int_{J_\varepsilon} |x|^{2m} d\mu_\psi \leq \int_{\mathbb{R}\setminus J_\varepsilon} \frac{|p(x)|^2}{(|a_m|+\varepsilon)^2} d\mu_\psi + \sup_{J_\varepsilon} |x|^{2m} \int_{\mathbb{R}} d\mu_\psi$$
$$\leq \frac{1}{(|a_m|+\varepsilon)^2} \int_{\mathbb{R}} |p(x)|^2 d\mu_\psi + \sup_{J_\varepsilon} |x|^{2m} \int_{\mathbb{R}} d\mu_\psi < +\infty ,$$

and so $\psi \in \Delta_{x^m}$. Thus $\Delta_p \subset \Delta_{x^m}$ and $\Delta_p = \Delta_{x^m}$. From (9.28) and (9.29) we have $\Delta_p = \Delta_{x^m} = D(T^m) = D(p(T))$, ending this part.
(e) Define the usual bounded maps $f_n := \chi_{F_n} \cdot f$ tending to f in $L^2(\mathsf{X}, \mu_\psi)$. By definition of integral, and by (iv) in Theorem 8.50(c):

$$\left(\phi \left| \int_{\mathsf{X}} f(x) dP(x)\ \psi \right.\right) = \lim_{n\to+\infty} \left(\phi \left| \int_{\mathsf{X}} f_n(x) dP(x)\ \psi \right.\right) = \lim_{n\to+\infty} \int_{\mathsf{X}} f_n(x) d\mu_{\phi,\psi}(x) .$$

Now we show

$$\lim_{n\to+\infty} \int_{\mathsf{X}} (f_n(x) - f(x)) d\mu_{\phi,\psi}(x) = 0 .$$

By Lemma 9.2 (same notations):

$$\left| \int_{\mathsf{X}} (f_n(x) - f(x)) d\mu_{\phi,\psi}(x) \right| = \left| \int_{\mathsf{X}} (f_n(x) - f(x)) h(x) d|\mu_{\phi,\psi}(x)| \right|$$
$$\leq \int_{\mathsf{X}} |f_n(x) - f(x)| d|\mu_{\phi,\psi}(x)| \leq ||\phi|| \sqrt{\int_{\mathsf{X}} |f_n(x) - f(x)|^2 d\mu_\psi(x)} \to 0$$

as $n \to +\infty$, by definition of $\int_X f(x)dP(x)\psi$. Uniqueness now follows. If $T : \Delta_f \to \mathsf{H}$ satisfies the same property of $\int_X f(x)dP(x)$, then $T' := T - \int_X f(x)dP(x)$ solves $(\phi|T'\psi) = 0$ for any $\phi \in \mathsf{H}$, irrespective of $\psi \in \Delta_f$. Choosing $\phi = T'\psi$ gives $||T'\psi|| = 0$ and so $T = \int_X f(x)dP(x)$.

Parts (f) and (g) were seen to at the beginning of this proof.

(h) We outline only the proof as it is elementary if tedious. P' is a PVM by direct inspection. If f is simple, assertion (9.23) is trivial. Using Definition 8.47 one generalises (9.23) to bounded measurable maps, so (9.23) extends by virtue of the definition of integral for unbounded f. □

Corollary 9.5. *Under the assumptions of Theorem 9.3, if $f : \mathsf{X} \to \mathbb{C}$ is measurable the following are equivalent.*

(a) $\Delta_f = \mathsf{H}$ *(i.e. $D(\int_X f(x)dP(x)) = \mathsf{H}$).*
(b) *f is essentially bounded with respect to the PVM P (Definition 8.42).*
(c) $\int_X f(x)dP(x) \in \mathfrak{B}(\mathsf{H})$.
Under either of (a), (b), (c), the estimate

$$||f||_\infty^{(P)} \leq \left|\left|\int_X f dP\right|\right| \tag{9.30}$$

holds. Then we can redefine f on a zero-measure set for P obtaining $||f||_\infty < +\infty$, and without altering $\int_X f dP$: that latter can be computed with Definition 8.47 and yields the same result.

Proof. Let us prove the equivalence of (a), (b), (c), and that any one implies (9.30). Properties (a) and (c) are equivalent by the closed graph theorem (2.95), for $\int_X f dP$ is closed by Theorem 9.4(a). Let us prove (b) implies (c). Define $F_n := \{x \in \mathsf{X} \,|\, |f(x)| < n\}$. If $f_n := \chi_{F_n} \cdot f$, then $f_n \to f$ pointwise as $n \to +\infty$. If f is essentially bounded, for n large enough $f - f_n$ is not 0 on a set $G_n \in \mathscr{B}(\mathsf{X})$ with $P(G_n) = 0$. Hence

$$\int_X f + (-f_n)dP = \int_X \chi_{G_n}(f - f_n)dP = \int_X f - f_n dP \int_X \chi_{G_n} dP$$
$$= \left(\int_X f - f_n dP\right) P(G_n) = 0\,.$$

By Theorem 9.4(c), $\int_X f(x)dP(x) = -\int_X (-f_n(x))dP(x)$ belongs to $\mathfrak{B}(\mathsf{H})$ ($-f_n$ being bounded by Theorem 9.3(c)). Now we prove (c) implies (b). To show (9.30) consider $f : \mathsf{X} \to \mathbb{C}$ measurable, with no boundedness assumption, and assume (c) (i.e. (a)). Take the usual sequence $f_n \in M_b(\mathsf{X})$. By (8.49):

$$||f_n||_\infty^{(P)} = \left|\left|\int_X f\chi_n dP\right|\right| = \left|\left|\int_X f dP \int_X \chi_{F_n} dP\right|\right| \leq \left|\left|\int_X f dP\right|\right| \left|\left|\int_X \chi_{F_n} dP\right|\right|$$
$$\leq \left|\left|\int_X f dP\right|\right| =: M < +\infty\,.$$

By construction $\{x \in \mathsf{X} \,|\, |f(x)| \geq M\} \subset \cup_{n\in\mathbb{N}} \{x \in \mathsf{X} \,|\, |f_n(x)| \geq M\}$. Monotonicity and sub-additivity imply

$$(\psi|P(\{x \in \mathsf{X} \,|\, |f(x)| \geq M\}\psi) \leq \sum_{n\in\mathbb{N}} (\psi|P(\{x \in \mathsf{X} \,|\, |f_n(x)| \geq M\})\psi) = 0\,,$$

which means $||f||_\infty^{(P)} \leq M < +\infty$, as required. The above assertion is valid even if we redefine f null on the zero-measure set $|f(x)| > N|$ for some finite $N > ||f||_\infty^{(P)}$. Then the last statement of the thesis immediately arises. □

The next definition is based on (9.7), and extends the integral in a PVM. We can also make use of Theorem 9.4(e) to obtain a more elegant, equivalent definition.

Definition 9.6. *Let* X *be a second-countable space,* H *a Hilbert space and* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ *a PVM.*

(a) *If* $f : \mathsf{X} \to \mathbb{C}$ *is measurable with* Δ_f *as in (9.6), the operator*

$$\int_{\mathsf{X}} f(x)dP(x) : \Delta_f \to \mathsf{H}$$

of (9.7) is called **integral of** f *with respect to the projector-valued measure P. Equivalently,* $\int_{\mathsf{X}} f(x)dP(x)$ *is the unique operator* $S : \Delta_f \to \mathsf{H}$ *such that*

$$(\phi\,|S\psi) = \int_X f(x)d\mu_{\phi,\psi}(x)\,, \quad \textit{for any } \phi \in H,\ \psi \in \Delta_f,$$

where the complex spectral measure $\mu_{\phi,\psi}$ *is defined in Theorem 8.50(c).*

(b) *For every* $E \subset \mathscr{B}(\mathsf{X})$*,* $f : \mathsf{X} \to \mathbb{C}$ *and* $g : E \to \mathbb{C}$ *measurable, the integrals*

$$\int_E f(x)\,dP(x) := \int_{\mathsf{X}} \chi_E(x) f(x)\,dP(x) \textit{ and } \int_E g(x)\,dP(x) := \int_{\mathsf{X}} g_0(x)\,dP(x)\,,$$

with $g_0(x) := g(x)$ *if* $x \in E$ *or* $g_0(x) := 0$ *if* $x \notin E$*, are respectively called* **integral of** f **on** E *and* **integral of** g **on** E *(in the PVM P).*

Remark 9.7. (1) By Theorem 9.3(c), the above extends Definition 8.47 for bounded maps.

(2) For any $f : \mathsf{X} \to \mathbb{C}$ measurable,

$$\int_{\mathsf{X}} f(x)dP(x) = \int_{supp(P)} f(x)dP(x) \quad \text{and so} \quad \int_{\mathsf{X}\setminus supp(P)} f(x)dP(x) = 0\,. \quad (9.31)$$

The proof is straighforward: $\chi_{supp(P)}$ is bounded, by definition its integral is ($\chi_{supp(P)}$ is simple):

$$\int_{supp(P)} 1dP := \int_{\mathsf{X}} \chi_{supp(P)} dP = P(supp(P)) = I\,,$$

where Proposition 8.43(c) was used in the last equality. Now the second identity in (9.18) gives

$$\int_{\mathsf{X}} f(x)dP(x) = \int_{\mathsf{X}} f(x)dP(x) \int_{\mathsf{X}} \chi_{supp(P)}(x)dP(x) = \int_{supp(P)} \chi_{supp(P)}(x) f(x)dP(x)$$
$$=: \int_{supp(P)} f(x)dP(x)\,.$$

The rest of (9.31) follows, similarly, by using $P(\mathsf{X} \setminus supp(P)) = 0$. ■

Examples 9.8. (1) Consider the spectral measure:

$$P : \mathscr{B}(N) \ni E \mapsto P_E = \sum_{z\in E} z(z|\ \ )$$

of Example 8.49(2) on a basis N of a separable Hilbert space H, and equip N with the second-countable topology of power sets. We are interested in writing an explicit formula for the integral of unbounded maps $f : N \to \mathbb{C}$ relying on definition (9.4). In the case under exam $\int_N |f(z)|^2 d\mu_\psi(z) < +\infty$ becomes $\sum_{z\in N} |f(z)|^2 |(z|\psi)|^2 < +\infty$. We aim to show

$$\int_N f(z)dP(z) = \text{s-}\sum_{z\in N} f(z)z(z|\ \ )$$

for f unbounded. This formula was proved in Example 8.49(2) for f bounded. Suppose $\{N_n\}_{n\in\mathbb{N}}$ are finite subsets in N, $N_{n+1} \supset N_n$ and $\cup_{n\in\mathbb{N}} N_n = N$. The sequence of bounded maps $f_n := \chi_{N_n} \cdot f$ converges in $L^2(N, \mu_\psi)$, for any $\psi \in \mathsf{H}$ such that $\sum_z |f(z)|^2 |(z|\psi)|^2 < +\infty$, simply by Lebesgue's dominated convergence. By definition (9.4) we have, if $\sum_{z\in N} |f(z)|^2 |(z|\psi)|^2 < +\infty$:

$$\int_N f(z)dP(z)\psi := \lim_{n\to+\infty} \int_N f_n(z)dP(z)\psi\,. \tag{9.32}$$

But f_n is bounded, so Example 8.49(2) guarantees

$$\int_N f_n(z)dP(z)\psi = \text{s-}\sum_{z\in N} f_n(z)(z|\psi) = \sum_{z\in N_n} f(z)z(z|\psi)\,,$$

where the sum is finite for N_n contains a finite number of points. Definition (9.32) reduces to

$$\int_N f(z)dP(z)\psi = \lim_{n\to+\infty} \sum_{z\in N_n} f(z)z(z|\psi)\,,$$

i.e.

$$\int_N f(z)dP(z) = \text{s-}\sum_{z\in N} f(z)z(z|\ \ )\,. \tag{9.33}$$

Later we will see a concrete example of an unbounded self-adjoint operator built with this type of spectral measure.

(2) Consider the spectral measure of Example 8.49(1). Take a Hilbert space $\mathsf{H} = L^2(\mathsf{X}, \mu)$, X second countable and μ positive, σ-additive on the Borel σ-algebra of X. The spectral measure on H we wish to consider is the following. For any $\psi \in L^2(\mathsf{X}, \mu)$, $E \in \mathscr{B}(\mathsf{X})$, let

$$(P(E)\psi)(x) := \chi_E(x)\psi(x)\,, \quad \text{for almost every } x \in \mathsf{X}\,. \tag{9.34}$$

With $\psi \in \mathsf{H}$, the measure μ_ψ is

$$\mu_\psi(E) = (\psi|P(E)\psi) = \int_E |\psi(x)|^2 d\mu(x)\,, \quad \text{for any } E \in \mathscr{B}(\mathsf{X})\,.$$

Consequently if $g : \mathsf{X} \to \mathbb{C}$ is measurable:

$$\int_{\mathsf{X}} g(x) d\mu_\psi(x) = \int_{\mathsf{X}} g(x)|\psi(x)|^2 d\mu(x) .$$

In Example 8.49(1) we saw that if $f : \mathsf{X} \to \mathbb{C}$ is measurable and bounded:

$$\left(\int_{\mathsf{X}} f(y)\, dP(y)\psi\right)(x) = f(x)\psi(x) \quad \text{for every } \psi \in L^2(\mathsf{X},\mu) \text{ and almost every } x \in \mathsf{X}. \tag{9.35}$$

This is valid for unbounded measurable maps, too, as long as $\psi \in \Delta_f$. If $f : \mathsf{X} \to \mathbb{C}$ is unbounded and measurable take a sequence of bounded measurable maps f_n such that $f_n \to f$, $n \to +\infty$, in $L^2(\mathsf{X}, \mu_\psi)$, with $\psi \in \Delta_f$. In other words, by the above expression for μ_ψ we take

$$\int_{\mathsf{X}} |f_n(x) - f(x)|^2 |\psi(x)|^2 d\mu(x) \to 0 \quad \text{as } n \to +\infty.$$

By (9.35):

$$||f \cdot \psi - f_n \cdot \psi||^2_{\mathsf{H}} = \int_{\mathsf{X}} |f(x) - f_n(x)|^2 |\psi(x)|^2 d\mu(x) \to 0 \quad \text{as } n \to +\infty.$$

Therefore the definition of integral in P implies that for any $\psi \in \Delta_f$, with $f : \mathsf{X} \to \mathbb{C}$ measurable and possibly unbounded:

$$\left(\int_{\mathsf{X}} f(x) dP(x)\psi\right)(y) = f(y)\psi(y) \quad \text{for almost every } y \in \mathsf{X}. \tag{9.36}$$

■

9.1.2 Von Neumann algebra of a bounded normal operator

Corollary 9.5 has an important consequence for the von Neumann algebra (see Definition 3.47 and the ensuing argument) generated by a bounded normal operator and the adjoint. (The result may be somehow generalised by looking at unbounded PVMs, as proven in Exercises 9.6–9.7.)

Theorem 9.9 (Von Neumann algebra generated by a bounded normal operator and its adjoint). *Take a normal operator $T \in \mathfrak{B}(\mathsf{H})$ with H separable. The subspace in $\mathfrak{B}(\mathsf{H})$ that commutes with every operator commuting with T and T^* (the von Neumann algebra generated by T, T^*) consists precisely of the operators $f(T, T^*)$ of Theorem 8.39, for $f \in M_b(\sigma(T))$.*

Proof. Indicate by $\mathfrak{M}$ the von Neumann algebra generated by T, T^*. We know that any $f(T, T^*)$, with $f : \sigma(T) \to \mathbb{C}$ measurable and bounded, belongs to $\mathfrak{M}$ by (iii) in Theorem 8.39(b) (in the sequel we will need Theorem 8.52, the spectral Theorem 8.54 and Theorem 8.56). Let us show the converse. Clearly $\mathfrak{M}$ coincides with the von Neumann algebra generated by the *-algebra with unit of complex polynomials in T, T^* (restricted to $\sigma(T)$, from now on always assumed). By the double commutant theorem (3.46) $\mathfrak{M}$ is the strong closure of complex polynomials in T, T^*. That is to say, if $B \in \mathfrak{M}$ there is a sequence of bounded measurable f_n (better: restrictions of polynomials to $\sigma(T)$) such that $f_n(T,T^*)x \to Bx$, $n \to +\infty$, for any $x \in \mathsf{H}$. We claim $B = f(T,T^*)$ for some bounded measurable f defined on $\sigma(T)$. As $g(T,T^*) = \int_{\sigma(T)} g dP^{(T)}$ by Theorem 8.52(d) for g bounded and measurable, let $\{\psi_\alpha\}_{\alpha\in\mathbb{N}} \subset \mathsf{H}$ be an orthonormal system (countable since H is separable – the finite case is alike) as the one of Theorem 8.56. As in the mentioned theorem, build orthogonal spaces $\overline{\mathsf{H}_{\psi_\alpha}} \subset \mathsf{H}$ corresponding to those vectors whose orthogonal sum is H. Each $\overline{\mathsf{H}_{\psi_\alpha}}$ is invariant under any $g(T,T^*)$, and isomorphic to $L^2(\sigma(T),\mu_\alpha)$, where $\mu_\alpha(E) := (\psi_\alpha|P^{(T)}(E)\psi_\alpha)$ are the usual positive probability measures, $P^{(T)}$ is the PVM of T and $E \subset \sigma(T)$ a Borel set. The vector ψ_α is described in $L^2(\sigma(T),\mu_\alpha)$ by the constant map 1. The operator $f_n(T,T^*)$ is described in $L^2(\sigma(T),\mu_\alpha)$ by the multiplication by f_n. Now look at $x := \sum_{\alpha\in\mathbb{N}} 2^{-\alpha/2}\psi_\alpha$. The sequence $f_n(T,T^*)x$ is a Cauchy sequence by assumption. Expanding $\mathsf{H} = \oplus_{\alpha\in\mathbb{N}}\overline{\mathsf{H}_{\psi_\alpha}}$, the inequality $||f_n(T,T^*)x - f_m(T,T^*)x||^2 < \varepsilon$, $n,m > N_\varepsilon$, is equivalent to $\sum_{\alpha\in\mathbb{N}} \int_{\sigma(T)} |f_n - f_m|^2 2^{-\alpha} d\mu_\alpha < \varepsilon$, $n,m > N_\varepsilon$, by Theorem 8.56(a). Let $\mu(E) := \sum_{\alpha\in\mathbb{N}} 2^{-\alpha}\mu_\alpha(E)$ be a bounded positive Borel measure. Then the previous condition says $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy in $L^2(\sigma(T),\mu)$, so there is a subsequence (called alike for simplicity) converging μ-almost everywhere to a measurable map $f \in L^2(\sigma(T),\mu)$, possibly unbounded. Since zero-measure sets for μ are so also for each μ_α, convergence holds almost everywhere for $P^{(T)}$ as well, by Theorem 8.56. (In fact for any $x \in \mathsf{H}$ and corresponding maps $g_{x,\alpha} \in L^2(\sigma_\alpha(T),\mu_\alpha)$ we have $(x|P^{(T)}(E)x) = \sum_{\alpha\in\mathbb{N}} \int_{\sigma(T)} \chi_E g_{x,\alpha} d\mu_\alpha = 0$ if $\mu_\alpha(E) = 0$ for any $\alpha \in \mathbb{N}$.) So we may define a closed operator $A := \int_{\sigma(T)} f dP^{(T)}$, with dense domain Δ_f. Call $\mathscr{D} \subset \Delta_f$ the linear space dense in H of finite combinations of $\phi_\alpha \in \overline{\mathsf{H}_{\psi_\alpha}}$ associated to bounded measurable maps in the respective $L^2(\sigma(T),\mu_\alpha)$. By linearity $A\restriction_{\mathscr{D}} = B\restriction_{\mathscr{D}}$ as $||A\phi_\alpha - B\phi_\alpha|| \leq ||(A - f_n(T,T^*))\phi_\alpha|| + ||f_n(T,T^*)\phi_\alpha - B\phi_\alpha||$, where the last term is infinitesimal as $n \to +\infty$ by construction, whereas the penultimate term squared is smaller than $C\sum_{\alpha\in F} \int |f - f_n|^2 d\mu_\alpha$, $C \geq 0$ finite, F finite, $\int |f - f_n|^2 d\mu_\alpha \to 0$ as $n \to +\infty$, as we know. Since B is bounded and $\mathscr{D}$ dense, closing $A\restriction_{\mathscr{D}} = B\restriction_{\mathscr{D}}$ gives $\overline{A\restriction_{\mathscr{D}}} = B$. But $A = \overline{A} \supset \overline{A\restriction_{\mathscr{D}}}$, and B is defined on H, whence $A = B$. Recalling $A = \int_{\sigma(T)} f dP^{(T)}$, by Corollary 9.5 $||f||_\infty^{(P)} < +\infty$. We may redefine f on a zero-measure set for $P^{(T)}$ without changing $A = \int_{\sigma(T)} f dP^{(T)}$. Since f is now bounded, we can define $f(T,T^*)$ as in Theorem 8.39. Thus $f(T,T^*) = \int_{\sigma(T)} f dP^{(T)} = B$. □

9.1.3 Spectral decomposition of unbounded self-adjoint operators

The time is right to prove the *spectral decomposition theorem for unbounded self-adjoint operators*. We shall state it and prove it for unbounded self-adjoint operators, although it holds also for unbounded normal operators [Rud91].

Theorem 9.10 (Spectral decomposition of unbounded self-adjoint operators). *Let T be self-adjoint (possibly unbounded) on the Hilbert space* H.

(a) *There exists a unique PVM $P^{(T)} : \mathscr{B}(\mathbb{R}) \to \mathfrak{B}(\mathsf{H})$, such that*

$$T = \int_{\mathbb{R}} \lambda \, dP^{(T)}(\lambda) . \tag{9.37}$$

(b) *The following identity holds*

$$supp(P^{(T)}) = \sigma(T)\,, \tag{9.38}$$

and in particular:

(i) $\lambda \in \sigma_p(T) \;\Leftrightarrow\; P^{(T)}(\{x\}) \neq 0$;
(ii) $\lambda \in \sigma_c(T) \;\Leftrightarrow\; P^{(T)}(\{x\}) = 0$, *and for any open set* $A_x \subset \mathbb{R}$ *containing* x $P^{(T)}(A_x) \neq 0$;
(iii) *if* $\lambda \in \sigma(T)$ *is isolated, then* $\lambda \in \sigma_p(T)$;
(iv) *if* $\lambda \in \sigma_c(T)$, *then for any* $\varepsilon > 0$ *there exists* $\phi_\varepsilon \in D(T)$, $||\phi_\varepsilon|| = 1$ *with*

$$0 < ||T\phi_\varepsilon - \lambda\phi_\varepsilon|| \leq \varepsilon\,.$$

Proof. (a) Let V be the Cayley transform of T, a unitary operator by Theorem 5.34 because T is self-adjoint. If $\mathbb{S}^1 := \{(x,y) \in \mathbb{R}^2 \mid x^2+y^2 = 1\}$, define $\mathsf{X} := \mathbb{S}^1 \setminus \{(1,0)\}$ and denote $z = x+iy$. Put on X the topology induced by $\mathbb{R}^2$ (or $\mathbb{S}^1$, which is the same) and consider its Borel σ-algebra $\mathscr{B}(\mathsf{X}) \subset \mathscr{B}(\mathbb{S}^1)$. Let also $P_0^{(V)}$ be the spectral measure of V in $\mathbb{S}^1$, stemming from the spectral decomposition Theorem 8.54(a)'. Then

$$V = \int_{\mathbb{S}^1} z dP_0^{(V)}(x,y)\,. \tag{9.39}$$

The operator $I - V$ is one-to-one by (i) in Theorem 5.34(b), so $1 = 1 + i0 \notin \sigma_p(V)$. This in turn implies $P_0^{(V)}(\{(1,0)\}) = 0$ by (i) in Theorem 8.54(b). Consider orthogonal projectors

$$P^{(V)} : \mathscr{B}(\mathsf{X}) \ni E \mapsto P_0^{(V)}(E) \in \mathfrak{P}(\mathsf{H})\,,$$

where $\mathscr{B}(\mathsf{X}) \subset \mathscr{B}(\mathbb{S}^1)$. $P^{(V)}$ is a PVM on X by construction (cf. Definition 8.41); note that $P^{(V)}(\mathsf{X}) = I$ because $P_0^{(V)}(\{(1,0)\}) = 0$:

$$P^{(V)}(\mathsf{X}) := P_0^{(V)}(\mathsf{X}) = P_0^{(V)}(\mathbb{S}^1 \setminus \{(1,0)\}) = P_0^{(V)}(\mathbb{S}^1) - P_0^{(V)}(\{(1,0)\}) = I - 0 = I\,.$$

For the same reason the integral of a simple map s on $\mathbb{S}^1$ in $P_0^{(V)}$ coincides trivially with the integral of $s\restriction_\mathsf{X}$ in $P^{(V)}$. From the construction of the integral of bounded maps, taking $f \in M_b(\mathbb{S}^1)$, hence $f\restriction_\mathsf{X} \in M_b(\mathsf{X})$, it follows that $\int_\mathsf{X} f\restriction_\mathsf{X}\, dP^{(V)} = \int_\mathsf{X} f dP_0^{(V)}$. However we choose $\phi, \psi \in \mathsf{H}$, $E \subset \mathscr{B}(\mathbb{S}^1)$:

$$\mu_{\phi,\psi}^{(P^{(V)})}(E \setminus \{(1,0)\}) = (\phi|P^{(V)}(E \setminus \{(1,0)\})\psi) = (\phi|P_0^{(V)}(E \setminus \{(1,0)\})\psi)$$

$$= (\phi|P_0^{(V)}(E)\psi) = \mu_{\phi,\psi}^{(P_0^{(V)})}(E)\,,$$

(in the obvious notation), using the definition of integral of measurable maps we find $\int_\mathsf{X} f\restriction_\mathsf{X}\, dP^{(V)} = \int_\mathsf{X} f dP_0^{(V)}$ for any $f : \mathbb{S}^1 \to \mathbb{C}$ measurable. In particular, from (9.39) and dropping $\restriction_\mathsf{X}$, we obtain:

$$V = \int_\mathsf{X} z dP^{(V)}(x,y)\,. \tag{9.40}$$

Now define the real-valued, measurable unbounded map on X:

$$f(z) := i\frac{1+z}{1-z} \quad z \in \mathsf{X}, \tag{9.41}$$

and integrate it in the spectral measure $P^{(V)}$ on X, to get the operator (unbounded, in general):

$$T' := \int_{\mathsf{X}} f(z) dP^{(V)}(x,y). \tag{9.42}$$

As f ranges in the reals ($(x,y) \in \mathsf{X}$), T' must be self-adjoint by Theorem 9.4(b). The equation $f(z)(1-z) = i(1+z)$, by virtue of Theorem 9.4(c), implies:

$$T'(I-V) = i(I+V) \tag{9.43}$$

(it is easy to see that there is $=$ in (9.13)). In particular (9.43) implies $Ran(I-V) \subset \Delta_f =: D(T')$. From Theorem 5.34 we know

$$T(I-V) = i(I+V) \quad \text{and } D(T) = Ran(I-V) \subset \Delta_f.$$

Comparing with (9.43) allows to conclude T' is a self-adjoint extension of T. As $T = T^*$ has no proper self-adjoint extension (Proposition 5.17(c)), then $T = T'$. Hence

$$T = \int_{\mathsf{X}} f(z) dP^{(V)}(x,y). \tag{9.44}$$

The function $f : \mathsf{X} \to \mathbb{R}$ is actually bijective and so its range is $\mathbb{R}$. From Theorem 9.4(h) $\mathscr{B}(\mathbb{R}) \ni E \mapsto P^{(T)}(E) := P^{(V)}(f^{-1}(E))$ is a PVM on $\mathbb{R}$ and (9.44) may be written as in (9.37):

$$T = \int_{\mathbb{R}} \lambda dP^{(T)}(\lambda).$$

But this is exactly the spectral expansion we wanted. So let us pass to uniqueness of the measure solving (9.37). Let P' be a PVM on $\mathbb{R}$ with

$$T = \int_{\mathbb{R}} \lambda dP'(\lambda).$$

The Cayley transform, by Theorem 9.4(c), reads

$$V = (T-iI)(T+iI)^{-1} = \int_{\mathbb{R}} \frac{\lambda - i}{\lambda + i} dP'(\lambda).$$

Using statement (h) in the same theorem, with the same measurable $f : \mathsf{X} \to \mathbb{R}$ with measurable inverse of (9.41), we find

$$V = \int_{\mathsf{X}} z dP'(f(x,y)),$$

where $\mathscr{B}(\mathsf{X}) \ni F \mapsto Q(F) := P'(f(F))$ is a PVM on X which we can extend to a PVM on $\mathbb{S}^1$ by $Q_0(F) := Q(F \setminus \{(1,0)\})$, $F \in \mathscr{B}(\mathbb{S}^1)$. Thus

$$V = \int_{\mathbb{S}^1} z dQ_0(x,y).$$

By (9.39), then, as the spectral measure associated to a bounded normal operator is unique by Theorem 8.54, necessarily $Q_0(F) = P_0^{(V)}(F)$ for any Borel set in $\mathbb{S}^1$. Hence $Q(F) = P^{(V)}(F)$ for any Borel set of X. Therefore, for any $E \in \mathscr{B}(\mathbb{R})$, $Q(f^{-1}(E)) = P^{(V)}(f^{-1}(E))$, i.e. $P'(E) = P^{(T)}(E)$, as required.
(b) Now let us show $\sigma(T) = supp(P^{(T)})$, or equivalently, $\lambda_0 \in \rho(T) \Leftrightarrow \lambda_0 \notin supp(P^{(T)})$. So first of all we prove $\lambda_0 \notin supp(P^{(T)})$ implies $\lambda_0 \in \rho(T)$. In fact, there exists an open interval $(a,b) \subset \mathbb{R} \setminus supp(P^{(T)})$, $\lambda_0 \in (a,b)$. Hence $I = \int_{\mathbb{R}} \chi_{\mathbb{R}\setminus(a,b)} dP$, and from the last result in Theorem 9.4(c)

$$\int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda) = \int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda) \int_{\mathbb{R}} \chi_{\mathbb{R}\setminus(a,b)} dP =$$

$$\int_{\mathbb{R}} \chi_{\mathbb{R}\setminus(a,b)}(\lambda) \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda) .$$

By Theorem 9.4(c), as the last integrand is bounded,

$$R_{\lambda_0}(T) := \int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda) \in \mathfrak{B}(\mathsf{H}) .$$

Always by Theorem 9.4(c) (and keeping an eye on the domains of the products):

$$R_{\lambda_0}(T)(T-\lambda_0 I) = I\restriction_{D(T)} , \quad (T-\lambda_0 I)R_{\lambda_0}(T) = I .$$

The second is true everywhere on H, so $Ran(T-\lambda_0 I) = \mathsf{H}$. The operator $R_{\lambda_0}(T)$ is therefore the resolvent of T associated to λ_0 by Theorem 8.4(a), as the name suggests. By definition, then, $\lambda_0 \in \rho(T)$. Conversely let us prove $\lambda_0 \in \rho(T)$ implies $\lambda_0 \notin supp(P^{(T)})$. Under the assumptions on λ_0, $P^{(T)}(\{\lambda_0\}) = 0$, otherwise therewould be $\psi \in P^{(T)}_{\{\lambda_0\}}(\mathsf{H}) \setminus \{0\}$ such that (by Theorem 9.4(c)):

$$T\psi = \int_{\mathbb{R}} \lambda dP^{(T)}(\lambda) P^{(T)}_{\{\lambda_0\}} \psi = \int_{\mathbb{R}} \lambda dP^{(T)}(\lambda) \int_{\mathbb{R}} \chi_{\{\lambda_0\}}(\lambda) dP^{(T)}(\lambda)\psi$$

$$= \int_{\mathbb{R}} \lambda \chi_{\{\lambda_0\}}(\lambda) dP^{(T)}(\lambda)\psi = \int_{\mathbb{R}} \lambda_0 \chi_{\{\lambda_0\}}(\lambda) dP^{(T)}(\lambda)\psi = \lambda_0 P^{(T)}_{\{\lambda_0\}}\psi = \lambda_0\psi$$

and then $\psi \in \sigma_p(T)$, contradicting $\lambda_0 \in \rho(T)$. Furthermore, the resolvent exists (as T is closed and by Theorem 8.4(a, b)). This is the operator $R_{\lambda_0}(T) \in \mathfrak{B}(\mathsf{H})$ that satisfies

$$(T-\lambda_0 I)R_{\lambda_0}(T) = I \quad \text{and} \quad R_{\lambda_0}(T)(T-\lambda_0 I) = I\restriction_{D(T)} .$$

On the other hand Theorem 9.4(c) and $P^{(T)}(\{\lambda_0\}) = 0$ imply

$$\left(\int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda)\right)(T-\lambda_0 I) = I\restriction_{D(T)} , \quad (T-\lambda_0 I)\int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda) = I$$

(again, beware of domains). From the first we also see that the domain of $\int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda)$ is $D(T-\lambda_0 I)$, i.e. H. It does not really matter how one defines $\lambda \mapsto \frac{1}{\lambda-\lambda_0}$ at $\lambda = \lambda_0$, because $P^{(T)}(\{\lambda_0\}) = 0$. By uniqueness of the inverse, then,

$$\int_{\mathbb{R}} \frac{1}{\lambda-\lambda_0} dP^{(T)}(\lambda) = R_{\lambda_0}(T) ,$$

and the operator on the left is bounded. Now suppose by contradiction that $\lambda_0 \in supp(P^{(T)})$. Then any open set containing λ_0, in particular any interval $I_n := (\lambda_0 - 1/n, \lambda_0 + 1/n)$, must satisfy $P^{(T)}(I_n) \neq 0$. Take $\psi_n \in P^{(T)}_{I_n}(\mathsf{H}) \setminus \{0\}$ for any $n = 1, 2, \ldots$. Without loss of generality assume $||\psi_n|| = 1$. Using Theorem 9.4(f) we obtain

$$||R_{\lambda_0}(T)\psi_n||^2 = \left|\left| \int_{\mathbb{R}} \frac{1}{\lambda - \lambda_0} dP^{(T)}(\lambda)\psi_n \right|\right|^2 = \int_{I_n} \frac{1}{|\lambda - \lambda_0|^2} d\mu_{\psi_n}(\lambda)$$

$$\geq \inf_{I_n} \frac{1}{|\lambda - \lambda_0|^2} \int_{I_n} d\mu_{\psi_n}(\lambda) \geq \inf_{I_n} \frac{1}{|\lambda - \lambda_0|^2} = n^2 \to +\infty \quad \text{as } n \to +\infty.$$

We have reached the absurd that $R_{\lambda_0}(T)$ cannot be bounded. Therefore $\lambda_0 \notin supp(P^{(T)})$.

Let us prove (i). As above, if $P^{(T)}(\{x\}) \neq 0$ then $x \in \sigma_p(T)$. Suppose $x \in \sigma_p(T)$. By definition of Cayley transform, $((x-i)/(x+i)) \in \sigma_p(V)$. We may apply (i) in Theorem 8.54(b) to the normal V (unitary) replacing T. Then $P^{(V)}(\{\frac{x-i}{x+i}\}) \neq 0$. Looking at the way the PVM associated to T was obtained from the PVM of V, we see $P^{(T)}(x) = P^{(V)}(\{\frac{x-i}{x+i}\}) \neq 0$.

Now to (ii). By Proposition 8.7(a), $x \in \sigma_c(T)$ means: (1) $x \in \sigma(T)$ but (2) $x \notin \sigma_p(T)$. Assertion (1) implies $x \in supp(P^{(T)})$, so any open set A_x containing x must satisfy $P^{(T)}(A_x) \neq 0$. Number (2) is equivalent to $P^{(T)}(\{x\}) = 0$ (otherwise (i) would give a contradiction).

The proof of (iii) is immediate: if $x \in supp(P^{(T)})$ is an isolated point, then $P^{(T)}(\{x\}) \neq 0$, otherwise x could not belong to $supp(P^{(T)})$, and using (i) the claim follows.

At last, let us prove (iv). If $x \in \sigma_c(T)$, using (ii) on the intervals $I_n := (x - 1/n, x + 1/n)$, $n = 1, 2, \ldots$, we have $P^{(T)}(I_n) \neq 0$. So choose $\psi_n \in P^{(T)}_{I_n}(\mathsf{H})$ with $||\psi_n|| = 1$ for any n. Then

$$||T\psi_n - x\psi_n||^2 = \left(\int_{\mathbb{R}} (\lambda - x) dP^{(T)}(\lambda)\psi_n \middle| \int_{\mathbb{R}} (\lambda - x) dP^{(T)}(\lambda)\psi_n \right) =$$

$$\left(\int_{\mathbb{R}} (\lambda - x) dP^{(T)}(\lambda) P^{(T)}_{I_n}\psi_n \middle| \int_{\mathbb{R}} (\lambda - x) dP^{(T)}(\lambda)\psi_n \right).$$

Using Theorem 9.4(c) the last inner product is

$$\int_{\mathbb{R}} \chi_{I_n}(x)(\lambda - x)^2 d\mu_{\psi_n}(\lambda) \leq \int_{I_n} \sup_{I_n}(\lambda - x)^2 d\mu_{\psi_n}(\lambda) =$$

$$= n^{-2} \int_{I_n} d\mu_{\psi_n}(\lambda) = n^{-2} \int_{\mathbb{R}} d\mu_{\psi_n}(\lambda) = n^{-2} ||\psi_n||^2.$$

So for any $n = 1, 2, \ldots$ there is $\psi_n \neq 0$, $||\psi_n|| = 1$, with $||T\psi_n - x\psi_n|| \leq 1/n$. The claim follows, since $x \notin \sigma_p(T)$ by assumption, and $0 < ||T\psi_n - x\psi_n||$. □

After the spectral theorem we pass to a definition useful for the applications to QM.

Definition 9.11. *Consider a self-adjoint operator T on the Hilbert space H, and $f : \sigma(T) \to \mathbb{C}$ a measurable map. The operator:*

$$f(T) := \int_{\sigma(T)} f(x) dP^{(T)}(x) , \tag{9.45}$$

with domain

$$D(f(T)) = \Delta_f := \left\{ \psi \in \mathsf{H} \;\middle|\; \int_{\sigma(T)} |f(x)|^2 d\mu_\psi^{(T)}(x) < +\infty \right\} ,$$

where $\mu_\psi^{(T)}(E) := (\psi|P^{(T)}(E)\psi)$ for any $E \in \mathscr{B}(\sigma(T))$, is called **function of the operator** T.

Since $supp(P^{(T)}) = \sigma(T)$, the PVM $P^{(T)}$ associated to T can be thought as defined either on $\sigma(T)$ or on $\mathbb{R}$. Even when defined on $\sigma(T)$ only (precisely, on the Borel σ-algebra $\mathscr{B}(\sigma(T))$) we still have $supp(P^{(T)}) = \sigma(T)$ by the definition of support in the subspace $\sigma(T)$ of $\mathbb{R}$ with induced topology. The easy proof is left to the reader (one uses that $\lambda \in supp(P)$ iff $P(A) \neq 0$ for any open set A containg λ, for a generic PVM). Therefore we can view the right integral in (9.45) as living on $\mathbb{R}$, by extending f trivially (as zero) outside $\sigma(T)$ or directly taking a measurable $f : \mathbb{R} \to \mathbb{C}$:

$$f(T) := \int_{\mathbb{R}} f(x) dP^{(T)}(x) ,$$

with

$$D(f(T)) = \Delta_f := \left\{ \psi \in \mathsf{H} \;\middle|\; \int_{\mathbb{R}} |f(x)|^2 d\mu_\psi^{(T)}(x) < +\infty \right\} .$$

In the sequel we shall use the best viewpoint without further explanations. To the reader is also left the obvious check that the definition of $f(T)$ coincides with the known one (relying on the functional calculus for bounded measurable functions, cf. Chapter 8) when $T \in \mathfrak{B}(\mathsf{H})$, $f \in M_b(\sigma(T))$.

Remark 9.12. (1) The spectral theorem allows for a second decompostion of the spectrum of a self-adjoint operator $T : D(T) \to \mathsf{H}$, consisiting in the **discrete spectrum**

$$\sigma_d(T) := \left\{ \lambda \in \sigma(T) \;\middle|\; dim\left(P^{(T)}_{(\lambda-\varepsilon,\lambda+\varepsilon)}(\mathsf{H})\right) \text{ is finite for some } \varepsilon > 0 \right\} ,$$

plus the **essential spectrum** $\sigma_{ess}(T) := \sigma(T) \setminus \sigma_d(T)$.

It is not hard to see that $\lambda \in \sigma_d(T) \;\Leftrightarrow\; \lambda$ is an isolated point in $\sigma(T)$, and as such, λ is an eigenvalue for T with finite-dimensional eigenspace. By Theorem 9.10 $\sigma_d(T) \subset \sigma_p(T)$. In general, though, the opposite inclusion fails, for instance because there may be non-isolated points in $\sigma_p(T)$.

(2) A third spectral decomposition for $T : D(T) \to \mathsf{H}$ self-adjoint arises by splitting the Hilbert space into the closed span H_p of the eigenvectors and its orthogonal complement: $\mathsf{H} = \mathsf{H}_p \oplus \mathsf{H}_p^\perp$. Both $\mathsf{H}_p \cap D(T)$ and $\mathsf{H}_p^\perp \cap D(T)$ are easily T-invariant. With the obvious symbols:

$$T = T\restriction_{\mathsf{H}_p} \oplus T\restriction_{\mathsf{H}_p^\perp} .$$

One calls **purely continous spectrum** the set $\sigma_{pc}(T) := \sigma(T\restriction_{\mathsf{H}_p^\perp})$, where for simplicity $T\restriction_{\mathsf{H}_p^\perp}$ stands for $T\restriction_{D(T)\cap\mathsf{H}_p^\perp}$ here and in the sequel. Then $\sigma(T) = \overline{\sigma_p(T)} \cup \sigma_{pc}(T)$. Note that the latter is not necessarily a disjoint union, and in general $\sigma_{pc}(T) \neq \sigma_c(T)$.
(3) The fourth spectral decomposition of $T : D(T) \to \mathsf{H}$ on the Hilbert space H (and even on a normed space), is that into **approximate point spectrum**

$$\sigma_{ap}(T)$$

$$:= \left\{\lambda \in \sigma(T) \,|\, (T-\lambda I)^{-1} : Ran(T-\lambda I) \to D(T) \text{ does not exist or is not bounded}\right\}$$

and **purely residual spectrum** $\sigma_{pr}(T) := \sigma(T) \setminus \sigma_{ap}(T)$. The unboundedness of $(T-\lambda I)^{-1}$ is equivalent to the existence of $\delta > 0$ with $||(T-\lambda I)\psi|| \geq \delta||\psi||$ for any $\psi \in D(T)$, so we immediately see how the next result comes about, thereby justifying the names: $\lambda \in \sigma_{ap}(T) \Leftrightarrow$ there is a unit $\psi_\varepsilon \in D(T)$ such that

$$||T\psi - \lambda\psi|| \leq \varepsilon$$

for any $\varepsilon > 0$. For self-adjoint operators the above holds for any $\lambda \in \sigma_c(T)$ due to Theorem 8.54(b), but clearly also for $\lambda \in \sigma_p(T)$; since $\sigma(T) = \sigma_p(T) \cup \sigma_c(T)$ in this case, we conclude $\sigma_{ap}(T) = \sigma(T)$ and $\sigma_{pr}(T) = \varnothing$ for every self-adjoint operator.
(4) The last partial spectral classification for self-adjoint operators (cf. [ReSi80, vol. I] and[Gra04]) descends from Lebesgue's Theorem 1.76 on the decomposition of regular Borel measures on $\mathbb{R}$. If T is self-adjoint on the Hilbert space H and μ_ψ is the spectral measure of the vector ψ (Theorem 8.50(c)), we define the sets (all closed spaces):

$\mathsf{H}_{ac} := \{\psi \in \mathsf{H} \,|\, \mu_\psi$ is absolutely continuous for Lebesgue's measure$\}$;
$\mathsf{H}_{sing} := \{\psi \in \mathsf{H} \,|\, \mu_\psi$ is singular and continuous for Lebesgue's measure$\}$;
$\mathsf{H}_{pa} := \{\psi \in \mathsf{H} \,|\, \mu_\psi$ is purely atomic$\}$.

Then we define $\sigma_{ac}(T) := \sigma(T\restriction_{\mathsf{H}_{ac}})$, $\sigma_{sing}(T) := \sigma(T\restriction_{\mathsf{H}_{sing}})$ respectively called **absolutely continuous spectrum of** T and **singular spectrum of** T. It turns out that $\sigma_{ac}(T) \cup \sigma_{sing}(T) = \sigma_{pc}(T)$ and $\overline{\sigma_p(T)} = \sigma(T\restriction_{\mathsf{H}_{pa}})$.
(5) As $supp(P^{(T)}) = \sigma(T)$, definition (9.45) reads:

$$f(T) := \int_{\sigma(T)} f(x) dP^{(T)}(x) \,. \tag{9.46}$$

Likewise, the domain of $f(T)$ is

$$D(f(T)) = \left\{\psi \in \mathsf{H} \,\middle|\, \int_{\sigma(T)} |f(x)|^2 d\mu_\psi^{(T)}(x) < +\infty\right\},$$

as $supp(\mu_{\phi,\psi}) \subset supp(P^{(T)})$. Eventually, if $f : \mathbb{R} \to \mathbb{R}$ is measurable, decomposing $f(T)$ (self-adjoint by Theorem 9.4(b)) under the spectral Theorem 9.10 produces

$$\int_{\sigma(f(T))} \lambda dP^{(f(T))}(\lambda) = \int_{\sigma(T)} f(\lambda) dP^{(T)}(\lambda) = \int_{\sigma(f(T))} \lambda dP^{(T)}(f^{-1}(\lambda)) \,. \tag{9.47}$$

The last identity follows from Theorem 9.4(h). By uniqueness of the PVM associated to $f(T)$ we have

$$P^{f(T)}(E) = P^{(T)}(f^{-1}(E)) \quad \text{for any } E \in \mathscr{B}(\sigma(f(T))). \tag{9.48}$$

(6) Theorem 5.10(d) implies that for any self-adjoint T, the standard domain of a polynomial $p(T)$ and the domain of $p(T)$ thought of as function of T according to Definition 9.11 coincide. By definition of standard domain we also have, for any self-adjoint T:

$$D(T^m) \subset D(T^n)\,, \qquad \text{for any } 0 \leq n \leq m \text{ in } \mathbb{N}. \tag{9.49}$$

■

Functions of an operator enjoy properties that descend directly from 9.3 and 9.4. The next proposition specifies more features of the spectrum of $f(T)$. In order to stay general, we shall state the result for spectral measures that do not necessarily come from self-adjoint operators. But first a definition.

Definition 9.13. *If $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ is a PVM on the second-countable X and $f : \mathsf{X} \to \mathbb{C}$ is measurable, the* **essential rank of** f **with respect to** P*, ess ran$_P(f) \subset \mathbb{C}$, is the closed complement of the union of all open sets $A \subset \mathbb{C}$ such that $P(f^{-1}(A)) = 0$. I.e., $z \in$ ess ran$_P(f)$ $\Leftrightarrow$ $P(f^{-1}(A)) \neq 0$, with $z \in A$ and $A \subset \mathbb{C}$ open.*

If V is the union of said sets A then $P(f^{-1}(V)) = 0$, because V is the union of countably many sets A by Lindelöf's lemma, and PVMs are sub-additive.

Proposition 9.14. *Let $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ be a PVM on the second-countable X and $f : \mathsf{X} \to \mathbb{C}$ a measurable map. If $E_z := f^{-1}(\{z\})$, $z \in \mathbb{C}$ then:*

(a)

$$\sigma\left(\int_{\mathsf{X}} f dP\right) = ess\ ran_P(f)\,,$$

and in particular, for $z \in$ ess ran$_P(f)$:

(i) $P(E_z) \neq 0 \Rightarrow z \in \sigma_p\left(\int_{\mathsf{X}} f dP\right)$;
(ii) $P(E_z) = 0 \Rightarrow z \in \sigma_c\left(\int_{\mathsf{X}} f dP\right)$,

(hence $\sigma_r\left(\int_{\mathsf{X}} f dP\right) = \varnothing$ even if f is not essentially bounded).
(b) *If $f : \sigma(T) \to \mathbb{C}$ is continuous and T self-adjoint, $\sigma(f(T)) = \overline{f(\sigma(T))}$, with bar denoting closure.*
(c) *If $f : \sigma(T) \to \mathbb{C}$ is continuous and $T \in \mathfrak{B}(\mathsf{H})$ normal, $\sigma(f(T)) = f(\sigma(T))$.*
(d) *If $f : \sigma(T) \to \mathbb{C}$ is measurable and T as in (b) or (c), then $\sigma_p(f(T)) \supset f(\sigma_p(T))$ (not an equality, in general).*

Proof. (a) In the sequel $\Psi(f) := \int_{\mathsf{X}} f dP$ and we assume $z = 0$ without loss of generality. Let us prove (i). If $P(E_0) \neq 0$, there is $x \in P(E_0)(\mathsf{H})$ with $||x|| = 1$. Call $\chi := \chi_{E_0}$, so $f\chi = 0$ and $\Psi(f)\Psi(\chi) = 0$ by Theorem 9.4(c). As $\Psi(\chi) = P(E_0)$, $\Psi(f)x = \Psi(f)P(E_0)x = \Psi(f)\Psi(\chi)x = 0$, proving (i).

Now to (ii). By assumption $P(E_0) = 0$, but $P(F_n) \neq 0$ if $F_n := \{s \in \mathsf{X} \,|\, |f(s)| < 1/n\}$, $n = 1, 2, \ldots$, because $z \in$ ess ran$_P(f)$. Choose $x_n \in P(F_n)(\mathsf{H})$, $||x_n|| = 1$ and let $\chi_n := \chi_{F_n}$. As before $||\Psi(f)x_n|| = ||\Psi(f\chi_n)x_n|| \leq ||\Psi(f\chi_n)|| = ||f\chi_n||_\infty \leq 1/n$. Therefore $\Psi(f)x_n \to 0$, notwithstanding $||x_n|| = 1$. This shows that $\Psi(f)^{-1}$ (and, similarly, $(\Psi(f) - zI)^{-1}$) cannot be bounded if it exists. To show 0 (z in general) is in the continuous spectrum we need prove $Ker(\Psi(f)) = \{0\}$ and $\overline{Ran(\Psi(f))} = \mathsf{H}$. Suppose $\Psi(f)x = 0$ for some $x \in \Delta_f$. Then

$$\int_{\mathsf{X}} |f| d\mu_x = ||\Psi(f)x||^2 = 0\,.$$

Since $|f| > 0$ almost everywhere for μ_x, it follows $0 = \mu_x(\mathsf{X}) = ||x||^2$. That is, $Ker(\Psi(f)) = \{0\}$. To finish part (ii) we prove $\overline{Ran(\Psi(f))} = \mathsf{H}$. Since $\Psi(f)^* = \Psi(\overline{f})$, the same argument used above tells $Ker(\Psi(f)^*) = \{0\}$ and $\overline{Ran(\Psi(f))} = (Ker(\Psi(f)^*)^\perp)^\perp = \{0\}^\perp = \mathsf{H}$.

The remains the first assertion in (a). By (i)–(ii) we have ess $\mathrm{ran}_P(f) \subset \sigma(\Psi(f))$. For the opposite inclusion assume $z = 0 \notin$ ess $\mathrm{ran}_P(f)$ ($z \neq 0$ is analogous). Then $f' := 1/f$ is bounded and $ff' = 1$, so $\psi(f)\Psi(f') = I$ and $Ran(\Psi(f)) = \mathsf{H}$. Since $|f| > 0$, $\Psi(f)$ is one-to-one as in case (ii). Therefore $\Psi(f)^{-1} \in \mathfrak{B}(\mathsf{H})$ by the closed graph theorem. This ends the proof because it implies $0 \notin \sigma(\Psi(f))$ by Theorem 8.4(a).
(b) Recall $supp(P^{(T)}) = \sigma(T)$ (viewing $P^{(T)}$ defined on $\mathbb{C}$, or only on $\sigma(T)$ if we use the induced topology). Certainly $f(supp(P^{(T)})) \subset$ ran $\mathrm{ess}_P(f)$ (if $z \in f(supp(P^{(T)}))$, for any open set $A \subset \mathbb{C}$, $A \ni z$, the open set $f^{-1}(A)$ is in $supp(P^{(T)})$ and so $P^{(T)}(f^{-1}(A)) \neq 0$). Because ran $\mathrm{ess}_{P(T)}(f)$ is closed, we have $\overline{f(supp(P^{(T)}))} \subset$ ran $\mathrm{ess}_{P(T)}(f)$. If $z \in$ ran $\mathrm{ess}_{P(T)}(f)$ but $z \notin \overline{f(supp(P^{(T)}))}$, there would be an open set $A \ni z$ not intersecting $f(supp(P^{(T)}))$. Thus $P^{(T)}(f^{-1}(A)) = P^{(T)}(\varnothing) = 0$, which cannot be by definition of ran $\mathrm{ess}_{P(T)}(f)$.
(c) The statement is straighforward from (b). Note, if $T \in \mathfrak{B}(\mathsf{H})$, that the spectrum $\sigma(T)$ is compact and its image in $\mathbb{C}$ under the continuous f is compact, so closed.
(d). If $\lambda \in \sigma_p(T)$ then $P^{(T)}(\{\lambda\}) \neq \{0\}$. If $x \in P^{(T)}(\{\lambda\})(\mathsf{H}) \setminus \{0\}$, using Theorem 9.4(c) we get $f(T)x = f(T)\chi_{\{\lambda\}}(T)x = f(\lambda)x$, hence $f(\lambda) \in \sigma_p(f(T))$.
Here is an example where $\sigma_p(f(T)) \supsetneq f(\sigma_p(T))$ for $T = T^*$. Take $(a,b) \subset \sigma_c(T)$ so that $P^{(T)}((a,b)) \neq 0$. If f is measurable, it equals a constant c on (a,b) and $f(\lambda) < c$ outside the interval, so $c \in \sigma_p(f(T))$ by (i) in (a). But $c \notin f(\sigma_p(T)) < c$ and hence $\sigma_p(f(T)) \supsetneq f(\sigma_p(T))$. □

9.1.4 Example with pure point spectrum: the Hamiltonian of the harmonic oscillator

On the complex Hilbert space $L^2(\mathbb{R}, dx)$ (dx is the Lebesgue measure on $\mathbb{R}$) consider the operator

$$H_0 := -\frac{1}{2m}\left(P\restriction_{\mathscr{S}(\mathbb{R})}\right)^2 + \frac{m\omega^2}{2}\left(X\restriction_{\mathscr{S}(\mathbb{R})}\right)^2,$$

where X, P are the position and momentum operators for a particle moving on the real line, seen in Chapter 5. In other terms

$$H_0 := -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{m\omega^2}{2}x^2,$$

where x^2 stands for the multiplication by $\mathbb{R} \ni x \mapsto x^2$ and $\hbar$, ω, m are positive constants. Define $D(H_0) := \mathscr{S}(\mathbb{R})$, where $\mathscr{S}(\mathbb{R})$ is the Schwartz space of $\mathbb{R}$, i.e. the space of smooth complex functions that vanish at infinity, together with any derivative, faster than any negative power of x (see Example 2.87).

The numbers $\hbar$, ω, m have no mathematical relevance (and could be set to 1 in the sequel), yet it is their physical meaning that is important. The operator H_0 is called the **Hamiltonian of the one-dimensional harmonic oscillator** with charactersitic frequency $\omega/(2\pi)$, for a particle of mass m; $h := 2\pi\hbar$ is Planck's constant. Physically, $\overline{H_0}$ is the *energy observable* of the system under exam; in this section, though, we shall not be concerned with the physical background; we will just study the operator from a mathematical perspective, leaving any comment about the physics to Chapter 12, 13.

H_0 is evidently symmetric as Hermitian and because $\mathscr{S}(\mathbb{R})$ is dense in $L^2(\mathbb{R}, dx)$. H_0 admits self-adjoint extensions by von Neumann's criterion (Theorem 5.43), for it commutes with the (antiunitary) complex conjugation of $L^2(\mathbb{R}, dx)$. We will show H_0 is essentially self-adjoint, provide an explicit expression for it in terms of the spectral expansion of its unique self-adjoint extension $\overline{H_0}$, and also describe the spectrum.

Let us introduce three operators, called **creation operator**, **annihilation operator** and **number operator**:

$$A^\star := \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{\hbar}{m\omega}\frac{d}{dx}\right), \quad A := \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{\hbar}{m\omega}\frac{d}{dx}\right), \quad \mathscr{N} := A^\star A\,.$$

In this case, as well, we assume the operators are densely defined on $D(A) = D(A^\star) = D(\mathscr{N}) := \mathscr{S}(\mathbb{R})$. It should be clear that $A^\star \subset A^*$, justifying the notation, and $\mathscr{N}$ is further symmetric. Notice $\mathscr{S}(\mathbb{R})$ is dense and invariant under $H_0, A, A^\star$. Using $A, A^\star$ we will build eigenvectors for $\mathscr{N}$ and H_0 that form a basis in $L^2(\mathbb{R}, dx)$. As eigenvectors are obviously analytic vectors, by Nelson's criterion (Theorem 5.47) H_0, $\mathscr{N}$ are essentially self-adjoint on their domain $\mathscr{S}(\mathbb{R})$.

We start by observing that, by definition, the commutation relation

$$[A, A^\star] = I\,, \tag{9.50}$$

hold, where either side acts on the dense invariant space $\mathscr{S}(\mathbb{R})$. The proof is immediate. What is more, still by definition,

$$H_0 = \hbar\omega\left(A^\star A + \frac{1}{2}I\right) = \hbar\omega\left(\mathscr{N} + \frac{1}{2}I\right). \tag{9.51}$$

Consider the equation in $\mathscr{S}(\mathbb{R})$:

$$A\psi_0 = 0\,. \tag{9.52}$$

A solution is, easily,

$$\psi_0(x) = \frac{1}{\pi^{1/4}\sqrt{s}}\, e^{-\frac{x^2}{(2s)^2}}\,, \quad s := \sqrt{\frac{\hbar}{m\omega}}\,,$$

where the factor was chosen so to normalise $||\psi_0|| = 1$. The function ψ_0 is the first Hermite function introduced in Example 3.32(4), provided we use the variable $x' = x/s$ and consider the factor $1/\sqrt{s}$ not to destroy the normalisation. Now define vectors:

$$\psi_n := \frac{(A^\star)^n}{\sqrt{n!}}\psi_0 \tag{9.53}$$

for $n = 1, 2, \ldots$. Only using (9.52), (9.50) it is easy to prove by induction that

$$A\psi_n = \sqrt{n}\psi_{n-1}\,, \quad A^\star\psi_n = \sqrt{n+1}\psi_{n+1}\,, \quad (\psi_n|\psi_m) = \delta_{nm}\,, \tag{9.54}$$

$n, m \in \mathbb{N}$. The second identity actually follows from the definition of the ψ_n, whilst the first is proved like this:

$$A\psi_n = \frac{1}{\sqrt{n!}}A(A^\star)^n\psi_0 = \frac{1}{\sqrt{n!}}[A,(A^\star)^n]\psi_0 + \frac{1}{\sqrt{n!}}(A^\star)^n A\psi_0 = \frac{1}{\sqrt{n!}}[A,(A^\star)^n]\psi_0 + 0\,;$$

but (9.50) implies $[A,(A^\star)^n] = n(A^\star)^{n-1}$, substituting which above gives what needed. Here is the proof of the third identity (for $n \geq m$, the other case is similar):

$$(\psi_m|\psi_n) = \frac{1}{\sqrt{n!m}}(\psi_{m-1}|A(A^\star)^n\psi_0) = \frac{1}{\sqrt{n!m}}(\psi_{m-1}|[A,(A^\star)^n]\psi_0)$$

$$= \frac{n}{\sqrt{n!m}}(\psi_{m-1}|(A^\star)^{n-1}\psi_0) = \sqrt{\frac{n}{m}}(\psi_{m-1}|\psi_{n-1}) = \cdots = \sqrt{\frac{n!}{m!(n-m)!}}(\psi_0|\psi_{n-m})\,.$$

If $n = m$ the result is 1, otherwise 0, for

$$(\psi_0|\psi_{n-m}) = (n-m)^{-1/2}(\psi_0|A^\star\psi_{n-m-1}) = (n-m)^{-1/2}(A\psi_0|\psi_{n-m-1}) = 0\,.$$

The second equation in (9.54) (the normalisation is preserved when using $x' = x/s$ because of $1/\sqrt{s}$) is the recurrence relationship of Hermite functions mentioned in Example 3.32(4). Hence the ψ_n are (up to a multiplicative constant and a change of variable) Hermite functions, and so they are a basis of $L^2(\mathbb{R}, dx)$. The last equation in (9.54) implies $\{\psi_n\}_{n\in\mathbb{N}}$ is, as it should, an orthonormal system in $L^2(\mathbb{R}, dx)$; the first two tell

$$\mathscr{N}\psi_n = n\psi_n\,, \tag{9.55}$$

so by (9.51) the ψ_n are a basis of eigenvectors of H_0, as:

$$H_0\psi_n = \hbar\omega\left(n + \frac{1}{2}\right)\psi_n\,. \tag{9.56}$$

By the way this proves H_0 (but also $\mathscr{N}$) is unbounded, since the set $\{||H_0\psi|| \mid \psi \in D(H_0)\,, ||\psi|| = 1\}$ contains all numbers $\hbar\omega(n + 1/2)$, $n \in \mathbb{N}$. By Nelson's criterion (Theorem 5.47) the symmetric operators $\mathscr{N}$, H_0 are both essentially self-adjoint, since their domains contain a set $\{\psi_n\}_{n\in\mathbb{N}}$ of analytic vectors spanning a dense subset in $L^2(\mathbb{R}, dx)$.

To obtain the spectral decomposition of $\overline{H_0}$ consider $N = \{\psi_n\}_{n\in\mathbb{N}}$ as topological space with power-set topology, and consider the spectral measure

$$P : \mathscr{B}(N) \ni E \mapsto P_E = \sum_{\psi_n \in E} \psi_n (\psi_n | \ )$$

as in Example 9.8(1). Construct an analogous spectral measure on $\mathbb{R}$ with support on the eigenvalues of the ψ_n for H_0:

$$P'_F := \text{s-} \sum_{\hbar\omega(n+1/2)\in F} \psi_n (\psi_n | \ ) \quad \text{for } F \in \mathscr{B}(\mathbb{R}).$$

Take the measurable function $\phi : N \ni \psi_n \mapsto \hbar\omega\left(n+\frac{1}{2}\right)$. Using Theorem 9.4(c), for any measurable $f : \mathbb{R} \to \mathbb{C}$:

$$\int_{\mathbb{R}} f(x) dP'(x) = \int_N f(\phi(z)) dP(z) = \text{s-}\sum_{n\in\mathbb{N}} f\left(\hbar\omega\left(n+\frac{1}{2}\right)\right) \psi_n (\psi_n | \ )$$

where the last equality is (9.33). Taking f to be $\mathbb{R} \ni x \mapsto x$, we obtain the explicit expression

$$H'_0 := \int_{\mathbb{R}} x dP'(x) = \text{s-}\sum_{n\in\mathbb{N}} \hbar\omega\left(n+\frac{1}{2}\right) \psi_n (\psi_n | \ ) . \tag{9.57}$$

We claim $H'_0 = \overline{H_0}$. Let $< N >$ be the dense space spanned by finite combinations of the ψ_n. $H_0\restriction_{<N>}$ is still essentially self-adjoint by Nelson's criterion. Thus $\overline{H_0} = \overline{H_0\restriction_{<N>}}$, i.e. H_0 and $H_0\restriction_{<N>}$ have the same (unique) self-adjoint extension (their closure). On the other hand H'_0 is certainly a self-adjoint extension of $H_0\restriction_{<N>}$, because (9.57) implies

$$H'_0 \psi_n = \omega\left(n+\frac{1}{2}\right)\psi_n = H_0 \psi_n$$

for any n, and so $H'_0\restriction_{<N>} = H_0\restriction_{<N>}$. Therefore H'_0 must be the unique self-adjoint extension of $H_0\restriction_{<N>}$, hence of H_0. Then the spectral measure associated to $\overline{H_0}$ by the spectral decomposition Theorem 9.10 is

$$P'_F := \text{s-} \sum_{\hbar\omega(n+1/2)\in F} \psi_n (\psi_n | \ ) , \quad F \in \mathscr{B}(\mathbb{R});$$

we also have the spectral decomposition of $\overline{H_0}$ into

$$\overline{H_0} = \text{s-}\sum_{n\in\mathbb{N}} \hbar\omega\left(n+\frac{1}{2}\right) \psi_n (\psi_n | \ ) .$$

Eventually, using Theorem 9.10(b), from the latter we obtain

$$\sigma(\overline{H_0}) = \sigma_p(\overline{H_0}) = \left\{ \hbar\omega\left(n+\frac{1}{2}\right) \,\middle|\, n \in \mathbb{N} \right\} .$$

We must remark that the spectrum of $\overline{H_0}$ is a pure point spectrum and eigenspaces are all finite-dimensional, even though the operator itself is not compact (it is unbounded). Yet the first and second inverse powers of $\overline{H_0}$ are compact, for they are a Hilbert–Schmidt operator and a trace-class operator respectively (exercise).

The numbers in $\sigma_p(\overline{H_0})$ are, physically, the levels of total mechanical energy that a quantum oscillator may assume for given ω, m, in contrast to the classical case where the energy varies with continuity.

9.1.5 Examples with pure continuous spectrum: the operators position and momentum

We return to the operators position X_i (5.12)-(5.13) and momentum P_i (5.17)-(5.18), $i = 1,2,3$, on the Hilbert space $\mathsf{H} = L^2(\mathbb{R}^3, dx)$ with Lebesgue measure. In the sequel $x = (x_1, x_2, x_3)$. We saw they are self-adjoint, and at present we will determine their spectra and spectral expansion.

Start by the position operator X_1. The findings will work for X_2 and X_3 by swapping names. A PVM on $\mathbb{R}$ with values in $\mathfrak{B}(\mathsf{H}) = \mathfrak{B}(L^2(\mathbb{R}^3, dx))$ is

$$(P(E)\psi)(x_1,x_2,x_3) = \chi_E(x_1)\psi(x_1,x_2,x_3) \quad \text{for any } E \in \mathscr{B}(\mathbb{R}),\ \psi \in L^2(\mathbb{R}^3, dx). \tag{9.58}$$

If $\psi \in L^2(\mathbb{R}^3, dx)$, it is easy to see the measure μ_ψ on $\mathscr{B}(\mathbb{R})$ is defined by:

$$\mu_\psi(E) = \int_{E\times\mathbb{R}^2} |\psi(x_1,x_2,x_3)|^2 dx\,, \quad E \in \mathscr{B}(\mathbb{R}),$$

so

$$\int_{\mathbb{R}} g(y)d\mu_\psi(y) = \int_{E\times\mathbb{R}^2} f(x_1)\psi(x_1,x_2,x_3)dx \tag{9.59}$$

for $g : \mathbb{R} \to \mathbb{C}$ measurable. In analogy to Example 9.8(2) it is easy to check, for $f : \mathbb{R} \to \mathbb{C}$ measurable and $\psi \in \Delta_f$ (i.e. $\int_{\mathbb{R}} |f(x_1)\psi(x_1,x_2,x_3)|^2 dx < +\infty$), that

$$\left(\int_{\mathbb{R}} f(y)dP(y)\psi\right)(x_1,x_2,x_3) = f(x_1)\psi(x_1,x_2,x_3) \quad \text{a.e. for } (x_1,x_2,x_3) \in \mathbb{R}^3. \tag{9.60}$$

We can then introduce the operator X_1' associated, in (9.60), to the map $f := f_1 : \mathbb{R} \ni y \mapsto y$. It is self-adjoint by Theorem 9.4(b), as the map is real. By comparison with (5.13) we infer $\Delta_{f_1} = D(X_1)$, and from (9.60) we get

$$X_1'\psi = X_1\psi \quad \text{for any } \psi \in D(X_1).$$

The spectral decomposition 9.10 warrants uniqueness of the spectral measure, whence (9.58) is the spectral measure associated to the position operator X_1. The spectral expansion of X_i, $i = 1,2,3$, must thus be

$$\left(\int_{\mathbb{R}} y dP^{(X_i)}(y)\psi\right)(x_1,x_2,x_3) = (X_i\psi)(x_1,x_2,x_3) \quad \text{a.e. for } (x_1,x_2,x_3) \in \mathbb{R}^3, \tag{9.61}$$

where

$$(P^{(X_i)}(E)\psi)(x_1,x_2,x_3) = \chi_E(x_i)\psi(x_1,x_2,x_3) \quad \forall E \in \mathscr{B}(\mathbb{R}), \psi \in L^2(\mathbb{R}^3, dx). \tag{9.62}$$

This spectral measure allows to find the spectrum of X_i, $i = 1,2,3$. Applying (ii) in Theorem 9.10(b) immediately gives

$$\sigma(X_i) = \sigma_c(X_i) = \mathbb{R}\,. \tag{9.63}$$

Now to momenta. The argument is rather straighforward because of Proposition 5.31, since the Fourier-Plancherel transform is unitary. As such, it preserves spectra (Exercise 8.9), so if K_i are the position operators (as in Proposition 5.31):

$$\sigma(P_i) = \sigma(\hbar\hat{\mathscr{F}}^{-1}K_i\hat{\mathscr{F}}) = \hbar\mathbb{R} = \mathbb{R}\,,$$

i.e.

$$\sigma(P_i) = \sigma_c(P_i) = \mathbb{R}\,. \tag{9.64}$$

The spectral measure of P_i must be supported on the whole $\mathbb{R}$. The reader may prove easily, using Proposition 5.31 and Exercises 9.1–9.5, that the PVM associated to the momentum P_i is just

$$P^{(P_i)}(E) = \hat{\mathscr{F}}^{-1}P^{(K_i)}\hat{\mathscr{F}}\,, \quad E \in \mathscr{B}(\mathbb{R}) \tag{9.65}$$

where $P^{(K_i)}$ is the spectral measure of the position operator K_i.

9.1.6 Spectral representation of unbounded self-adjoint operators

The next *spectral representation* theorem generalises Theorem 8.56 to self-adjoint unbounded operators. The details are left as exercise, as they essentially replicate the proof of Theorem 8.56.

Theorem 9.15 (Spectral representation of unbounded self-adjoint operators). *Let H be a Hilbert space, $T : D(T) \to \mathsf{H}$ a self-adjoint operator in H, $P^{(T)}$ the PVM of T according to Theorem 9.10.*

(a) *H may be decomposed as a Hilbert sum (Definition 7.34) $\mathsf{H} = \oplus_{\alpha\in A}H_\alpha$ (A countable, at most, if H is separable), whose summands H_α are closed and orthogonal. Moreover:*

(i) *for any $\alpha \in A$, $T(\mathsf{H}_\alpha) \subset \mathsf{H}_\alpha$;*
(ii) *for any $\alpha \in A$ there exist a unique finite positive Borel measure μ_α on $\sigma(T) \subset \mathbb{R}$, and a surjective isometric operator $U_\alpha : \mathsf{H}_\alpha \to L^2(\sigma(T), \mu_\alpha)$, such that:*

$$U_\alpha\left(\int_{\sigma(T)} f(x)dP^{(T)}(x)\right)\restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = f\cdot$$

for any measurable $f : \sigma(T) \to \mathbb{C}$. In particular

$$U_\alpha T\restriction_{\mathsf{H}_\alpha} U_\alpha^{-1} = x\cdot\,,$$

where $f\cdot$ is the multiplication by f on $L^2(\sigma(T),\mu_\alpha)$:

$$(f\cdot g)(x) = f(x)g(x) \text{ a.e. on } \sigma(T) \text{ if } g\,, f\cdot g \in L^2(\sigma(T),\mu_\alpha)\,.$$

(b) *if $supp\{\mu_\alpha\}_{\alpha\in A}$ is the complementary set to the numbers $\lambda\in\mathbb{R}$ for which there exists an open set $A_\lambda\subset\mathbb{R}$ with $A_\lambda\ni\lambda$, $\mu_\alpha(A_\lambda)=0$ for any $\alpha\in A$, then*

$$\sigma(T) = supp\{\mu_\alpha\}_{\alpha\in A}\,.$$

(c) *If H is separable, there exist a measure space $(\mathsf{M}_T,\Sigma_T,\mu_T)$, $\mu_T(\mathsf{M}_T)<+\infty$, a map $F_T: M_T\to\mathbb{R}$ and a unitary operator $U_T:\mathsf{H}\to L^2(\mathsf{M}_T,\mu_T)$ such that:*

$$\left(U_T T U_T^{-1} f\right)(m) = F_T(m)f(m)\,,\quad f\in L^2(\mathsf{M}_T,\mu_T),\ U_T^{-1}f\in D(T). \tag{9.66}$$

Proof. The proof mimicks Theorem 8.56 for T self-adjoint and any H_ψ. Apart from the obvious adaptations, it suffices to replace bounded measurable maps $M_b(\sigma(T))$ with the space $L^2(\sigma(T),\mu_\psi)$, paying attention to domains. □

9.1.7 Joint spectral measures

The final notion of this section is the *joint spectral measure* of a set of self-adjoint operators with commuting spectral measures.

Theorem 9.16. *Let $\mathbf{A} := \{A_1,A_2,\dots,A_n\}$ be a set of self-adjoint operators (even unbounded) on the separable Hilbert space H, and suppose the associated spectral measures $P^{(A_k)}$ commute:*

$$P^{(A_k)}(E)P^{(A_h)}(E') = P^{(A_h)}(E')P^{(A_k)}(E)\,,\quad E,E'\in\mathscr{B}(\mathbb{R}),\ h,k\in\{1,2,\dots,n\}.$$

Then there exists a unique PVM $P^{(\mathbf{A})}:\mathscr{B}(\mathbb{R}^n)\to\mathfrak{B}(\mathsf{H})$ such that

$$P^{(\mathbf{A})}(E^{(1)}\times\cdots\times E^{(n)}) = P^{(A_1)}(E^{(1)})\cdots P^{(A_n)}(E^{(n)})\,,\quad E^{(k)}\in\mathscr{B}(\mathbb{R}),\ k=1,\dots,n. \tag{9.67}$$

$P^{(\mathbf{A})}$ is the **joint spectral measure** *of $A_1,A_2,\dots,A_n$ and $supp(P^{(\mathbf{A})})$ is the* **joint spectrum** *of* $\mathbf{A}$.
For any measurable $f:\mathbb{R}\to\mathbb{C}$:

$$\int_{\mathbb{R}^n} f(x_k(x))dP^{(\mathbf{A})}(x) = \int_{\mathbb{R}} f(x_k)dP^{(A_k)}(x_k) = f(A_k)\,,\quad k=1,2,\dots,n, \tag{9.68}$$

where $x_k(x)$ is the kth component of $x=(x_1,x_2,\dots,x_k\dots,x_n)\in\mathbb{R}^n$.

Proof. We need a couple of technical lemmas.

Lemma 9.17. *Let H be a separable Hilbert space, $\{P_\alpha\}_{\alpha\in A}\subset\mathfrak{P}(\mathsf{H})$ an infinite family of orthogonal projectors such that $P_\alpha P_{\alpha'} = P_{\alpha'}P_\alpha = P_\beta$ for any $\alpha,\alpha'\in A$ and some $\beta\in A$ depending on α,α'. Let $\mathsf{M}_a := P_\alpha(\mathsf{H})$, $\mathsf{M} := \cap_{\alpha\in A}\mathsf{M}_\alpha$ and P_M be the orthogonal projector onto M.*

(a) *There is a countable subfamily $\{\mathsf{M}_{\alpha_m}\}_{m\in\mathbb{N}}$ such that $\cap_{m\in\mathbb{N}}\mathsf{M}_{\alpha_m} = \mathsf{M}$.*
(b) $(\psi|P_\mathsf{M}\psi) = \inf_{\alpha\in A}(\psi|P_\alpha\psi)$ *for any $\psi\in\mathsf{H}$.*

Proof. (a) We have $\mathsf{H}\setminus\mathsf{M} = \cup_{\alpha\in A}(\mathsf{H}\setminus\mathsf{M}_\alpha)$, where the $\mathsf{H}\setminus\mathsf{M}_\alpha$ are an open covering of $\mathsf{H}\setminus\mathsf{M}$. As H is separable, it is second countable (see Remark 2.82(2), noting the topology of H is induced by the norm distance). By Theorem 1.8 we can take a countable subcovering $\mathsf{H}\setminus\mathsf{M} = \cup_{m\in\mathbb{N}}(\mathsf{H}\setminus\mathsf{M}_{\alpha_m})$. Now we take complements in H, and obtain (a).
(b) As $\mathsf{M}\subset\mathsf{M}_\alpha$, by Proposition 7.16(a) $P_\mathsf{M} \le P_{\mathsf{M}_\alpha}$, so $(\psi|P_\mathsf{M}\psi) \le \inf_{\alpha\in A}\{(\psi|P_\alpha\psi)\}$ for any $\psi\in\mathsf{H}$. By part (a) and the second equation in (ii) of Theorem 7.18(b) we have $(\psi|P_\mathsf{M}\psi) = \lim_{N\to+\infty}(\psi|P_{\alpha_1}\cdots P_{a_N}\psi) = \lim_{N\to+\infty}(\psi|P_{\beta_N}\psi)$, where $I \ge P_{\beta_N} \ge P_{\beta_{N+1}} \ge \cdots \ge 0$ by construction, because of Proposition 7.16(a). Therefore $(\psi|P_\mathsf{M}\psi) = \inf_{N\in\mathbb{N}}(\psi|P_{\beta_N}\psi) \ge \inf_{\alpha\in A}\{(\psi|P_\alpha\psi)\}$. But the opposite inequality holds too, so (b) is proved. □

Lemma 9.18. *Let $\mathscr{A}$ be an algebra (Definition 1.30) or a σ-algebra of subsets in* X *. If $P : \mathscr{A} \to \mathfrak{P}(\mathsf{H})$,* H *Hilbert space, satisfies (c) and (d) in Definition 8.41 (the latter if $\cup_n E_n \in \mathscr{A}$ algebra), then it also satisfies (a) and (b) of that definition.*

Proof. The proof is the same as for Proposition 8.43(a). □

It is easy to see that finite unions of products $E^{(1)}\times\cdots\times E^{(n)}$, with $E^{(k)}\in\mathscr{B}(\mathbb{R})$, form an *algebra of sets*, denoted $\mathscr{B}_0(\mathbb{R}^n)$; the same can be obtained by taking *disjoint* finite unions of products (just decompose further in case of non-empty intersections). The σ-algebra generated by $\mathscr{B}_0(\mathbb{R}^n)$ contains countable unions of products of open balls in $\mathbb{R}$: as $\mathbb{R}^n$ is second countable, the σ-algebra includes all open sets in $\mathbb{R}^n$, so *a fortiori* the Borel σ-algebra $\mathscr{B}(\mathbb{R}^n)$, and then it must concide with the latter.

If $S = \cup_{j=1}^N E_j^{(1)}\times\cdots\times E_j^{(n)} \in \mathscr{B}_0(\mathbb{R}^n)$ with $(E_j^{(1)}\times\cdots\times E_i^{(n)})\cap(E_i^{(1)}\times\cdots\times E_j^{(n)}) = \varnothing$, $i\neq j$, define:

$$Q(S) := \sum_{j=1}^{N} P^{(A_1)}(E_j^{(1)})\cdots P^{(A_n)}(E_j^{(n)})\,.$$

Since $P^{(A_k)}(E_j^{(k)})$ are commuting orthogonal projectors, every $Q(S)$ is an orthogonal projector that commutes with every other $Q(S')$. It is not hard to prove $\mathscr{B}_0(\mathbb{R}^n)\ni S\mapsto Q(S)$ satisfies $Q(\varnothing)=0$, $Q(\mathbb{R}^n)=I$, and s-$\sum_{n\in\mathbb{S}}Q(S_n)\in\mathscr{P}(\mathsf{H})$ exists when $S_k\cap S_h=\varnothing$, $h\neq k$; moreover the result is $Q(\cup_{k\in\mathbb{N}}S_k)$ if $\cup_{k\in\mathbb{N}}S_k\in\mathscr{B}_0(\mathbb{R}^n)$. Applying Lemma 9.18 gives $Q(S_1)Q(S_2)=Q(S_1\cap S_2)$ if $S_1,S_2\in\mathscr{B}_0(\mathbb{R}^n)$.

If $R\in\mathscr{B}(\mathbb{R}^n)$ let $P^{(\mathbf{A})}(R)$ be the projector onto the intersection of all projection spaces of $\sum_k Q(S_k)$, for any family $\{S_k\}_{k\in\mathbb{N}}\subset\mathscr{B}_0(\mathbb{R}^n)$ such that $S_k\cap S_h=\varnothing$ for $h\neq k$, $\cup_{n\in\mathbb{N}}S_k\supset R$. By construction $P^{(\mathbf{A})}(\mathbb{R}^n)=I$: if $\cup_{k\in\mathbb{N}}S_k=\mathbb{R}^n$, for $\mathbb{R}\in\mathscr{B}_0(\mathbb{R}^n)$, σ-additivity implies $\sum_k Q(S_k)=Q(\mathbb{R}^n)=I$. The latter projects onto H, so $P(\mathbb{R}^n)=I$. Using Lemma 9.17, with $\psi\in\mathsf{H}$:

$$(\psi|P^{(\mathbf{A})}(R)\psi)$$

$$=\inf\left\{\left(\psi\left|\sum_{k\in\mathbb{N}}Q(S_k)\psi\right.\right)\middle|\bigcup_{k\in\mathbb{N}}S_k\supset R\,,\{S_k\}_{k\in\mathbb{N}}\subset\mathscr{B}_0(\mathbb{R}^n)\,,S_k\cap S_h=\varnothing\text{ for }k\neq h\right\}.$$

As consequence of Theorem 1.41, for $\psi \in \mathsf{H}$, $\mathscr{B}(\mathbb{R}^n) \ni E \mapsto (\psi|P^{(\mathbf{A})}(R)\psi)$ defines a positive σ-additive finite measure on $\mathscr{B}(\mathbb{R}^n)$, unique extension of $\mathscr{B}_0(\mathbb{R}^n) \ni S \mapsto (\psi|Q(S)\psi)$. In other words, it is the only positive σ-additive measure ν_ψ on $\mathscr{B}(\mathbb{R}^n)$ such that $\nu_\psi(E_j^{(1)} \times \cdots \times E_j^{(n)}) = (\phi|P^{(A_1)}(E_j^{(1)}) \cdots P^{(A_n)}(E_j^{(n)})\psi)$, for any $E^{(k)} \in \mathscr{B}(\mathbb{R})$. Using the polarisation formula, $\mathscr{B}(\mathbb{R}^n) \ni E \mapsto (\psi|P(R)\phi)$ is, for $\psi, \phi \in \mathsf{H}$, a complex measure on $\mathscr{B}(\mathbb{R}^n)$. Therefore $\mathscr{B}(\mathbb{R}^n) \ni E \mapsto P(R)$ satisfies (a),(b),(c), (d) in Definition 8.41: (a) holds because $P^{(\mathbf{A})}(R)$ is a projector, (c) by construction and (d) by σ-additivity of $\mathscr{B}(\mathbb{R}^n) \ni E \mapsto (\psi|P^{(\mathbf{A})}(R)\phi)$. Eventually, (b) follows from Lemma 9.18. The identity $P^{(\mathbf{A})}(E^{(1)} \times \cdots \times E^{(n)}) = P^{(A_1)}(E^{(1)}) \cdots P^{(A_n)}(E^{(n)})$ implies $P^{(\mathbf{A})}(\Pi_k^{-1}(E^{(k)})) = P^{(A_k)}(E^{(k)})$ for any $E^{(k)} \in \mathscr{B}(\mathbb{R})$, where $\Pi_k : \mathbb{R}^n \to \mathbb{R}$ is the kth canonical projection of $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$. Using Theorem 9.4(h) with $\phi := \Pi_k$ and the spectral Theorem 9.10 for A_k gives

$$\int_{\mathbb{R}^n} f(\Pi_k(x))dP^{(\mathbf{A})}(x) = \int_{\mathbb{R}} f(x_k)dP^{(A_k)}(x_k) = f(A_k)\,, \quad k = 1, 2, \ldots, n.$$

Now, every PVM $P' : \mathscr{B}(\mathbb{R}) \to \mathfrak{P}(\mathsf{H})$ satisfying $P'(E^{(1)} \times \cdots \times E^{(n)}) = P^{(A_1)}(E^{(1)}) \cdots P^{(A_n)}(E^{(n)})$ for any $E^{(k)} \in \mathscr{B}(\mathbb{R})$ must also solve $(\psi|P'(E^{(1)} \times \cdots \times E^{(n)})\psi) = (\psi|P^{(A_1)}(E^{(1)}) \cdots P^{(A_n)}(E^{(n)})\psi)$; as positive measures doing just that are unique, we have $(\psi|P'(R)\psi) = (\psi|P^{(\mathbf{A})}(R)\psi)$ for any $R \in \mathscr{B}(\mathbb{R}^n)$ and any $\psi \in \mathsf{H}$. Therefore $P^{(\mathbf{A})} = P'$, since the previous relation, by polarisation, implies $(\psi|P'(R)\phi) = (\psi|P^{(\mathbf{A})}(R)\phi)$ for $\psi, \phi \in \mathsf{H}$. □

An exhaustive discussion on joint spectral measures, their integrals, and the meaning in QM can be found in [Pru81] and [BeCa81]. In analogy to Theorem 9.9 we could prove what follows ([BeCa81], and Exercise 9.7 for $n = 1$).

Proposition 9.19. *Let $\mathbf{A} = \{A_1, \ldots, A_n\}$ be a collection of self-adjoint operators on the separable Hilbert space H whose spectral measures commute. The von Neumann algebra $\mathbf{A}''$ (the set of operators in $\mathfrak{B}(\mathsf{H})$ commuting with operators in $\mathfrak{B}(\mathsf{H})$ that commute with all spectral measures) coincides with the collection of operators $f(A_1, \ldots, A_n) := \int_{supp(P^{(\mathbf{A})})} f(x_1, \ldots, x_n) dP^{(\mathbf{A})}$ with $f : supp(P^{(\mathbf{A})}) \to \mathbb{C}$ measurable and bounded.*

If f is real-valued, $f(A_1, \ldots, A_n)$ is self-adjoint, and hence interpreted as observable, function of the observables $A_1, \ldots, A_n$ of the quantum system. This corresponds to the notion of Remark 7.47(2).

9.2 Exponential of unbounded operators: analytic vectors

This section is short and technical. We go back to *analytic vectors*, introduced at the end of Chapter 5, and uncover other properties in the light of the theory developed since. The results will be used at various places in the rest of the book.

An interesting general problem is this. If A is a self-adjoint operator on the Hilbert space H, the exponential e^{zA} can be defined as function of A (Definition 9.11). We expect, in some cases, to be able to employ the Taylor expansion:

$$e^{zA} = \sum_{n=0}^{+\infty} \frac{z^n}{n!} A^n ,$$

defining the left-hand side with Definition 9.11. If $A \in \mathfrak{B}(\mathsf{H})$ the above identity does hold, provided we understand the expansion in the uniform topology, as is easy to see (Exercise 8.16). If A is not bounded, the issue is subtler and the above equation makes no sense in the uniform topology. As Nelson clarified, it has a meaning in the strong topology and over a dense subspace in the Hilbert space, which is a core for A: this is the space of analytic vectors for A as we shall prove in Proposition 9.21.

Let A be an operator with domain $D(A)$ on the Hilbert space H. Recall (Definition 5.44) that a vector $\psi \in D(A)$ such that $A^n \psi \in D(A)$ for any $n \in \mathbb{N}$ ($A^0 := I$) is called a C^∞ **vector for** A. The subspace of C^∞ vectors for A is written $C^\infty(A)$. $\psi \in C^\infty(A)$ is an **analytic vector for** A if

$$\sum_{n=0}^{+\infty} \frac{||A^n \psi||}{n!} t^n < +\infty , \quad \text{for some } t > 0. \tag{9.69}$$

Recall also Nelson's Theorem 5.47, for which a symmetric operator on a Hilbert space is essentially self-adjoint if its domain contains analytic vectors whose finite combinations are dense.

Notation 9.20. If A is an operator on H with domain $D(A)$, we shall indicate with $\mathscr{A}(A)$ the subset in $C^\infty(A)$ of elements satisfying (9.69). ■

The next proposition discusses useful properties of analytic vectors, in particular the exponential of (self-adjoint) unbounded operators.

Proposition 9.21. *Let A be an operator on the Hilbert space* H.

(a) *$\mathscr{A}(A)$ is a vector space.*
(b) *If A is closable:*

$$\mathscr{A}(A) \subset \mathscr{A}(\overline{A}) .$$

(c) (i) *For any $c \in \mathbb{C}$, defining $A + cI$ on its standard domain:*

$$\mathscr{A}(A + cI) = \mathscr{A}(A) ;$$

(ii) *for any $c \in \mathbb{C} \setminus \{0\}$, defining cA on its standard domain:*

$$\mathscr{A}(cA) = \mathscr{A}(A) ;$$

(iii) *if A is Hermitian, defining A^2 on its standard domain:*

$$\mathscr{A}(A^2) \subset \mathscr{A}(A) .$$

(d) *If A is self-adjoint and $\psi \in \mathscr{A}(A) \cap D(e^{zA})$, viewing e^{zA} as in Definition 9.11:*

$$e^{zA}\psi = \sum_{n=0}^{+\infty} \frac{z^n}{n!} A^n \psi \quad \textit{for any } z \in \mathbb{C},\ |z| \le t, \textit{ satisfying (9.69) for } \psi . \tag{9.70}$$

If $Re z = 0$*, equation (9.70) holds for any* $\psi \in \mathscr{A}(A)$*, provided* $|z| \leq t$ *solves (9.69) for given* ψ*.*
(e) *If A is self-adjoint, viewing* e^{zA} *as in Definition 9.11:*

$$e^{isA}(\mathscr{A}(A)) \subset \mathscr{A}(A), \quad s \in \mathbb{R}.$$

(f) *If A is self-adjoint,* $D(A)$ *contains a dense subset of* $\mathscr{A}(p(A))$ *for any* $t > 0$ *in (9.69) and any complex polynomial* $p(A)$ *of A.*

Proof. (a) The claim follows from the estimate

$$||A^n(a\psi + b\phi)|| \leq |a|\,||A^n\psi|| + |b|\,||A^n\phi||,$$

$\psi, \phi \in \mathscr{A}(A)$, by choosing $t > 0$ small enough to satisfy (9.69) for ψ and ϕ.
(b) This is a direct consequence of definitions, for $\overline{A}$ is an extension of A and so $\overline{A}^n$ extends A^n.
(c) To prove (i), note that if $t > 0$ satisfies (9.69) for ψ:

$$+\infty > M \geq e^{|tc|} \sum_{k=0}^{+\infty} \frac{t^k\,||A^k\psi||}{k!} = \sum_{p=0}^{+\infty}\sum_{k=0}^{+\infty} \frac{|tc|^p}{p!} \frac{||t^kA^k\psi||}{k!}.$$

By Fubini–Tonelli on the counting product measure of $\mathbb{N}$ we may compute the product of the series (integral in the product measure) as the double integral on the right of M in the chain:

$$M \geq \sum_{n=0}^{+\infty}\sum_{k=0}^{n} \frac{|tc|^{n-k}}{(n-k)!} \frac{||t^kA^k\psi||}{k!} \geq \sum_{n=0}^{+\infty} \frac{t^n}{n!} \left|\left| \sum_{k=0}^{n} \frac{n!c^{n-k}A^k\psi}{k!(n-k)!} \right|\right| = \sum_{n=0}^{+\infty} \frac{t^n}{n!} ||(A+cI)^n\psi||.$$

Therefore $\mathscr{A}(A+cI) \supset \mathscr{A}(A)$. Now define $A' := A + cI$, so $A = A' + c'I$ and $c' = -c$. It follows that $\mathscr{A}(A' + c'I) \supset \mathscr{A}(A')$, which is equivalent to $\mathscr{A}(A) \supset \mathscr{A}(A+cI)$, so $\mathscr{A}(A) = \mathscr{A}(A+cI)$. Property (ii) is obvious by definition, so let us see to (iii). By construction $C^\infty(A) = C^\infty(A^2)$. Since A is Hermitian and $\sqrt{x} \leq 1 + x$ for $x \geq 0$, in $C^\infty(A)$ we have:

$$||A^n\psi|| = \sqrt{(\psi|A^{2n}\psi)} \leq \sqrt{||\psi||}\sqrt{||A^{2n}\psi||} \leq \sqrt{||\psi||}(1 + ||(A^2)^n\psi||).$$

The claim is thus true, since for $t > 0$:

$$\sum_{n=0}^{+\infty} \frac{t^n}{n!}||A^n\psi|| \leq \sqrt{||\psi||}\sum_{n=0}^{+\infty} \frac{t^n}{n!}||(A^2)^n\psi|| + \sqrt{||\psi||}\sum_{n=0}^{+\infty} \frac{t^n}{n!}$$

$$= \sqrt{||\psi||}\left(\sum_{n=0}^{+\infty} \frac{t^n}{n!}||(A^2)^n\psi|| + e^t\right).$$

(d) For some $\phi \in \mathsf{H}$, $\mu_{\phi,\psi}$ is the complex measure $\mu_{\phi,\psi}(E) := \left(\phi|P^{(A)}(E)\psi\right)$, and for any $\chi \in \mathsf{H}$, $\mu_\chi(E) := \left(\chi|P^{(A)}(E)\chi\right)$ is the usual positive finite spectral measure. Using decomposition $d\mu_{\phi,\psi} = hd|\mu_{\phi,\psi}|$ with $|h| = 1$ (Examples 2.45), for $\psi \in D(f(A))$

we have

$$\left|\int_{\mathbb{R}} f\, d|\mu_{\phi,\psi}|\right| = \left|\int_{\mathbb{R}} f\overline{h}\, d\mu_{\phi,\psi}\right| = \left|\left(\phi \left| \int_{\mathbb{R}} f\overline{h}\ dP^{(A)}\psi\right.\right)\right| \leq ||\phi|| \left|\left|\int_{\mathbb{R}} f\overline{h}\, dP^{(A)}\psi\right|\right|$$

$$= ||\phi|| \sqrt{\int_{\mathbb{R}} |f|^2 d\mu_\psi}\,.$$

If $z \in \mathbb{C}$ and $|z| \leq t$ then, using Lemma 9.2 and Theorem 9.4(e):

$$\sum_{n=0}^{+\infty} \int_{\sigma(A)} \left|\frac{z^n}{n!} x^n\right| d|\mu_{\phi,\psi}(x)| = \sum_{n=0}^{+\infty} \left|\frac{z^n}{n!}\right| \int_{\sigma(A)} |x^n| d|\mu_{\phi,\psi}(x)|$$

$$\leq \sum_{n=0}^{+\infty} \frac{t^n}{n!} ||\phi|| \left(\int_{\sigma(A)} x^{2n} d\mu_\psi(x)\right)^{1/2} = \sum_{n=0}^{+\infty} ||\phi|| \frac{t^n}{n!}\, ||A^n\psi|| < +\infty\,,$$

where (9.69) is needed in the last passage. Then Fubini–Tonelli implies, for $|z| \leq t$, we may swap sum and integral:

$$\int_{\sigma(A)} \sum_{n=0}^{+\infty} \frac{z^n}{n!} x^n d\mu_{\phi,\psi}(x)\,.$$

Thus for $|z| \leq t$, if ψ belongs to the domain of e^{zA} (cf. Definition 9.11) and by virtue of Theorem 9.4(e):

$$(\phi|e^{zA}\psi) = \int_{\sigma(A)} e^{zx} d\mu_{\phi,\psi} = \int_{\sigma(A)} \sum_{n=0}^{+\infty} \frac{z^n}{n!} x^n d\mu_{\phi,\psi} = \sum_{n=0}^{+\infty} \frac{z^n}{n!} \int_{\sigma(A)} x^n d\mu_{\phi,\psi} = \sum_{n=0}^{+\infty} \frac{z^n}{n!} (\phi|A^n\psi)\,.$$

By (9.69) the series

$$\sum_{n=0}^{+\infty} \frac{z^n}{n!} A^n\psi$$

converges in H, and the scalar product is continuous, so the above identity reads

$$(\phi|e^{zA}\psi) = \left(\phi \left| \sum_{n=0}^{+\infty} \frac{z^n}{n!} A^n\psi \right.\right)\,.$$

As ϕ is arbitrary, we have (9.70). In case $Re z = 0$, i.e. $z = is$, $s \in \mathbb{R}$, the map $\mathbb{R} \ni x \mapsto e^{isx}$ is clearly bounded, so $e^{isA} \in \mathfrak{B}(\mathsf{H})$ (the domain is H) by Corollary 9.5.
(e) If A is self-adjoint, by Theorem 9.4(c) $e^{isA}(e^{isA})^* = (e^{isA})^* e^{isA} = I$, so e^{isA} is unitary. Using Theorem 9.4(c) taking $\psi \in \mathscr{A}(A) \subset C^\infty(A)$ we obtain $A^n e^{isA}\psi = e^{isA}A^n\psi$, but e^{isA} is unitary and $||e^{isA}A^n\psi|| = ||A^n\psi||$, whence the claim follows.
(f) Consider the spectral decomposition $A = \int_{\mathbb{R}} x dP^{(A)}(x)$, partition the real line $\mathbb{R} = \cup_{n\in\mathbb{Z}}(n, n+1]$ and take its closed, pairwise-orthogonal subspaces $\mathsf{H}_n = P_n(\mathsf{H})$, where we define projectors $P_n := \int_{(n,n+1]} 1 dP^{(A)}(x)$. Choosing a basis $\{\psi_k^{(n)}\}_{k\in K_n} \subset \mathsf{H}_n$ for any n, the union of all bases is a basis of H. Notice $supp(\mu_{\psi_k^{(n)}}) \subset (n, n+1]$ by definition of μ_ϕ (Theorem 8.50). From Theorem 9.4(e) every $\psi_k^{(n)}$ belongs in $D(A)$, since

$\int_{\mathbb{R}} |x|^2 d\mu_{\psi_k^{(n)}}(x) = \int_{(n,n+1]} |x|^2 d\mu_{\psi_k^{(n)}}(x) \leq |n+1|^2$, Moreover (9.69) holds for any $t > 0$, as $||A^m \psi_k^{(n)}||^2 = \int_{(n,n+1]} |x|^{2m} d\mu_{\psi_k^{(n)}}(x) \leq |n+1|^{2m} ||\psi_k^{(n)}||^2$. Finite linear combinations are, by construction, a dense subspace in H, and analytic for A (for any $t > 0$) by (a).

Now take a complex polynomial $p_N(x) = \sum_{k=0}^N x^n$ of degree N, and define $p_N(A)$ on the domain $D(p_N(A)) = D(A^N)$ (Theorem 9.4(d)). We will check every $\psi_k^{(n)}$ is analytic for the closed (self-adjoint if p_N is real) $p_N(A)$ by Theorem 9.4. Choose one of unit norm, say ψ, and suppose its spectral measure μ_ψ has support in some $(-L, L]$. Then $||A^k \psi|| \leq L^k ||\psi|| = L^k$. Therefore

$$||p_N(A)\psi|| = \left|\left| \sum_{k=0}^N a_k A^k \psi \right|\right| \leq \sum_{k=0}^N |a_k| ||A^k \psi|| = \sum_{k=0}^N |a_k| L^k .$$

In a similar manner:

$$||p_N(A)^n \psi|| = \left|\left| \sum_{k_1,\ldots,k_n=0}^N a_{k_1} \cdots a_{k_n} A^{k_1+\cdots+k_n} \psi \right|\right|$$
$$\leq \sum_{k_1,\ldots,k_n=0}^N |a_{k_1}| \cdots |a_{k_n}| ||A^{k_1+\cdots+k_n} \psi|| \leq \sum_{k_1,\ldots,k_n=0}^N |a_{k_1}| \cdots |a_{k_n}| L^{k_1+\cdots+k_n} .$$

We conclude that if $M_L := \sum_{k=0}^N |a_k| L^k$, then

$$||p_N(A)^n \psi|| \leq M_L^n \quad \text{and} \quad \sum_{n=0}^{+\infty} \frac{t^n}{n!} ||p_N(A)^n \psi|| \leq e^{tM_L}$$

and so ψ (by (a), any combination of such vectors) is analytic for $p_N(A)$, for any $t > 0$. □

9.3 Strongly continuous one-parameter unitary groups

The goal of this section is to prove *Stone's theorem*, one of the most important results in view of the applications to QM (and not only that). To state it we will present some preliminary results about one-parameter groups of unitary operators, and in particular an importan theorem due to von Neumann.

9.3.1 Strongly continuous one-parameter unitary groups, von Neumann's theorem

Definition 9.22. *Let* H *be a Hilbert space. A collection* $\{U_t\}_{t\in\mathbb{R}} \subset \mathfrak{B}(\mathsf{H})$ *is called a* **one-parameter group (of operators)** *on* H *if*

$$U_0 = I \quad \textit{and} \quad U_t U_s = U_{t+s} \textit{ for any } t,s \in \mathbb{R} . \tag{9.71}$$

A one-parameter group $\{U_t\}_{t\in\mathbb{R}} \subset \mathfrak{B}(\mathsf{H})$ *is said:*

(a) One-parameter unitary group *if U_t is unitary for any $t \in \mathbb{R}$.*
(b) Weakly continuous at $t_0 \in \mathbb{R}$, *or* **strongly continuous at** $t_0 \in \mathbb{R}$, *if the mapping $t \mapsto U_t$ is continuous at t_0 in the weak, resp. strong, topology, where $\mathbb{R}$ is standard.*
(c) Weakly continuous *or* **strongly continuous** *if it is weakly, or strongly, continuous at each point of $\mathbb{R}$.*

By (9.71), if the U_t are unitary:

$$(U_t)^* = U_t^{-1} = U_{-t}\,, \quad \text{for any } t \in \mathbb{R}\,. \tag{9.72}$$

Proposition 9.23. *Let $\{U_t\}_{t\in\mathbb{R}}$ be a one-parameter unitary group on the Hilbert space $(\mathsf{H}, (\cdot|\cdot))$. The following assertions are equivalent.*

(a) $(\psi|U_t\psi) \to (\psi|\psi)$ *as $t \to 0$ for any $\psi \in \mathsf{H}$.*
(b) $\{U_t\}_{t\in\mathbb{R}}$ *is weakly continuous at $t = 0$.*
(c) $\{U_t\}_{t\in\mathbb{R}}$ *is weakly continuous.*
(d) $\{U_t\}_{t\in\mathbb{R}}$ *is strongly continuous at $t = 0$.*
(e) $\{U_t\}_{t\in\mathbb{R}}$ *is strongly continuous.*

Proof. First, let us number the properties.
(1) $\{U_t\}_{t\in\mathbb{R}}$ is weakly continuous at $t = 0$.
(2) $(\psi|U_t\psi) \to (\psi|\psi)$ as $t \to 0$ for any $\psi \in \mathsf{H}$.
(3) $\{U_t\}_{t\in\mathbb{R}}$ is strongly continuous at $t = 0$.
(4) $\{U_t\}_{t\in\mathbb{R}}$ is strongly continuous.
(5) $\{U_t\}_{t\in\mathbb{R}}$ is weakly continuous.
We will show $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (5) \Rightarrow (1)$.

$(1) \Rightarrow (2)$. Weak continuity at $t = 0$ implies, when $t \to 0$, that $(\psi|U_t\psi) \to (\psi|U_0\psi) = (\psi|\psi)$ and $(U_t\psi|\psi) \to (U_0\psi|\psi) = (\psi|\psi)$ by conjugation.

$(2) \Rightarrow (3)$. Strong continuity at $t = 0$ amounts to saying, for any $\psi \in \mathsf{H}$,

$$||U_t\psi - U_0\psi|| \to 0$$

as $t \to 0$. Since $U_0 = I$, squaring and writing norms via inner products transforms the above into

$$(U_t\psi|U_t\psi) - (\psi|U_t\psi) - (U_t\psi|\psi) + (\psi|\psi) \to 0\,.$$

U_t unitary implies $(U_t\psi|U_t\psi) = (\psi,\psi)$, so the identity reads

$$(\psi|\psi) - (\psi|U_t\psi) - (U_t\psi|\psi) + (\psi,\psi) \to 0\,, \quad \text{as } t \to 0\,.$$

But as we said at the beginning, the latter holds if (2) does.

$(3) \Rightarrow (4)$. If $\psi \in \mathsf{H}$:

$$U_t\psi - U_{t_0}\psi = U_t(\psi - U_t^{-1}U_{t_0}\psi) = U_t(\psi - U_{t_0-t}\psi)\,,$$

where (9.72) was used. As U_t is unitary, for any $\psi \in \mathsf{H}$:

$$||U_s\psi - U_{t_0}\psi|| = ||U_s(\psi - U_{t_0-s}\psi)|| = ||\psi - U_{t_0-s}\psi||\,.$$

Under strong continuity at $t = 0$, since $t_0 - s \to 0$ for $s \to t_0$, we find $||U_s\psi - U_{t_0}\psi|| \to 0$. Hence strong continuity at $t = 0$ forces strong continuity at any $t_0 \in \mathbb{R}$.

(4) $\Rightarrow$ (5). Obvious because strong convergence implies weak convergence.

(5) $\Rightarrow$ (1). True by definition. $\square$

Here is another property of unitary groups.

Proposition 9.24. *Let $\{U_t\}_{t\in\mathbb{R}}$ be a one-parameter unitary group on the Hilbert space $(\mathsf{H}, (\cdot|\cdot))$, and $\mathscr{H} \subset \mathsf{H}$ a subset such that:*

(a) *The finitely-generated span $< \mathscr{H} >$ is dense in H.*
(b) *$\{U_t\}_{t\in\mathbb{R}}$ satisfies $(\psi|U_t\psi) \to (\psi|\psi)$, as $t \to 0$, for any $\psi \in \mathscr{H}$.*
Then $\{U_t\}_{t\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group.

Proof. The same argument used in Proposition 9.23 gives that $(\phi_0|U_t\phi_0) \to (\phi_0|\phi_0)$, as $t \to 0$, for $\phi_0 \in \mathscr{H}$ implies $||U_t\phi_0 - \phi_0|| \to 0, t \to 0$. If, more generally, $\phi \in < \mathscr{H} >$ then $\phi = \sum_{i\in I} c_i\phi_{0i}$ where I is finite and $\phi_{0i} \in \mathscr{H}$. Hence as $t \to 0$

$$||U_t\phi - \phi|| = \left|\left|U_t\left(\sum_i c_i\phi_{0i}\right) - \sum_i c_i\phi_{0i}\right|\right| = \left|\left|\sum_i c_i(U_t\phi_{0i} - \phi_{0i})\right|\right|$$

$$\leq \sum_i |c_i|||U_t\phi_{0i} - \phi_{0i}|| \to 0\,.$$

By Proposition 9.23 it now suffices to extend this to H. That is to say, $||U_t\phi - \phi|| \to 0$, as $t \to 0$, for any $\phi \in < \mathscr{H} >$ implies $||U_t\psi - \psi|| \to 0$, $t \to 0$, for any $\psi \in \mathsf{H}$. As $< \mathscr{H} >$ is dense, for any given $\psi \in \mathsf{H}$ there is a sequence $\{\psi_n\}_{n\in\mathbb{N}} \subset \langle\mathscr{H}\rangle$ with $\psi_n \to \phi$, $n \to +\infty$. If $\{t_m\}_{m\in\mathbb{N}}$ is a real infinitesimal sequence, by the triangle inequality

$$||U_{t_m}\psi - \psi|| \leq ||U_{t_m}\psi - U_{t_m}\phi_n|| + ||U_{t_m}\phi_n - \phi_n|| + ||\phi_n - \psi||$$

for any given $n \in \mathbb{N}$. Since the U_{t_m} are unitary and the norm non-negative, that means

$$0 \leq ||U_{t_m}\psi - \psi|| \leq ||U_{t_m}\phi_n - \phi_n|| + 2||\phi_n - \psi||\,. \tag{9.73}$$

For fixed n, $||U_{t_m}\phi_n - \phi_n|| \to 0$, $m \to +\infty$, by assumption, so:

$$\limsup_m ||U_{t_m}\phi_n - \phi_n|| = \liminf_m ||U_{t_m}\phi_n - \phi_n|| = \lim_{m\to+\infty} ||U_{t_m}\phi_n - \phi_n|| = 0\,.$$

By (9.73), for any $n \in \mathbb{N}$:

$$0 \leq \limsup_m ||U_{t_m}\psi - \psi|| \leq 2||\phi_n - \psi||\,, \quad 0 \leq \liminf_m ||U_{t_m}\psi - \psi|| \leq 2||\phi_n - \psi||\,.$$

On the other hand for n big enough we can make $||\phi_n - \psi||$ infinitesimal, so:

$$\limsup_m ||U_{t_m}\psi - \psi|| = \liminf_m ||U_{t_m}\psi - \psi|| = 0\,.$$

Therefore

$$\lim_{m\to+\infty} ||U_{t_m}\psi - \psi|| = 0\,.$$

As $\psi \in \mathsf{H}$ and the $\{t_m\}_{m\in\mathbb{N}}$ are arbitrary, for any $\psi \in \mathsf{H}$ we have:

$$\lim_{t\to 0} ||U_t\psi - \psi|| = 0\,,$$

ending the proof. □

The theory developed thus far puts us in the position to prove an importan result due to von Neumann, which shows how strong continuity of one-parameter unitary groups is, actually, not such restrictive a fact in separable Hilbert spaces.

Theorem 9.25 (Von Neumann). *Let $\{U_t\}_{t\in\mathbb{R}}$ be a one-parameter unitary group on the Hilbert space $(\mathsf{H}, (\cdot|\cdot))$. If H is separable, $\{U_t\}_{t\in\mathbb{R}}$ is strongly continuous iff the map $\mathbb{R} \ni t \mapsto (U_t\psi|\phi)$ is measurable for any $\psi, \phi \in \mathsf{H}$.*

Proof. Obviously if the group is strongly continuous then any $\mathbb{R} \ni t \mapsto (U_t\psi|\phi)$ is measurable, being continuous. We show the converse. Suppose every such map is Borel measurable, hence Lebesgue measurable. By Schwarz's inequality and $||U_t|| = 1$ follows that these maps are bounded. Given $a \in \mathbb{R}$, $\psi \in \mathsf{H}$,

$$\mathsf{H} \ni \phi \mapsto \int_0^a (U_t\psi|\phi)dt$$

is a bounded linear functional with norm not exceeding $|a|\,||\psi||$ by Schwarz and $||U_t|| = 1$. Riesz's Theorem 3.16 provides $\psi_a \in \mathsf{H}$ such that

$$(\psi_a|\phi) = \int_0^a (U_t\psi|\phi)dt\,, \quad \text{for any } \phi \in \mathsf{H}.$$

So

$$(U_b\psi_a|\phi) = (\psi_a|U_{-b}\phi) = \int_0^a (U_t\psi|U_{-b}\phi)dt = \int_0^a (U_{t+b}\psi|\phi)dt = \int_b^{a+b} (U_t\psi|\phi)dt\,.$$

Splitting the integral in the obvious manner:

$$|(U_b\psi_a|\phi) - (\psi_a|\phi)| = \left|\int_b^{a+b} (U_t\psi|\phi)dt - \int_0^a (U_t\psi|\phi)dt\right|$$

$$\leq \left|\int_b^0 (U_t\psi|\phi)dt\right| + \left|\int_a^{a+b} (U_t\psi|\phi)dt\right| \leq 2b||\phi||\,||\psi||.$$

Then $(U_b\psi_a|\phi) \to (\psi_a|\phi)$, as $b \to 0$, and so by conjugation:

$$\lim_{t\to 0}(\phi\,|\,U_b\psi_a) \to (\phi\,|\,\psi_a)\,.$$

We are done if we can prove that the set $\{\psi_a\,|\,\psi \in \mathsf{H}, a \in \mathbb{R}\}$ finitely generates a dense space in H, by the previous proposition and choosing $\phi = \psi_a$. Let $\phi \in \{\psi_a\,|\,\psi \in \mathsf{H}, a \in$

$\mathbb{R}\}^\perp$ and $\{\psi^{(n)}\}_{n\in\mathbb{N}}$ be a countable basis for H, which we have by separability (the finite-dimensional case is the same). For any $n \in \mathbb{N}$:

$$0 = (\psi_a^{(n)}|\phi) = \int_0^a (U_t\psi^{(n)}|\phi)dt\,, \quad a \in \mathbb{R}\,,$$

implying (Theorem 1.75(b)) $\mathbb{R} \ni t \mapsto (U_t\psi^{(n)}|\phi)$ is null almost everywhere. Call $S_n \subset \mathbb{R}$ the set where the map *does not* vanish, and fix $t_0 \in \mathbb{R} \setminus \bigcup_{n\in\mathbb{N}} S_n$. The latter exists for $\bigcup_{n\in\mathbb{N}} S_n$ cannot coincide $\mathbb{R}$: the former, in fact, has zero measure as countable union of zero-measure sets. (This is the point where we need the basis to be countable, i.e. separability.) Then $(U_{t_0}\psi^{(n)}|\phi) = 0$ for any n, forcing $\phi = 0$ because U_{t_0} is unitary and $\{U_{t_0}\psi^{(n)}\}_{n\in\mathbb{N}}$ is a basis. Since $\{\psi_a \mid \psi \in \mathsf{H},\, a \in \mathbb{R}\}^\perp = \{0\}$, the span of $\{\psi_a \mid \psi \in \mathsf{H},\, a \in \mathbb{R}\}$ is dense, as required, and the theorem is proved. □

Remarks 9.26. In the statement we may substitute Borel measurability with measurability for the Lebesgue σ-algebra. If Lebesgue measurability holds, in fact, the proof does not change and so the group is strongly continuous. Under strong continuity Borel measurability is granted, so also Lebesgue measurability. ■

9.3.2 One-parameter unitary groups generated by self-adjoint operators and Stone's theorem

This section contains the celebrated *Stone's theorem*, that describes strongly continuous one-parameter unitary groups obtained by exponentiating self-adjoint operators. Later we will use these groups to provide a necessary and sufficient condition for the spectral measures of self-adjoint operators to commute.

Before all this we need a technical result, which we state separately given its usefulness in many contexts. As usual, dx is Lebesgue's measure on $\mathbb{R}^n$ and $\chi_{[a,b]}$ the characteristic function of $[a,b]$.

Proposition 9.27. *Let* H *be a complex Hilbert space and* $\{V_t\}_{t\in\mathbb{R}^n} \subset \mathfrak{B}(\mathsf{H})$ *a family of operators satisfying either one of the following conditions:*

(i) $s\text{-}\lim_{t\to t_0} V_t = V_{t_0}$*, for any* $t_0 \in \mathbb{R}^n$*;*
(ii) *there exists* $C \geq 0$ *such that* $||V_t|| \leq C$ *for any* $t \in \mathbb{R}^n$*.*

Then for any $f \in L^1(\mathbb{R}^n, dx)$ *there is a unique operator on* $\mathfrak{B}(\mathsf{H})$*, denoted* $\int_{\mathbb{R}^n} f(t)V_t\,dt$*, such that:*

$$\left(\phi \left| \int_{\mathbb{R}^n} f(t)V_t\,dt\;\psi\right.\right) = \int_{\mathbb{R}^n} f(t)\,(\phi|V_t\psi)\,dt\,, \quad \phi,\psi \in \mathsf{H}. \tag{9.74}$$

If $f \in L^1(\mathbb{R}^n, dx)$ *has compact (essential) support, condition (i) is enough to guarantee the existence of* $\int_{\mathbb{R}^n} f(t)V_t dt$*.*

The latter satisfies:

(a) *for any* $\psi \in \mathsf{H}$*:*

$$\left|\left|\int_{\mathbb{R}^n} f(t)V_t\, dt\psi\right|\right| \leq \int_{\mathbb{R}^n} |f(t)|\,||V_t\psi||\, dt\,. \tag{9.75}$$

(b) *If* $A \in \mathfrak{B}(\mathsf{H})$*:*

$$A\int_{\mathbb{R}^n} f(t)V_t dt = \int_{\mathbb{R}^n} f(t)AV_t dt \quad \textit{and} \quad \int_{\mathbb{R}^n} f(t)V_t dtA = \int_{\mathbb{R}^n} f(t)V_t A dt\,. \tag{9.76}$$

(c) *Let, for* $n = 1$, $\int_s^t f(\tau)V_\tau d\tau := \int_{\mathbb{R}} g(\tau)f(\tau)V_\tau d\tau$ *where* $g = \chi_{[s,t]}$ *if* $s \geq t$ *and* $g = -\chi_{[t,s]}$ *if* $t \leq s$*. Then*

(i) $\mathbb{R}^2 \ni (s,t) \mapsto \int_s^t f(\tau)V_\tau d\tau$ *is continuous in the uniform topology;*

(ii) f *continuous implies* $\displaystyle s\text{-}\frac{d}{dt}\int_s^t f(\tau)V_\tau d\tau = f(t)V_t \quad \forall s,t \in \mathbb{R}\,.$ (9.77)

Proof. Let $\psi, \phi \in \mathsf{H}$ and $f : \mathbb{R}^n \to \mathbb{C}$ be a map in $L^1(\mathbb{R}^n, dx)$. Consider the integral

$$I(\phi,\psi) := \int_{\mathbb{R}^n} f(t)\,(\phi|V_t\psi)\, dt\,.$$

It is well defined as $\mathbb{R}^n \ni t \mapsto (\phi|V_t\psi)$ is continuous, since $\{V_t\}_{t\in\mathbb{R}^n}$ is weakly continuous, and bounded by (ii) from Schwarz's inequality. Hence

$$|I(\phi,\psi)| \leq ||f||_1 C||\psi||\,||\phi||\,.$$

Since $\mathsf{H} \ni \psi \mapsto I(\phi,\psi)$ is linear and we have the above inequality, Riesz's theorem gives, for any $\phi \in \mathsf{H}$, a unique $\Phi_\phi \in \mathsf{H}$ such that

$$I(\phi,\psi) = (\Phi_\phi|\psi)\,, \quad \text{for any } \psi \in \mathsf{H}.$$

The map $\mathsf{H} \ni \phi \mapsto T\phi := \Phi_\phi$ is linear, and by construction

$$|(\psi|T\phi)| = |(T\phi|\psi)| = |(\Phi_\phi|\psi)| = |I(\phi,\psi)| \leq ||f||_1 C||\psi||\,||\phi||\,, \quad \text{with } \phi,\psi \in \mathsf{H}.$$

Choosing $\psi = T\phi$ shows T, and hence its adjoint $\int_{\mathbb{R}^n} f(t)V_t\, dt$, are bounded. By construction (9.74) holds, and the argument ensures uniqueness. From (9.74) follows

$$\left|\left(\phi\left|\int_{\mathbb{R}^n} f(t)V_t\, dt\ \psi\right.\right)\right| \leq \int_{\mathbb{R}^n} |f(t)|\,|(\phi|V_t\psi)|\, dt \leq \int_{\mathbb{R}^n} |f(t)|\,||V_t\psi||\, dt\, ||\phi||,$$

and taking $\phi = \int_{\mathbb{R}^n} f(t)V_t\, dt\psi$ leads to (9.75). Identity (9.76) follows from (9.74). In case the essential support of f is in a compact set K we can equivalently define $I(\psi,\phi)$ integrating on it and then proceeding as before. In such a case the constant C of (ii) ($t \in K$) automatically exists. By continuity, in fact, whichever $\psi \in \mathsf{H}$ we take there is $C_\psi \geq 0$ such that $||V_t\psi|| \leq C_\psi$ if $t \in K$. By Banach–Steinhaus this implies $C \geq 0$ exists with $||V_t|| \leq C$ if $t \in K$. So let us prove (c). Let $[a,b]$ be big enough so that $[a,b]\times[a,b]$ contains an open neigihbourhood of (t,s), to which (t',s') belongs. From (a) we have

$$\left|\left|\int_t^s f(\tau)V(\tau)d\tau\psi - \int_{t'}^{s'} f(\tau)V(\tau)d\tau\psi\right|\right| \leq (|t-t'|+|s-s'|)\sup_{\tau\in[a,b]}|f(\tau)|\sup_{\tau\in[a,b]}||V_\tau\psi||\,.$$

Since $||V_\tau|| \leq C < +\infty$ for $\tau \in [a,b]$, and as $\int_t^s - \int_{t'}^{s'} = \int_t^{t'} + \int_{t'}^{s} - \int_{t'}^{s'} = \int_t^{t'} + \int_s^{s'}$, taking the least upper bound over $||\psi|| \leq 1$ produces

$$\left|\left|\int_t^s f(\tau)V(\tau)d\tau - \int_{t'}^{s'} f(\tau)V(\tau)d\tau\right|\right| \leq (|t-t'|+|s-s'|)\sup_{\tau\in[a,b]}|f(\tau)|C\,,$$

whence continuity in uniform topology. As for the second property, by strong continuity of $t \mapsto f(t)V_t$, as $h \to 0$, we have

$$\left|\left|\frac{1}{h}\int_\tau^{\tau+h} f(t)V_t dt\,\psi - f(\tau)V_\tau\psi\right|\right| = \left|\left|\frac{1}{h}\left[\int_\tau^{\tau+h}(f(t)V_t - f(\tau)V_\tau)dt\right]\psi\right|\right|$$

$$\leq \frac{\left|\int_\tau^{\tau+h} dt\right|}{|h|}\sup_{|t'-\tau|\leq h}||f(t')V_{t'}\psi - f(\tau)V_\tau\psi|| = \sup_{|t'-\tau|\leq h}||f(t')V_{t'}\psi - f(\tau)V_\tau\psi|| \to 0\,.$$

□

Remarks 9.28. As exercise the reader might prove **Stone's formula**, valid for a self-adjoint operator $T : D(T) \to \mathsf{H}$ with spectral measure $P^{(T)}$:

$$\frac{1}{2}(P^{(T)}(\{a\}) + P^{(T)}(\{b\})) + P^{(T)}((a,b)) = s\text{-}\lim_{\varepsilon\to 0^+}\frac{1}{2\pi i}\int_a^b \frac{1}{T-\lambda-i\varepsilon} - \frac{1}{T-\lambda+i\varepsilon}d\lambda\,.$$

The integral is understood in the sense of Proposition 9.27. ■

It is time to pass to *Stone's theorem*. This name actually refers to assertion (b), the only non-elementary statement.

Theorem 9.29 (Stone). *Let* H *be a Hilbert space.*

(a) *If* $A : D(A) \to \mathsf{H}$*,* $D(A)$ *dense in* H*, is a self-adjoint operator and* $P^{(A)}$ *its spectral measure, then the operators*

$$U_t = e^{itA} := \int_{\sigma(A)} e^{i\lambda t}\,dP^{(A)}(\lambda)\,, \quad t \in \mathbb{R}\,,$$

form a strongly continuous one-parameter unitary group. Moreover:

(i) *the limit*

$$s\text{-}\left.\frac{dU_t}{dt}\right|_{t=0} \psi := \lim_{t\to 0} \frac{U_t\psi - \psi}{t} \tag{9.78}$$

exists in $\mathsf{H} \Leftrightarrow \psi \in D(A)$*;*

(ii) *if* $\psi \in D(A)$*:*

$$s\text{-}\left.\frac{dU_t}{dt}\right|_{t=0} \psi = iA\psi\,. \tag{9.79}$$

(b) *If* $\{U_t\}_{t\in\mathbb{R}}$ *is a strongly continuous one-parameter unitary group on* H*, there exists a unique self-adjoint operator* $A : D(A) \to \mathsf{H}$ *(with* $D(A)$ *dense in* H*) such that*

$$e^{itA} = U_t\,, \quad \textit{for any } t \in \mathbb{R}. \tag{9.80}$$

Proof. (a) If $t \in \mathbb{R}$, $\mathbb{R} \ni \lambda \mapsto e^{it\lambda}$ is trivially bounded, so $e^{itA} \in \mathfrak{B}(\mathsf{H})$ by Corollary 9.5. Theorem 9.4(c) implies ($t \in \mathbb{R}$) $e^{itA}(e^{itA})^* = (e^{itA})^* e^{itA} = I$, making e^{itA} unitary. To prove strong continuity it is enough to check $(\psi|U_t\psi) \to (\psi|\psi)$ for any $\psi \in \mathsf{H}$ as $t \to 0$, by Proposition 9.23. This is true, by Theorem 9.4(f) and since the domain of e^{itA} is all of H, because:

$$(\psi|U_t\psi) = \int_{\sigma(A)} e^{it\lambda} d\mu_\psi(\lambda) \to \int_{\sigma(A)} 1 d\mu_\psi(\lambda) = (\psi|\psi) \quad \text{as } t \to 0\,.$$

We used that $e^{it\lambda} \to 1$ and so Lebesgue's dominated convergence applies, as $|e^{it\lambda}| = 1$ for any t and the constant 1 is integrable as μ_ψ is finite.

Let us prove (i)–(ii). If $\psi \in D(A)$, from Theorem 9.4(c) we compute

$$\left|\left|\frac{U_t - I}{t}\psi - iA\psi\right|\right|^2 = \int_{\sigma(A)} \left|\frac{e^{i\lambda t} - 1}{t} - i\lambda\right|^2 d\mu_\psi(\lambda)\,. \tag{9.81}$$

On the other hand $|e^{i\lambda t} - 1| = 2|\sin(\lambda t/2)| \le |\lambda t|$, so

$$\left|\frac{e^{i\lambda t} - 1}{t} - i\lambda\right|^2 \le 4|\lambda|^2\,.$$

The map $\mathbb{R} \ni \lambda \mapsto |\lambda|^2$ is integrable in μ_ψ by definition of $D(A) \ni \psi$. At last,

$$\left|\frac{e^{i\lambda t} - 1}{t} - i\lambda\right|^2 \to 0 \quad \text{as } t \to 0, \text{ for any } \lambda \in \mathbb{R}.$$

The dominated convergence theorem on the right side of (9.81) gives

$$\left|\left|\frac{U_t - I}{t}\psi - iA\psi\right|\right| \to 0 \quad \text{as } t \to 0, \text{ for any } \psi \in D(A).$$

To finish we show that $\frac{U_t\psi-\psi}{t} \to \phi_\psi \in \mathsf{H}$, $t \to 0$, implies $\psi \in D(A)$. The set of $\psi \in \mathsf{H}$ for which the limit exists is a subspace $D(B)$ in H containing $D(A)$, and as such is dense. The mapping $\psi \mapsto iB\psi := \phi_\psi$ defines an operator with dense domain $D(B)$. If $\psi, \psi' \in D(B)$, using $U_t^* = U_{-t}$:

$$(\psi|B\psi') = \left(\psi\left|-i\lim_{t\to 0}\frac{U_t\psi'-\psi'}{t}\right.\right) = -i\lim_{t\to 0}\left(\psi\left|\frac{U_t\psi'-\psi'}{t}\right.\right)$$

$$= -i\lim_{t\to 0}\left(\left.\frac{U_{-t}\psi-\psi}{t}\right|\psi'\right) = \left(-i\lim_{t\to 0}\left.\frac{U_{-t}\psi-\psi}{-t}\right|\psi'\right) = (B\psi|\psi')\ .$$

Hence B is a symmetric extension of A. But A is self-adjoint, so $B = A$ by Proposition 5.17(d); thus any vector ψ for which the limit of $\frac{U_t\psi-\psi}{t}$ exists as $t \to 0$ lives in $D(A)$. This concludes part (a).

(b) A's uniqueness is immediate. If there were two self-adjoint operators A, A' with $e^{itA} = U_t = e^{itA'}$ for any $t \in \mathbb{R}$, (i)–(ii) in (a) would force $A = A'$. Let us manifacture a self-adjoint operator A satisfying $U_t = e^{itA}$ for a given strongly continuous one-parameter unitary group. Specialise Proposition 9.27 to a strongly continuous one-parameter unitary group $V_t = U_t$. Call $\mathscr{D}$ the space of vectors of the form $\int_{\mathbb{R}} f(t)U_t\,dt\phi$, $\phi \in \mathsf{H}$, with arbitrary $f \in \mathscr{D}(\mathbb{R})$ (smooth complex functions on $\mathbb{R}$ with compact support). This vector space $\mathscr{D}$ is called **Gårding space**. Equation (9.74) easily implies the invariance $U_s\mathscr{D} \subset \mathscr{D}$ for any $s \in \mathbb{R}$, i.e.

$$U_s\int_{\mathbb{R}} f(t)U_t\,dt\psi = \int_{\mathbb{R}} f(t)U_{t+s}\,dt\psi = \int_{\mathbb{R}} f(t-s)U_t dt\psi \quad \text{for any } \psi \in \mathsf{H}. \tag{9.82}$$

Let us show, if $\psi \in \mathscr{D}$, that $\frac{U_t\psi-\psi}{t} \to \psi_0 \in \mathsf{H}$ as $t \to 0$. Suppose $\psi = \int_{\mathbb{R}} f(t)U_t\,dt\phi$. A few computations involving (9.82) and the definition of $\int_{\mathbb{R}} f(t)U_t\,dt\phi$, yield

$$\left|\left|\frac{U_t\psi-\psi}{t} - \int_{\mathbb{R}} f'(s)U_s\,ds\phi\right|\right|^2$$

$$= \left(\int_{\mathbb{R}}\left(\frac{f(s-t)-f(s)}{t} - f'(s)\right)U_s\,ds\phi\left|\int_{\mathbb{R}}\left(\frac{f(r-t)-f(r)}{t} - f'(r)\right)U_r\,dr\phi\right.\right)$$

$$= \int_{\mathbb{R}} ds\int_{\mathbb{R}} dr\overline{h_t(s)}h_t(r)\,(\phi|U_{r-s}\phi)\ ,$$

where

$$h_t(s) := \frac{f(s-t)-f(s)}{t} - f'(s)\,.$$

For any $t \in \mathbb{R}$, the function $s \mapsto h_t(s)$ has support contained in a compact set and is C^∞ (hence bounded). As $(r,s) \mapsto (\phi|U_{r-s}\phi)$ is also bounded, we may interpret the integral using the product Lebesgue measure:

$$\left|\left|\frac{U_t\psi-\psi}{t} - \int_{\mathbb{R}} f'(t)U_t\,dt\phi\right|\right|^2 = \int_{\mathbb{R}\times\mathbb{R}} dsdr\overline{h_t(s)}h_t(r)\,(\phi|U_{r-s}\phi)\ . \tag{9.83}$$

Now: the integrand is pointwise infinitesimal as $t \to 0$, the maps

$$(s,r) \mapsto \overline{h_t(s)} h_t(r) \, (\phi | U_{r-s} \phi)$$

all have support in one large-enough compact set if t varies in a bounded interval around 0, and they are, there, uniformly bounded by some constant not depending on t ($(t,s,r) \mapsto \overline{h_t(s)} h_t(r) \, (\phi | U_{r-s} \phi)$ is jointly continuous in its variables). By all this we apply dominated convergence obtaining that both sides in (9.83) vanish as $t \to 0$. Therefore, for $\psi \in \mathscr{D}$ we have proven $\frac{U_t \psi - \psi}{t} \to \psi_0 \in \mathsf{H}$ as $t \to 0$. The map $\psi \mapsto iS\psi := \psi_0$ is clearly linear. Continuing as in part (a) one can see S is Hermitian. As a matter of fact S is symmetric since $\mathscr{D}$ is dense, which is what we prove next. Given $\phi \in \mathsf{H}$ consider the sequence of $\int_{\mathbb{R}} f_n(t) U_t \, dt \phi$, where $f_n \in \mathscr{D}(\mathbb{R})$ satisfy $f_n \geq 0$, supp $f_n \subset [-1/n, 1/n]$ and $\int_{\mathbb{R}} f_n(s) ds = 1$. Then

$$\left|\left| \int_{\mathbb{R}} f_n U_t dt \psi - \psi \right|\right| = \left|\left| \int_{\mathbb{R}} f_n U_t dt \psi - \int_{\mathbb{R}} f_n dt \psi \right|\right| = \left|\left| \int_{\mathbb{R}} f_n (U_t - I) dt \psi \right|\right|$$

$$\leq \int_{\mathbb{R}} |f_n(t)| \, ||(U_t - I)\psi|| \, dt$$

where we used (9.75) on $V_t = U_t - I$. Since

$$\int_{\mathbb{R}} |f_n(t)| \, ||(U_t - I)\psi|| \, dt \leq \int_{-1/n}^{1/n} |f_n(t)| \, dt \sup_{t \in [-1/n, 1/n]} ||(U_t - I)\psi||$$

$$= \sup_{t \in [-1/n, 1/n]} ||(U_t - I)\psi||$$

and $\sup_{t \in [-1/n, 1/n]} ||(U_t - I)\psi|| \to 0$ as $n \to \infty$, the U_t being strongly continuous, we conclude

$$\mathscr{D} \ni \int_{\mathbb{R}} f_n(t) U_t \, dt \phi \to \phi \in \mathsf{H}, \quad n \to \infty.$$

Hence $\mathscr{D}$ is dense in H and S is symmetric. Now we prove it is essentially self-adjoint on $\mathscr{D}$. If $\psi_\pm \in Ran(S \pm iI)^\perp$, then for any $\chi \in \mathscr{D}$ (recall $U_t \mathscr{D} \subset \mathscr{D}$):

$$\frac{d}{dt} (\psi_\pm | U_t \chi) = \lim_{h \to 0} \left(\psi_\pm \left| \frac{U_h U_t \chi - U_t \chi}{h} \right. \right) = (\psi_\pm | iSU_t \chi)$$

$$= i (\psi_\pm | (S \pm iI) U_t \chi) \pm (\psi_\pm | U_t \chi) = \pm (\psi_\pm | U_t \chi)$$

and $F_\pm(t) := (\psi_\pm | U_t \chi)$ is of the form $F_\pm(0) e^{\pm t}$. If we want it bounded ($||U_t|| = 1$ for any $t \in \mathbb{R}$), necessarily $F_\pm(0) = 0$ and $\psi_\pm = 0$, in turn implying $\overline{Ran(S \pm iI)} = \mathsf{H}$. By Theorem 5.19 that means $S : \mathscr{D} \to \mathsf{H}$ is essentially self-adjoint. Now let $\overline{S}$ be the self-adjoint extension of S. To finish observe that if $V_t := e^{it\overline{S}}$, for any $\psi, \phi \in \mathscr{D}$:

$$\frac{d}{dt} (\psi | (V_t)^* U_t \, \phi) = \frac{d}{dt} (V_t \psi | U_t \phi) = (iSV_t \psi | U_t \phi) + (V_t \psi | iSU_t \phi)$$

$$= -(V_t \psi | iSU_t \phi) + (V_t \psi | iSU_t \phi) = 0 \, .$$

Thus $(\psi | (V_t)^* U_t \phi) = (\psi | I \phi)$. As $\mathscr{D}$ is dense, $(V_t)^* U_t = I$, i.e. $U_t = e^{it\overline{S}}$ for any $t \in \mathbb{R}$. □

Corollary 9.30. *If A is self-adjoint on the Hilbert space H and $\mathscr{D}_0 \subset D(A)$ is dense such that $e^{itA}\mathscr{D}_0 \subset \mathscr{D}_0$ for any $t \in \mathbb{R}$, then $A\restriction_{\mathscr{D}_0}$ is essentially self-adjoint, i.e. $\mathscr{D}_0$ is a core of A.*

Proof. Take $\psi \in \mathscr{D}_0 \subset D(A)$. Then $U_t\psi = e^{itA}\psi$ is differentiable and its derivative is $iAU_t\psi$. Going through the final part of Stone's proof and replacing Gårding's space $\mathscr{D}$ with $\mathscr{D}_0$ proves the claim. □

Now comes a related technical, and useful, elementary result.

Proposition 9.31. *Let A be self-adjoint on the Hilbert space H, and define $U_t := e^{itA}$, $t \in \mathbb{R}$. For any measurable $f : \sigma(A) \to \mathbb{C}$:*

$$U_t f(A)\psi = f(A)U_t\psi\,, \quad \forall t \in \mathbb{R}\; \forall \psi \in D(f(A))\,. \tag{9.84}$$

Proof. $\psi \in D(f(A))$ if and only if $\int_{\sigma(A)} |f(\lambda)|^2 d\mu_\psi(\lambda) < +\infty$. On the other hand the measures μ_ψ and $\mu_{U_t\psi}$ are the same, since

$$(U_t\psi|P^{(A)}(E)U_t\psi) = (\psi|U_t^*P^{(A)}(E)U_t\psi)\,,$$

but $U_t^*P^{(A)}(E)U_t = P^{(A)}(E)$ from (9.13)–(9.14) in Theorem 9.4(c) (recall all integrals refer to bounded maps so the operators are defined on the entire space). In conclusion $\psi \in D(f(A)) \;\Leftrightarrow\; U_t\psi \in D(f(A))$. Conversely $f(A)\psi \in D(U_t) = \mathsf{H}$ holds trivially, since U_t is unitary. With this, using (9.13)–(9.14) in Theorem 9.4(c), we get $U_t f(A)\psi = f(A)U_t\psi$ for any $\psi \in D(f(A))$, i.e. (9.84). □

The next definition will be fundamental for physical applications, as we will see in Chapter 12 and 13.

Definition 9.32. *Let H be a Hilbert space, $\{U_t\}_{t\in\mathbb{R}} \subset \mathfrak{B}(\mathsf{H})$ a strongly continuous one-parameter unitary group. The unique self-adjoint operator A on H fulfilling (9.80) is called* **(self-adjoint) generator of** $\{U_t\}_{t\in\mathbb{R}} \subset \mathfrak{B}(\mathsf{H})$.

In general the self-adjoint generator A is unbounded. It is bounded – and so defined on all H – precisely when $\{U_t\}_{t\in\mathbb{R}}$ is continuous at $t = 0$ (and almost everywhere) in the uniform topology. See Exercise 9.8.

Stone's theorem has a host of useful corollaries, and here is one.

Corollary 9.33. *If $A : D(A) \to \mathsf{H}$, $D(A)$ dense in H, is self-adjoint (in general unbounded) on the Hilbert space H, and $U : \mathsf{H} \to \mathsf{H}_1$ is an isomorphism (surjective isometry), then*

$$Ue^{isA}U^{-1} = e^{isUAU^{-1}}\,, \quad s \in \mathbb{R}\,.$$

The same holds, in particular, when $\mathsf{H} = \mathsf{H}_1$ and U is unitary.

Proof. The operator UAU^{-1} is clearly self-adjoint on $UD(A)$ by definition. Hence the strongly continuous one-parameter unitary group $\{e^{isUAU^{-1}}\}_{s\in\mathbb{R}}$ is well defined.

As U is an isomorphism, $\{Ue^{isA}U^{-1}\}_{s\in\mathbb{R}}$ too is a strongly continuous one-parameter unitary group on H_1. Furthermore, if $\psi = U^{-1}\phi \in U^{-1}D(A)$ then

$$\lim_{s\to 0}\frac{Ue^{isA}U^{-1}\psi-\psi}{s} = \lim_{s\to 0}\frac{Ue^{isA}\phi-U\phi}{s} = U\lim_{s\to 0}\frac{e^{isA}\phi-\phi}{s} = iUA\phi = UAU^{-1}\psi\,.$$

By Stone's theorem the generator of $\{Ue^{isA}U^{-1}\}_{s\in\mathbb{R}}$ is a self-adjoint extension of UAU^{-1}; but the latter is already self-adjoint, so the generator of $\{Ue^{isA}U^{-1}\}_{s\in\mathbb{R}}$ is UAU^{-1} itself, and

$$Ue^{isA}U^{-1} = e^{isUAU^{-1}}\,, \quad s\in\mathbb{R}\,.$$ □

Remarks 9.34. In a sense Stone's theorem can be viewed as a special case in a larger picture arising from the *Hille–Yosida theorem* [Rud91], which has had momentous consequences in mathematical physics, esp. concerning the applications of the theory of *semigroups*. Let us remind that, in a Banach space $(\mathsf{X}, ||\ ||)$, a **strongly continuous semigroup of operators** $\{Q_t\}_{t\in[0,+\infty)}$ is a collection of operators $Q_t \in \mathfrak{B}(\mathsf{X})$ such that: (a) $Q(0) = I$, (b) $Q_{t+s} = Q_tQ_s$ for $s,t\in[0,+\infty)$, and (c) $||Q_t\psi - \psi|| \to 0$ as $t\to 0$ for any $\psi\in\mathsf{X}$. In this context one proves [Rud91] that every strongly continuous semigroup admits one **generator**, i.e. an operator A on X completely determined by the demand

$$\frac{d}{dt}Q_t\psi = -AQ_t\psi = -Q_tA\psi\,, \quad \psi\in D(A).$$

(The derivative is computed in the norm of X.) It turns out $D(A)\subset\mathsf{X}$ is dense.

If we look at the subcase of normal (bounded) operators $\{Q_t\}_{t\in[0,+\infty)}$ on a Hilbert space $\mathsf{X}=\mathsf{H}$, then [Rud91]: (1) the semigroup has a generator A, (2) A is normal (unbounded, in general), and (3)

$$Q_t = e^{-tA}\,,$$

where the right-hand side is defined by the integral of

$$\sigma(A)\ni\lambda\mapsto e^{-t\lambda}$$

in the PVM of the spectral decomposition of A (extending Theorem 9.10 to unbounded normal operators [Rud91]). Eventually, (4) the real part of the spectrum of A is lower bounded, i.e. there is $\gamma\in\mathbb{R}$ such that $\gamma < Re(\lambda)$ for any $\lambda\in\sigma(A)$. ■

9.3.3 Commuting operators and spectral measures

As a concluding result we prove a theorem on commuting spectral measures of self-adjoint operators, which relies on the generated one-parameter groups. For *bounded* self-adjoint operators the spectral measures commute if and only if the operators themselves commute, an easy consequence of the spectral theorem (see also Corollary 9.36). For *unbounded* operators, instead, there are domain-related issues and the criterion cannot be used. Using unitary groups is a simple way to overcome this problem. The next result is widely applied in QM.

Theorem 9.35. *Let A, B be (in general unbounded) operators on the Hilbert space* H *with A self-adjoint.*

(i) *If B is self-adjoint and calling* $P^{(A)}$, $P^{(B)}$ *the respective spectral measures, the following statements are equivalent.*
(a) *For any Borel sets* $E, E' \subset \mathbb{R}$*:*

$$P^{(A)}(E)P^{(B)}(E') = P^{(B)}(E')P^{(A)}(E)\,.$$

(b) *For any Borel set* $E \subset \mathbb{R}$ *and any* $s \in \mathbb{R}$*:*

$$P^{(A)}(E)e^{-isB} = e^{-isB}P^{(A)}(E)\,.$$

(c) *For any* $t, s \in \mathbb{R}$*:*

$$e^{-itA}e^{-isB} = e^{-isB}e^{-itA}\,.$$

(d) *For any real* $t \in \mathbb{R}$*:*

$$e^{-itA}D(B) \subset D(B) \text{ and } \quad e^{-itA}B\psi = Be^{-itA}\psi\,, \quad \text{if } \psi \in D(B).$$

(ii) *Under either of the above four conditions:*

$$AB\psi = BA\psi \qquad \text{if } \psi \in D(AB)\cap D(BA)$$
$$(A\varphi|\, B\psi) - (B\varphi|A\psi) = 0 \text{ if } \psi,\varphi \in D(A)\cap D(B).$$

(iii) *If* $B \in \mathfrak{B}(\mathsf{H})$ *(not necessarily self-adjoint) and* $P^{(A)}$ *is the PVM of A, the following are equivalent.*
(e) $BA\varphi = AB\varphi$ *for any* $\varphi \in D(A)$.
(f) $Bf(A)\psi = f(A)B\psi$ *for any* $\psi \in D(f(A))$ *and any* $f : \sigma(A) \to \mathbb{R}$ *measurable.*
(g) $BP^{(A)}(E) = P^{(A)}(E)B$ *for any Borel set* $E \subset \mathbb{R}$.

Proof. (i) Using Definition 9.11 the identity in (b) reads

$$\int_{\mathbb{R}} e^{-it\lambda}\, dP^{(A)}_\lambda \int_{\mathbb{R}} e^{-is\mu}\, dP^{(B)}_\mu = \int_{\mathbb{R}} e^{-is\mu}\, dP^{(B)}_\mu \int_{\mathbb{R}} e^{-it\lambda}\, dP^{(A)}_\lambda\,, \quad t,s \in \mathbb{R}, \tag{9.85}$$

where the standard definition of integral of a *bounded* measurable map in a spectral measure was employed, by Theorem 9.4(a). That (a) implies (c) is immediate by definition of integral of a bounded map in a spectral measure (Chapter 8) working in the strong topology. Let us prove (c) $\Rightarrow$ (b) $\Rightarrow$ (a). For the former implication, from (9.85), if $U_s := e^{-isB}$, $\psi, \phi \in \mathsf{H}$, $s \in \mathbb{R}$ are fixed, we have $\left(\psi \middle| \int_{\mathbb{R}} e^{-it\lambda}\, dP^{(A)}_\lambda U_s \phi\right) = \left(U_s^*\psi \middle| \int_{\mathbb{R}} e^{-it\lambda}\, dP^{(A)}_\lambda \phi\right)$ for any $t \in \mathbb{R}$, i.e.

$$\int_{\mathbb{R}} e^{-it\lambda}\, d\mu^{(A)}_{\psi, U_s\phi}(\lambda) = \int_{\mathbb{R}} e^{-it\lambda}\, d\mu^{(A)}_{U_s^*\psi,\phi}(\lambda)\,, \tag{9.86}$$

where we introduced complex measures as in Theorem 8.50(c). The above can be transformed in integrals for *finite* positive measures by Theorem 1.86. Next, using

Fubini–Tonelli in (9.86) we can say that if f is the Fourier transform of a map in the Schwartz space $\mathscr{S}(\mathbb{R})$ (see Chapter 3):

$$\int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t)\, e^{-it\lambda} dt \right) d\mu^{(A)}_{\psi, U_s\phi}(\lambda) = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(t)\, e^{-it\lambda} dt \right) d\mu^{(A)}_{U_s^*\psi,\phi}(\lambda)\,.$$

As the Fourier transform maps $\mathscr{S}(\mathbb{R})$ to itself bijectively, the identity becomes

$$\int_{\mathbb{R}} g(\lambda)\, d\mu^{(A)}_{\psi, U_s\phi}(\lambda) = \int_{\mathbb{R}} g(\lambda)\, d\mu^{(A)}_{U_s^*\psi,\phi}(\lambda)\,, \quad g \in \mathscr{S}(\mathbb{R}). \tag{9.87}$$

If $h \in C_c(\mathbb{R})$ (continuous with compact support), the sequence

$$g_n(x) := \sqrt{\frac{n}{4\pi}} \int_{\mathbb{R}} e^{-n(x-y)^2/4} h(y) dy$$

satisfies $g_n \in \mathscr{D}(\mathbb{R})$ and converges uniformly to h as $n \to +\infty$. As $g_n \in \mathscr{D}(\mathbb{R}) \subset \mathscr{S}(\mathbb{R})$ and $g_n \to h \in C_c(\mathbb{R})$ in sup norm, and measures are finite, (9.87) implies

$$\int_{\mathbb{R}} h(\lambda)\, d\mu^{(A)}_{\psi, U_s\phi}(\lambda) = \int_{\mathbb{R}} h(\lambda)\, d\mu^{(A)}_{U_s^*\psi,\phi}(\lambda)\,, \quad h \in C_c(\mathbb{R}). \tag{9.88}$$

Riesz's Theorem 2.48 for complex measures ensures the measures involved in the integrals above coincide. By their explicit expression (Theorem 8.50(c)):

$$\left(\psi \middle| P^{(A)}(E) U_t \phi\right) = \left(U_s^*\psi \middle| P^{(A)}(E)\phi\right) \quad \text{for any Borel set } E \subset \mathbb{R} \text{ and any } s \in \mathbb{R}. \tag{9.89}$$

As ψ, ϕ are arbitrary, obvious manipulations give (b):

$$P^{(A)}(E) e^{-isB} = e^{-isB} P^{(A)}(E) \quad \text{for any Borel set } E \subset \mathbb{R}, \text{ any } s \in \mathbb{R}. \tag{9.90}$$

Now we can prove (b) implies (a). Iterating the procedure that leads to (b) knowing (c), where now e^{-isB} replaces e^{-itA} and the unitary U_s is replaced by the projector $P^{(A)}(E)$, we obtain that (9.90) implies (a): $P^{(A)}_E P^{(B)}_{E'} = P^{(B)}_{E'} P^{(A)}_E$ for any pair of Borel sets $E, E' \subset \mathbb{R}$.

To finish (i) there remains to show (d) is equivalent to one of the preceding statements. If (c) holds, by Stone's theorem and the continuity of e^{-itA} (d) follows immediately. On the other hand (d) implies (c), let us see why. First of all (d) amounts to $e^{-itA} B e^{itA} = B$, so exponentiating gives $e^{-is(e^{-itA} B e^{itA})} = e^{-isB}$. For any given $s \in \mathbb{R}$ the strongly continuous one-parameter unitary groups $t \mapsto e^{-is(e^{-itA} B e^{itA})}$ and $t \mapsto e^{-itA} e^{-isB} e^{itA}$ have the same generator, so they coincide by Stone's theorem. Hence $e^{-itA} e^{-isB} e^{itA} = e^{-isB}$, i.e. (c).

Let us prove (ii). For the first assertion, take $\psi \in D(AB) \cap D(BA)$ and look at (c) in (i): $e^{-itA} e^{-isB} \psi = e^{-isB} e^{-itA} \psi$. Differentiating in t at the origin, Stone's theorem gives $A e^{-isB} \psi = e^{-isB} A \psi$. Now we differentiate in s at the origin. The right side gives $-iBA\psi$ by Stone. On the left we can move the derivative past A, as $A = A^*$ is closed and because the limit exists. Hence $-iAB\psi = -iBA\psi$, as we wanted. Now we prove the second assertion, assuming again (c). If $\psi \in D(A)$ and $\varphi \in D(B)$, $(e^{itA}\psi|e^{-isB}\varphi) = (e^{isB}\psi|e^{-itA}\varphi)$. Differentiating in t and s at $t = s = 0$ proves the claim by Stone's theorem.

Assertion (iii) goes like this. It is obvious that (f) implies (e) and (g) (choose $f = \chi_E$). So we show (e) $\Rightarrow$ (f). First we prove (e) forces B to commute with e^{-itA} for any $t \in \mathbb{R}$. For this we shall use (d), (f) in Proposition 9.21. Let ψ be analytic for A and for every power in the dense set that exists by Proposition 9.21(f). As B is bounded, using Proposition 9.21(d):

$$Be^{-itA}\psi = \sum_{n=0}^{+\infty} \frac{(-it)^n}{n!} BA^n\psi = \sum_{n=0}^{+\infty} \frac{(-it)^n}{n!} A^n B\psi = e^{-itA}B\psi\,.$$

In the last two equalities we used $BA\psi = AB\psi$ repeatedly, plus $||A^nB\psi|| = ||BA^n\psi|| \leq ||B||||A^n\psi||$, so $B\psi$ is analytic for A. But ψ moves in a dense set and the operators B, e^{-itA} are continuous, so $Be^{-itA} = e^{-itA}B$.

If B is bounded and commutes with every e^{-itA}, B commutes with the spectral measure of A. The proof is similar to the proof that, in (i), (c) implies (b): we just have to replace U_s by B. Hence by definition of $g(A)$, *if g is bounded* (and so is $g(A)$) then $Bg(A) = g(A)B$. At this point notice

$$\mu^{(A)}_{B\psi}(E) = (B\psi|P^{(A)}(E)B\psi) = (P^{(A)}B\psi|P^{(A)}(E)B\psi)$$

$$= (BP^{(A)}\psi|BP^{(A)}(E)\psi) \leq ||B||^2\mu^{(A)}_\psi(E) \tag{9.91}$$

so $\psi \in D(f(A))$ implies $B\psi \in D(f(A))$. Applying the definition of $f(A)$ for f measurable unbounded, and taking a sequence of bounded measurable maps f_n converging to f in $L^2(\sigma(A),\mu_\psi)$, we have the claim by taking the limit as $n \to +\infty$ of $Bf_n(A)\psi = f_n(A)B\psi$, for any $n \in \mathbb{N}$, since B is continuous (the equality holds for f_n is bounded). At last, (g) implies (9.91), and (e) follows from the previous argument with $f(x) = x$. □

Corollary 9.36. *Consider two self-adjoint operators $A : D(A) \to \mathsf{H}$, $B \in \mathfrak{B}(\mathsf{H})$ on the Hilbert space H. They commute, i.e. $AB\psi = BA\psi$ for any $\psi \in D(A)$, if and only if their spectral measures commute.*

Proof. If the operators commute, (g) holds in (iii) above. Apply (iii) giving B the role of A and $P^{(A)}(E)$ the role of B; then $P^{(B)}(F)P^{(A)}(E) = P^{(A)}(E)P^{(B)}(F)$ for any Borel sets $E,F \subset \mathbb{R}$. Conversely if the spectral measures commute, by definition of integral of a bounded PVM follows $BP^{(A)}(E) = P^{(A)}(E)B$ for any Borel set $E \subset \mathbb{R}$. Now (iii) implies $AB\psi = BA\psi$ for any $\psi \in D(A)$. □

Here is another useful technical consequence.

Corollary 9.37. *Let A be self-adjoint on the Hilbert space H and $B_0 : D(B_0) \to \mathsf{H}$ essentially self-adjoint. If*

$$e^{-itA}D(B_0) \subset D(B_0)\,, \quad e^{-itA}B_0\phi = B_0e^{-itA}\phi\,, \quad \forall t \in \mathbb{R}\,, \forall\phi \in D(B_0)\,,$$

then A and $B := \overline{B_0}$ satisfy (a), (b), (c), (d) in Theorem 9.35(i).

Proof. It suffices to note that by definition of closure, using the continuity of $e^{-it}A$, the self-adjoint operator $B : D(B) \to \mathsf{H}$ satisfies

$$e^{itA}D(B) \subset D(B)\,, \quad e^{-itA}B\phi = Be^{-itA}\phi\,, \quad \forall t \in \mathbb{R}, \forall \phi \in D(B)\,.$$

Then part (i) in Theorem 9.35 produces the claim. □

Exercises

9.1. Consider a PVM $P : \mathscr{B}(\mathsf{X}) \ni E \mapsto P(E) \in \mathfrak{B}(\mathsf{H})$ and a unitary operator (isometric and onto) $V : \mathsf{H} \to \mathsf{H}'$, where H is a complex Hilbert space. Prove

$$P' : \mathscr{B}(\mathsf{X}) \ni E \mapsto P'(E) := VP(E)V^{-1} \in \mathfrak{B}(\mathsf{H}')$$

is a PVM.

9.2. In relationship to Exercise 9.1, prove the following facts.

(i) If $f : \mathsf{X} \to \mathbb{C}$ is measurable then $\psi \in \Delta_f \;\Leftrightarrow\; V\psi \in \Delta'_f$, where Δ'_f is the domain of the integral of f in P'.
(ii) $V \int_{\mathsf{X}} f(x)dP(x)V^{-1} = \int_{\mathsf{X}} f(x)dP'(x)$.

9.3. If H is a Hilbert space, prove $T \in \mathfrak{B}(\mathsf{H})$ is of trace class ($T \in \mathfrak{B}_1(\mathsf{H})$) $\Leftrightarrow$ $\sum_{u\in N} |(u|Tu)| < +\infty$ for any basis $N \subset \mathsf{H}$.

Solution. Suppose $\sum_{u\in N} |(u|Tu)| < +\infty$ for any basis N. Assume, first, $T = T^*$. By the spectral theorem $T = \int_{\mathbb{R}} \lambda dP^{(T)}(\lambda)$. Define $T_- := \int_{(-\infty,0)} \lambda dP^{(T)}(\lambda)$ and $T_+ := \int_{[0,+\infty)} \lambda dP^{(T)}(\lambda)$. Clearly $T_\pm \in \mathfrak{B}(\mathsf{H})$ ($|\lambda\chi_{(-\infty,0)}| \leq |\lambda|$, $|\lambda\chi_{[0,+\infty)}| \leq |\lambda|$, Corollary 9.5 holds, and λ is bounded on the support of $P^{(T)}$ by the spectral theorem for self-adjoint bounded operators). Moreover $\pm T_\pm \geq 0$ by Theorem 10.37. Further, $\mathsf{H} = \mathsf{H}_- \oplus \mathsf{H}_+$ is a closed orthogonal sum where $\mathsf{H}_- := P^{(T)}((-\infty,0))\mathsf{H}$, $\mathsf{H}_+ := P^{(T)}([0,+\infty))\mathsf{H}$. Let $N_- \subset \mathsf{H}_-$, $N_+ \subset \mathsf{H}_+$ be bases. $N := N_- \cup N_+$ is a basis of H. As $-T_-, T_+ \geq 0$ and $T_\pm u = 0$ if $u \in \mathsf{H}_\mp$, we have

$$+\infty > \sum_{u\in N} |(u|Tu)| = \sum_{u\in N_-} |(u|T_-u)| + \sum_{u\in N_+} |(u|T_+u)| = \sum_{u\in N_-} -(u|T_-u) + \sum_{u\in N_-} (u|T_+u)$$

$$= \sum_{u\in N_-} (u||T_-|u) + \sum_{u\in N_-} (u||T_+|u) = \sum_{u\in N} (u||T_-|u) + \sum_{u\in N} (u||T_+|u)\,.$$

Therefore $T_\pm \in \mathfrak{B}_1(\mathsf{H})$ by Definition 4.30, so $T = T_+ + T_- \in \mathfrak{B}_1(\mathsf{H})$ by Theorem 4.32(b). In case T is not self-adjoint, we can decompose $T = A + iB$, with $A := \frac{1}{2}(T + T^*)$, $B := \frac{1}{2i}(T - T^*)$, A, B self-adjoint. For any given basis $N \subset \mathsf{H}$, $|(u|Tu)| = |(u|Au) + i(u|Bu)| = \sqrt{|(u|Au)|^2 + |(u|Bu)|^2} \geq |(u|Au)|\,, |(u|Bu)|$, with $u \in N$. Applying the result proved above for self-adjoint operators gives $A, B \in \mathfrak{B}_1(\mathsf{H})$, so $T \in \mathfrak{B}_1(\mathsf{H})$. If, instead, $T \in \mathfrak{B}_1(\mathsf{H})$, for any basis $\sum_{u\in N} |(u|Tu)| < +\infty$ by Proposition 4.34.

9.4. Prove that (iv) in Theorem 9.10(b) can be strengthened to: *Let* $T : D(T) \to \mathsf{H}$ *be self-adjoint on the Hilbert space* H. *Then* $\lambda \in \sigma_c(T)$ *is equivalent to asking:* $0 < ||T\phi - \lambda\phi||$, $\forall \phi \in D(T)$ *with* $||\phi|| = 1$, *and for any* $\varepsilon > 0$ *there is* $\phi_\varepsilon \in D(T)$, $||\phi_\varepsilon|| = 1$, *such that*

$$||T\phi_\varepsilon - \lambda\phi_\varepsilon|| \leq \varepsilon .$$

Hint. The second condition amounts to saying λ does not belong to $\sigma_p(T)$, so $(T - \lambda I)^{-1} : Ran(T - \lambda I) \to D(T)$ exists. Then $\lambda \in \sigma(T) = \sigma_p(T) \cup \sigma_c(T)$. Can $(T - \lambda I)^{-1}$ be bounded?

9.5. Consider a space $L^2(\mathsf{X}, \mu)$ with μ positive and finite on the Borel σ-algebra of a space X. Let $f : \mathsf{X} \to \mathbb{R}$ be an arbitrary real, measurable, and locally L^2 map (i.e. $f \cdot g \in L^2(\mathsf{X}, \mu)$ for any $g \in C_c(\mathsf{X})$). Consider the operator on $L^2(\mathsf{X}, \mu)$:

$$T_f : h \mapsto f \cdot h$$

where $D(T_f) := \{h \in L^2(\mathsf{X}, \mu) \,|\, f \cdot h \in L^2(\mathsf{X}, \mu)\}$. Prove T_f is self-adjoint. Then show

$$\sigma(T_f) = \text{ess ran}(f) .$$

For $f : \mathsf{X} \to \mathbb{R}$, ess ran$(f)$ is the **essential rank** of the measurable map f, defined by $\mathbb{R} \ni v \in \text{ran ess}(f) \Leftrightarrow \mu\left(f^{-1}(v - \varepsilon, v + \varepsilon)\right) > 0$ for any $\varepsilon > 0$.

Hint. The domain of T_f is dense because f is locally L^2, and the self-adjointness comes from computing $T_f^* = T_f$. The second part goes along these lines. $\lambda \in \rho(T_f)$ $\Leftrightarrow$ the resolvent $R_\lambda(T_f)$ exists on $L^2(\mathsf{X}, \mu)$ and is bounded, i.e. there is $M > 0$ such that $||R_\lambda(T_f)h|| \leq M$ for any unit $h \in L^2(\mathsf{X}, \mu)$. That is to say, $\lambda \in \rho(T_f)$ if and only if:

$$\int_{\mathsf{X}} \frac{|h(x)|^2}{|f(x) - \lambda|^2} d\mu(x) < M \quad \text{for any unit } h \in L^2(\mathsf{X}, \mu) .$$

If $\lambda \notin$ ess ran(f), by definition of essential rank and $\mu(\mathsf{X}) < +\infty$ we see that the condition holds, so $\lambda \notin$ ess ran(f) implies $\lambda \notin \sigma(T_f)$. If $\lambda \in$ ess ran(f), still by definition of essential rank we build a sequence of unit h_n such that $\int_{\mathsf{X}} \frac{|h(x)|^2}{|f(x)-\lambda|^2} d\mu(x) > 1/n^2$ for any $n = 1, 2, \ldots$. Hence $\lambda \in$ ess ran(f) implies $\lambda \in \sigma(T_f)$.

9.6. Consider a PVM $P : \mathscr{B}(\mathbb{C}) \to \mathsf{H}$ with H separable. Prove $A \in \mathfrak{B}(\mathsf{H})$ has the form $A = \int_{\mathbb{C}} f dP$ for some $f \in M_b(\mathbb{C})$ $\Leftrightarrow$ it commutes with every $B \in \mathfrak{B}(\mathsf{H})$ satisfying $BP(E) = P(E)B$ for any $E \in \mathscr{B}(\mathbb{C})$.

Solution. The converse statement is known, so we just prove the necessary part of the '$\Leftrightarrow$'. Divide $supp(P)$ in a disjoint collection, at most countable, of bounded sets E_n, and H in the corresponding orthogonal sum $\mathsf{H} = \oplus_n \mathsf{H}_n$, $\mathsf{H}_n := P(E_n)(\mathsf{H})$. A makes every H_n invariant, since $AP(E_n) = P(E_n)A$ by assumption. If $A_n := A\restriction_{\mathsf{H}_n} : \mathsf{H}_n \to \mathsf{H}_n$, then $A\psi = \sum_n A_n\psi$ for any $\psi \in \mathsf{H}$. Moreover (see Corollary 9.36) A_n commutes with any operator in $\mathfrak{B}(\mathsf{H}_n)$ that commutes with the bounded normal $T_n := \int_{E_n} z dP(z)$ and its adjoint. By Theorem 9.9, $A_n = \int_{E_n} f_n dP$ for some $f_n \in M_b(E_n)$. Define

$f(z) := f_n(z)$ on $z \in E_n$, for any $z \in \mathbb{C}$. Then $f_n \to f$ (the f_n are null outside E_n) in $L^2(\mathbb{C}, \mu_\psi)$ by dominated convergence if $\psi \in \Delta_f$. Therefore $A\psi = \int_{\mathbb{C}} f dP\psi$ for $\psi \in \Delta_f$, by definition of $\int_{\mathbb{C}} f dP$. As A is bounded, Corollary 9.5 implies f must be bounded, $\Delta_f = \mathsf{H}$ and $A = \int_{\mathbb{C}} f dP$.

9.7. Let H be separable and $T : \mathscr{D}(T) \to \mathsf{H}$ self-adjoint on H (not necessarily bounded). Prove that $A \in \mathfrak{B}(\mathsf{H})$ has the form $A = f(T)$, for some $f : \mathbb{R} \to \mathbb{C}$ measurable and bounded, $\Leftrightarrow$ A commutes with every $B \in \mathfrak{B}(\mathsf{H})$ such that $BT\psi = TB\psi$ for any $\psi \in \mathscr{D}(T)$.

Solution. If $P^{(T)}$ is the PVM of T, $BP^{(T)}(E) = P^{(T)}(E)B \Leftrightarrow BT\psi = TB\psi$ for any $\psi \in \mathscr{D}(T)$. The claim boils down to proving $A = \int f dP^{(T)}$, f bounded, iff A commutes with any $B \in \mathfrak{B}(\mathsf{H})$ commuting with $P^{(T)}$. Exercise 9.6 does exactly that.

9.8. If A is the self-adjoint generator of a strongly continuous one-parameter unitary group $U_t = e^{itA}$, prove A is bounded, and hence it is defined on the whole Hilbert space if and only if $||U_t - I|| \to 0$, as $t \to 0$.

Hint. Passing to the spectral representation of A, we have $||U_t - I|| = ||f_t||_\infty$ where $f_t(\lambda) = |e^{it\lambda} - 1|$. As $(a,b) \ni \lambda \mapsto f_t(\lambda)$ tends λ-uniformly to 0, as $t \to 0$ if and only if a, b are finite, the claim follows.

9.9. Show formula (9.72) for any one-parameter unitary group $\mathbb{R} \ni t \mapsto U_t$.

Solution. Applying U_{-t} on the right to $U_t^{-1}U_t = I$, and using the second relation in (9.71), we have $U_t^{-1}U_{t-t} = U_{-t}$. Using the first identity in (9.71), $U_t^{-1} = U_{-t}$. Eventually $U_t^{-1} = (U_t)^*$, as U_t is unitary.

9.10. Consider the operators $A, A^\star$ of Section 9.1.4. Prove they are closable, and that

$$\overline{A^\star A} = \overline{A^* A} = \overline{\overline{A}^* \overline{A}}\,.$$

9.11. Study the polar decomposition $\overline{A} = UP$ for the operator A of section 9.1.4. Prove that U satisfies

$$U\psi_n = \psi_{n-1}$$

if $n \geq 1$ and $\{\psi_n\}_{n\in\mathbb{N}}$ is the basis of $L^2(\mathbb{R}, dx)$ defined in Section 9.1.4.

9.12. Consider the operators A and $A^\star$ of Section 9.1.4 and the basis $\{\psi_n\}_{n\in\mathbb{N}}$ of $L^2(\mathbb{R}, dx)$. Compute $e^{\overline{\alpha A + \overline{\alpha} A^\star}}\psi_n$ with $\alpha \in \mathbb{C}$ given.

10

Spectral Theory III: applications

Particles are solutions to a differential equation.
Werner Karl Heisenberg

In this chapter we examine applications of the theory of unbounded operators in Hilbert spaces, where spectral theory, as developed in Chapters 8 and 9, plays a paramount technical role during the proofs. The final part of the chapter presents a series of classical results about certain operators of interest in Quantum Mechanics, in particular regarding self-adjointness and spectral lower bounds.

Section one is devoted to the study of abstract differential equations in Hilbert spaces.

The second section pertains the notion of *Hilbert tensor product* of Hilbert spaces and of operators (typically unbounded), plus their spectral properties. We apply this to one example, the orbital angular momentum of a quantum particle.

We extend the polar decomposition theorem to closed, densely-defined unbounded operators in the third section. The properties of operators of the form A^*A, with A densely defined and closed, are examined, together with square roots of unbounded positive self-adjoint operators.

Section four contains the statement, the proof and a few direct applications of the *Kato-Rellich theorem*, which gives criteria for a self-adjoint operator, perturbed by a symmetric operator, to be still self-adjoint operator, and establishes lower bounds for the perturbed spectrum.

10.1 Abstract differential equations in Hilbert spaces

Looking at spectral theory from the right angle allows to tackle the issue of existence and uniqueness of solutions to certain PDEs that are important in physics. Recall [Sal08] that a second-order linear differential equation in $u \in C^2(\Omega;\mathbb{R})$, with given open set $\Omega \subset \mathbb{R}^n$ and continuous real maps a_{ij}, b_i, c, has the form:

$$\sum_{i,j=1}^{n} a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^{n} b_i(x)\frac{\partial u}{\partial x_i} + c(x)u = 0\,.$$

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_10,

Equations of this kin are classified, *pointwise at each* $x \in \Omega$, into (a) *elliptic*, (b) *parabolic* or (c) *hyperbolic* according to the spectrum of the symmetric matrix $a_{ij}(x)$. Respectively, type (a) occurs when the eigenvalues have the same sign $\pm$, (b) when there is some null eigenvalue, (c) when no eigenvalue vanishes and there are eigenvalues opposite in sign. An equation is said *elliptic*, *parabolic* or *hyperbolic* if such at each point $x \in \Omega$.

By a smart coordinate choice around each point in Ω, the equation can be written as:

$$a(t,y)\frac{\partial^2}{\partial t^2}u(t,y) + \sum_{i,j=1}^{n-1} a'_{ij}(t,y)\frac{\partial^2 u}{\partial y_i \partial y_j} + b(t,y)\frac{\partial u}{\partial t}$$

$$+\sum_{i=1}^{n-1} b'_i(t,y)\frac{\partial u}{\partial y_i} + c(t,y)u(t,y) = 0\,.$$

For elliptic equations (e.g. *Poisson's equation*) $a(t,y)$ is never zero and has the same sign of the eigenvalues (all non-zero) of the symmetric matrix $a'_{ij}(t,y)$. Parabolic equations (e.g. the *heat equation*) have $a(t,y) = 0$. Hyperbolic equations (e.g. *d'Alembert's equation*) are such that $a(t,y)$ has opposite sign to some eigenvalues (none zero) of the symmetric matrix $a'_{ij}(t,y)$.

Supposing all functions we consider are *complex-valued* we shall study the theory of these PDEs under a different point of view. The above will be considered "abstract differential equations" in Hilbert spaces and with respect to suitable topologies. The variable t will be regarded as a parameter upon which the solutions depends: this will give a curve in the Hilbert space. The differential operators determined by the matrix a'_{ij} and the vector b'_i will become operators acting on a subspace in a Hilbert space $L^2(\Omega, dy)$ containing the support of the solution curve. We could, as a matter of fact, use a completely abstract Hilbert space H, without a mention of coordinates. Solutions will therefore be functions $t \mapsto u_t \in \mathsf{H}$ in the Hilbert space. This generalisation will allow us to treat equations that do not befit the classical trichotomy (like the *Schrödinger equation*), equations of degree higher than the second, and equations that cannot be classified within the above scheme, like equations where the differential operator given by the matrix a'_{ij} is formally replaced by a square root. For instance

$$a\frac{\partial u}{\partial t} + b\sqrt{-\frac{\partial^2}{\partial x^2}}\,u = 0\,.$$

Notation 10.1. If $J \subset \mathbb{R}$ is an interval, H a Hilbert space, $S \subset \mathsf{H}$ a subspace (closed or not) and $k = 0, 1, 2, \ldots$ is fixed, we let

$$C^k(J;S)$$
$$:= \left\{ J \ni t \mapsto u(t) \in S \;\middle|\; J \ni t \mapsto \frac{d^j u}{dt^j} \text{ exists and is continuous for } j = 0, 1 \ldots, k \right\},$$

where derivative and continuity refer to the topology of H.
We shall also write $C(J;S) := C^0(J;S)$. ■

Remark 10.2. **(1)** Of course $C(J;S)$ is a complex vector space.
(2) It is easy to prove, by inner product/norm continuity and Schwarz's inequality that:

$$\frac{d}{dt}(u(t)|v(t)) = \left(\frac{du}{dt}\middle|v(t)\right) + \left(u(t)\middle|\frac{dv}{dt}\right) \tag{10.1}$$

for every $t \in J$, when $u, v : J \to \mathsf{H}$ are differentiable everywhere on J (in particular continuous on J).
(3) If $\mathsf{H} = L^2(\Omega, dx)$, $\Omega \subset \mathbb{R}^n$ open, and we take a family of maps $u_t \in \mathscr{L}^2(\Omega, dx)$, $(t, x) \in J \times \Omega$ for a given oper interval $J \subset \mathbb{R}$, the existence of the derivative at t forces the existence in $L^2(\Omega, dx)$, under rather weak hypotheses. For example

Proposition 10.3. *Let $\Omega \subset \mathbb{R}^n$ be bounded and open, $J \subset \mathbb{R}$ an open interval and $\{u_t\}_{t\in J} \subset \mathscr{L}^2(\Omega, dx)$ a family defined on Ω.*
If the maps $u = u_t(x)$ are differentiable in t for every $x \in \Omega$ and $|\frac{\partial u_t}{\partial t}| \leq M$ for some $M \in \mathbb{R}$ and any $t \in J$, then (viewing $\{u_t\}_{t\in J} \subset L^2(\Omega, dx)$ for the derivative):

$$\exists \, \frac{du_t}{dt} \quad \textit{for every } t \in J \textit{ and } \quad \frac{du_t}{dt} = \frac{\partial u_t}{\partial t} \quad \textit{a.e. at } x \textit{ for any } t \in J,$$

where the derivative is computed as usual.

(This generalises to higher derivatives in the obvious way.)

Proof. Note $\Omega \ni x \mapsto \frac{\partial u_t}{\partial t}$ is measurable for any $t \in J$ as pointwise limit of measurable functions. For any given $t \in J$, the mean value theorem says that for every $x \in \Omega$ and some $x'(x, t, h) \in [t - h, t + h]$:

$$\int_\Omega \left| \frac{u_{t+h}(x) - u_t(x)}{h} - \frac{\partial u_t}{\partial t}(x) \right|^2 dx = \int_\Omega \left| \frac{\partial u_t}{\partial t}(x'(x, t, h)) - \frac{\partial u_t}{\partial t}(x) \right|^2 dx\,.$$

The right integrand is smaller, uniformly with respect to h, than the constant M in $L^2(\Omega, dx)$, since Ω has finite Lebesgue measure. Since the integrand is pointwise infinitesimal as $h \to 0$, dominated convergence proves the claim. □

Having Ω bounded can be dropped in favour of a uniform estimate in t, of the sort $|\frac{\partial u_t}{\partial t}(x)| \leq |g_{t_0}(x)|$ with $g_{t_0} \in \mathscr{L}^2(\Omega, dx)$, holding around every given $t_0 \in J$. ■

10.1.1 The abstract Schrödinger equation (with source)

The first equation we study is Schrödinger's equation, for which we allow a source term to be present. The equation will be considered abstractly, in a Hilbert space, and without referring to physics. We shal return to it in Chapter 13, when the physical meaning of the sourceless case will be discussed. For the standard theory of PDEs, Schrödinger's equation has the following structure (numerical coefficients apart, whose great relevance is neglected for the time being):

$$-i\frac{\partial}{\partial t}u_t(x) + (A_0 u_t)(x) = S(t, x) \tag{10.2}$$

where $J \subset \mathbb{R}$ a fixed open interval, $\Omega \subset \mathbb{R}^n$ a given open set,

$$A_0 := -\Delta_x + V(x) : D(A_0) \to L^2(\Omega, dx) \tag{10.3}$$

is defined on some domain $D(A_0) \subset C^2(\Omega)$, $V : \Omega \to \mathbb{R}$ and $S : J \times \Omega \to \mathbb{C}$ are given maps, say continuous, and finally Δ_x is the usual Laplacian on $\mathbb{R}^n$.

A function $u = u(t,x)$ is called **classical solution** to (10.2) if it is defined for $(t,x) \in J \times \Omega$, of class C^1 in t and C^2 in $x_1, \ldots, x_n$, and of course if it solves (10.2) on its domain.

If the functions in $D(A_0)$ decay quickly outside compact sets in Ω and first derivatives are bounded, the operator A_0 is certainly Hermitian, as is clear integrating by parts. At least for $\Omega := \mathbb{R}^n$, we expect that choosing $D(A_0)$ properly will make A_0 essentially self-adjoint in $L^2(\Omega, dx)$. We already know that for $\Omega := \mathbb{R}^n$ and $V := 0$, the operator A_0 of (10.3) is essentially self-adjoint on the domain $D(A_0) := \mathscr{S}(\mathbb{R}^n)$ (Exercises 5.11 and 5.12); as we shall see, the same holds on $D(A_0) := \mathscr{D}(\mathbb{R}^n)$. We will discuss more general cases, with $V \neq 0$, in Section 10.4.

Assuming $A := \overline{A_0}$ is self-adjoint leads to a different interpretation of equation (10.2), where A_0 is replaced by any self-adjoint operator and t-differentiation is defined in the Hilbert topology.

Let us fix an open interval $J \subset \mathbb{R}$, $J \ni 0$. If $A : D(A) \to \mathsf{H}$ is a self-adjoint operator on the Hilbert space H and $J \ni t \mapsto S_t \in \mathsf{H}$ a given map in $C(J; \mathsf{H})$, the abstract **Schrödinger equation** with source term is:

$$-i\frac{d}{dt}u_t + Au_t = S_t \tag{10.4}$$

where $u \in C^1(J; D(A))$ is the unknown. As mentioned, the derivative is computed in the topology of H. The *source* is the function $S = S_t$. If $S_t = 0$ for any $t \in J$, equation (10.4) is as usual called *homogeneous*.

The **Cauchy problem for the Schrödinger equation**, resp. with source or homogeneous, is the problem of finding a function $u \in C^1(J; \mathsf{H})$ solving (10.4) with or without source plus the **initial condition**:

$$u_0 = v \in D(A)\,. \tag{10.5}$$

Remarks 10.4. If A_0 is of the form (10.3) and essentially self-adjoint, we can consider a classical solution $u = u(t,x)$ to (10.2), for which $u(t,\cdot) \in D(A_0)$ for any $t \in J$. Under assumptions of the kind of Proposition 10.3, u also solves the abstract equation (10.4), as $D(A) = D(\overline{A_0}) \supset D(A_0)$. Therefore classical solutions are abstract solutions, under mild assumptions. ■

The first result establishes the uniqueness of the solution to the abstract Schrödinger equation with any initial condition.

Proposition 10.5. *If $u = u_t$ solves the homogeneous equation (10.4):*

$$||u_t|| = ||u_0|| \quad \textit{for any } t \in J. \tag{10.6}$$

Hence if a solution to the Cauchy problem (10.4)-(10.5) exists, with $S_t \neq 0$ in general, it is unique.

Proof. From (10.1) and (10.4), for $S_t = 0$:

$$\frac{d}{dt}||u_t||^2 = \frac{d}{dt}(u_t|u_t) = i(Au_t|u_t) - i(u_t|Au_t) = 0$$

because A is self-adjoint. So $||u_t|| = ||u_0||$. Uniqueness follows immediately because if u, u' both solve the Cauchy problem (same S_t), then $J \ni t \mapsto u_t - u'_t$ solves (10.4) with $S_t = 0$ and initial condition $u_0 = 0$, so $u_t - u'_t = 0$ for every $t \in J$. □

We are interested in existence now. Actually, we already have everything we need, because Stone's theorem (Theorem 9.29) implies existence in the homogeneous case:

Proposition 10.6. *A solution to the homogeneous Cauchy problem (10.4)-(10.5) has the form:*

$$u_t = e^{-itA}v\,, \quad t \in J\,,$$

where the exponential is understood in spectral sense.

Proof. Immediate consequence of Theorem 9.29. □

Remarks 10.7. If $v \notin D(A)$ we can still define $u_t := e^{-itA}v$, because the domain of the unitary operator e^{-itA} is H. The map $J \ni t \mapsto u_t$ *does not* solve the homogeneous Schrödinger equation. But trivially

$$\frac{d}{dt}(z|u_t) + i(Az|u_t) = 0 \quad \text{for any } z \in D(A),\, t \in J, \tag{10.7}$$

by Stone's theorem, because the inner product is continous and e^{-itA} is unitary, implying $(z|e^{-itA}v) = (e^{itA}z|v)$. From (10.7) this $J \ni t \mapsto u_t$ is called a **weak solution** to the homogeneous Schrödinger equation. ■

It should be clear that the solution set to the Schrödinger equation *with source* (10.4) – if non-empty – consists of functions

$$J \ni t \mapsto u_t^{(0)} + s_t\,,$$

where: s is an arbitrary, but fixed, solution to the non-homogeneous equation (10.4), and $u^{(0)}$ is free in the vector space of *homogeneous* solutions. A solution to the equation with source satisfying the zero initial condition can be written as:

$$s_t = e^{-tiA}\int_0^t e^{\tau iA}S_\tau d\tau\,,$$

assuming something on $S \in C(J;\mathsf{H})$. We can prove the next theorem.

Theorem 10.8. *Let $A : D(A) \to \mathsf{H}$ be a self-adjoint operator on the Hilbert space H, $J \subset \mathbb{R}$ an open interval with $0 \in J$. If:*

(i) $v \in D(A)$*;*

(ii) *$J \ni t \mapsto S_t$ is continuous in the topology of* H*;*
(iii) *$S_t \in D(A)$ for any $t \in J$;*
(iv) *$J \ni t \mapsto AS_t$ is continuous in the topology of* H*;*

there exists a unique solution $J \ni t \mapsto u_t \in C^1(J;D(A))$ to the Cauchy problem

$$\begin{cases} \frac{du_t}{dt} + Au_t = S_t\,, \\ u_0 = v\,. \end{cases} \tag{10.8}$$

The solution has the form:

$$u_t = e^{-itA}v + e^{-tiA}\int_0^t e^{\tau iA} S_\tau d\tau\,, \quad t \in J\,. \tag{10.9}$$

If $J \ni t \mapsto L_t \in \mathfrak{B}(\mathsf{H})$ is continuous in the strong topology and $J \ni t \mapsto \psi_t \in \mathsf{H}$ is continuous, the vector $\int_a^b L_\tau \psi_\tau d\tau \in D(A)$, $a,b \in J$, is by definition the unique element in H *satisfying:*

$$\left(u \middle| \int_a^b L_\tau \psi_\tau d\tau\right) = \int_a^b (u|L_\tau \psi_\tau)\, d\tau\,. \tag{10.10}$$

Proof. By continuity of $t \mapsto \psi_t$, of the scalar product and Schwarz's inequality on the right-hand side of (10.10):

$$\left|\left|\left(u \middle| \int_a^b L_\tau \psi_\tau d\tau\right)\right|\right| \leq K_{a,b}||u|| \quad \text{for any } u \in \mathsf{H}$$

for some constant $K_{a,b} \geq 0$. By Riesz's representation Theorem 3.16 the vector $\int_a^b L_\tau \psi_\tau d\tau \in \mathsf{H}$ is well defined. Schwarz's inequality implies

$$\left|\left|\int_a^b L_\tau \psi_\tau d\tau\right|\right| \leq \int_a^b ||L_\tau \psi_\tau|| d\tau\,, \tag{10.11}$$

namely:

$$\left|\left|\int_a^b L_\tau \psi_\tau d\tau\right|\right|^2 = \int_a^b \int_a^b (L_\tau \psi_\tau | L_s \psi_s)\, ds d\tau = \left|\int_a^b \int_a^b (L_\tau \psi_\tau | L_s \psi_s)\, ds d\tau\right|$$

$$\leq \int_a^b \int_a^b |(L_\tau \psi_\tau | L_s \psi_s)|\, ds d\tau \leq \int_a^b \int_a^b ||L_\tau \psi_\tau||\,||L_s \psi_s|| ds d\tau = \left(\int_a^b ||L\tau \psi_\tau|| d\tau\right)^2 .$$

Proposition 10.5 grants us uniqueness, so we just need existence. We will show the right side of (10.8) solves the Cauchy problem (10.9). By definition $\int_0^t e^{\tau iA} S_\tau d\tau$ is the null vector if $t = 0$, so the right-hand side of (10.9) satisfies $u_0 = v$. We claim $u_t \in D(A)$. We know $e^{itA}v \in D(A)$ by Proposition 10.6. In reality u_t belongs $D(A) = D(A^*)$, since

$$\left(Ax \middle| \int_0^t e^{\tau iA} S_\tau d\tau\right) = \int_0^t (Ax | e^{\tau iA} S_\tau) = \int_0^t (x | e^{\tau iA} AS_\tau) = \left(x \middle| \int_0^t e^{\tau iA} AS_\tau d\tau\right),$$

by definition of adjoint, $A = A^*$, $S_t \in D(A)$ and with $x \in D(A)$. We have proved that if $S_\tau \in D(A)$:

$$A \int_0^t e^{\tau i A} S_\tau d\tau = \int_0^t e^{\tau i A} A S_\tau d\tau \,. \tag{10.12}$$

Summing up we have $u_t \in D(A)$ since $e^{itA}v \in D(A)$, $\int_0^t e^{\tau i A} S_\tau d\tau \in D(A)$ and so $e^{-itA} \int_0^t e^{\tau i A} S_\tau d\tau \in D(A)$ since $e^{-itA}(D(A)) \subset D(A)$ for one-parameter unitary groups generated by self-adjoint operators.

Now we show u_t is a solution. The first term on the right in (10.9) admits derivative $-iAe^{-itA}v$ by Stone's theorem. We want to prove the derivative of the second term equals $-iAe^{-tiA} \int_0^t e^{\tau i A} S_\tau d\tau + S_t$. If so, u_t solves the problem. Define $t' := t + h$ and:

$$\Phi_t := \int_0^t e^{\tau i A} S_\tau d\tau.$$

The derivative of the second term on the right in (10.9) is the limit for $h \to 0$ of:

$$h^{-1}\left(e^{it'A}\Phi_{t'} - e^{itA}\Phi_t\right) = h^{-1}e^{it'A}\left(\Phi_{t'} - \Phi_t\right) + h^{-1}\left(e^{itA} - e^{it'A}\right)\Phi_t \,.$$

The last term converges to $-iAe^{-tiA}\Phi_t$ by Stone's theorem, since $\Phi_t \in D(A)$. As for the first term:

$$h^{-1}e^{it'A}\left(\Phi_{t'} - \Phi_t\right) = h^{-1}e^{itA}\left(\Phi_{t'} - \Phi_t\right) + h^{-1}\left(e^{it'A} - e^{itA}\right)\left(\Phi_{t'} - \Phi_t\right) \,.$$

By continuity of e^{itA} the first terms converges to

$$e^{itA}\left(\lim_{h \to 0} \frac{1}{h} \int_t^{t'} e^{-i\tau A} S_\tau d\tau\right) = e^{itA} e^{-itA} S_t = S_t \,,$$

where, by (10.11), we used:

$$h^{-1}\left|\left|\int_t^{t'} e^{-i\tau A} S_\tau d\tau - e^{-itA} S_t\right|\right| = h^{-1}\left|\left|\int_t^{t'} \left(e^{-i\tau A} S_\tau - e^{-itA} S_t\right) d\tau\right|\right|$$

$$\leq h^{-1} \int_t^{t'} \left|\left|e^{-i\tau A} S_\tau - e^{-itA} S_t\right|\right| d\tau \leq \sup_{\tau \in [t,t']} \left|\left|e^{-i\tau A} S_\tau - e^{-itA} S_t\right|\right| \to 0$$

as $h \to 0$, since $\tau \mapsto e^{-i\tau A} S_\tau$ is continuous from

$$\left|\left|e^{-i\tau A} S_\tau - e^{-itA} S_t\right|\right|^2 = ||S_\tau||^2 + ||S_t||^2 - 2Re\left(S_\tau \left| e^{i\tau A} e^{-itA} \, S_t\right.\right) \,.$$

The last thing to prove is

$$R_h := h^{-1}\left(e^{it'A} - e^{itA}\right)\left(\Phi_{t'} - \Phi_t\right) \to 0 \,, \quad h \to 0 \,.$$

Set $\Psi_{t'} := \Phi_{t'} - \Phi_t$:

$$||R_h|| = \left|\left|\frac{e^{-ihA} - I}{h}\Psi_{t'}\right|\right| \leq \left|\left|\frac{e^{-ihA} - I}{h}\Psi_{t'} + iA\Psi_{t'}\right|\right| + ||iA\Psi_{t'}|| \,. \tag{10.13}$$

The last term tends to zero as $h \to 0$ ($t' \to t$), since $\tau \mapsto ||AS_\tau||$ is continuous:

$$||A\Psi_{t'}|| = \left|\left|A\int_t^{t'} e^{\tau iA} S_\tau d\tau\right|\right| \leq \int_t^{t'} ||e^{\tau iA} AS_\tau|| d\tau = \int_t^{t'} ||AS_\tau|| d\tau \to 0 \text{ as } t' \to t.$$

The first term on the right in (10.13), using the spectral measure of Ψ_t, reads:

$$\sqrt{\int_{\mathbb{R}} \lambda^2 \left|\frac{e^{-ih\lambda}-1}{h\lambda} + i\right|^2 d\mu_{\Psi_{t'}}(\lambda)}.$$

Since:

$$\left|\frac{e^{-ih\lambda}-1}{h\lambda} + i\right|^2 = \left(1 + 2\frac{1-\cos h\lambda}{h\lambda} - 2\frac{\sin h\lambda}{h\lambda}\right) < 5\,,$$

we have

$$\left|\left|\frac{e^{-ihA}-I}{h}\Psi_{t'} + iA\Psi_{t'}\right|\right| \leq \sqrt{5}\sqrt{\int_{\mathbb{R}} \lambda^2 d\mu_{\Psi_{t'}}(\lambda)} = \sqrt{5}||A\Psi_{t'}|| \to 0\,, \quad t' \to t$$

as seen above.

So we proved u_t is a solution. Eventually we need to show it belongs to $C^1(J;D(A))$. By the equation and the assumptions on S_t, that means $t \mapsto Au_t$ is continuous. By definition of u_t, known properties of integrals in a PVM and (10.12) it follows:

$$Au_t = e^{-itA}Av + e^{-itA}\int_0^t e^{i\tau A} AS_\tau d\tau\,.$$

The map $t \mapsto e^{-itA}(Av)$ is clearly continuous, while

$$\left|\left|e^{-it'A}\int_0^{t'} e^{i\tau A}AS_\tau d\tau - e^{-itA}\int_0^t e^{i\tau A}AS_\tau d\tau\right|\right|$$

$$\leq \left|\left|\int_t^{t'} e^{i\tau A}AS_\tau d\tau\right|\right| + \left|\left|\left(e^{-it'A} - e^{-itA}\right)\int_0^t e^{i\tau A}AS_\tau d\tau\right|\right|$$

$$\leq \int_t^{t'} ||AS_\tau|| d\tau + \left|\left|\left(e^{-it'A} - e^{-itA}\right)\int_0^t e^{i\tau A}AS_\tau d\tau\right|\right| \to 0$$

as $t' \to t$, since $t \mapsto ||AS_t||$ is continuous by assumption and $s \mapsto e^{isA}$ is strongly continuous. $\square$

Example 10.9. Under the hypotheses of the previous theorem, suppose $S_t := e^{\alpha t}\psi$, with $\psi \in D(A)$ and $\alpha \in \mathbb{R}\setminus\{0\}$ a given constant. The Cauchy problem (10.8) is solved by:

$$u_t = e^{-itA}v + i(A - i\alpha I)^{-1}(e^{-itA} - e^{\alpha t}I)\psi.$$

The resolvent $(A - i\alpha I)^{-1}$ is a well-defined operator in $\mathfrak{B}(\mathsf{H})$ because $\sigma(A) \subset \mathbb{R}$. To arrive at

$$\int_0^t e^{i\tau A}e^{\alpha\tau}\psi d\tau = i(iA - i\alpha I)^{-1}(I - e^{itA+\alpha I})\psi\,,$$

implying the formula, notice that by definition

$$\left(\phi\left|\int_0^t e^{i\tau A}e^{\alpha\tau}\psi d\tau\right.\right) = \int_0^t \left(\phi\left|e^{i\tau A}e^{\alpha\tau}\psi\right.\right)d\tau = \int_0^t\int_{\mathbb{R}} e^{i\tau(\lambda+\alpha)}e^{\alpha\tau}d\mu_{\phi,\psi}d\tau\,.$$

As the complex measure $\mu_{\phi,\psi}$ is finite, $[0,t]$ has finite measure and $[0,t]\times\mathbb{R}\ni(\tau,\lambda)\mapsto e^{i\tau(\lambda+\alpha)}$ is bounded, we can swap the two integrals by the theorem of Fubini–Tonelli (after decomposing $\mu_{\phi,\psi} = h|\mu_{\phi,\psi}|$, $|h|=1$). Thus by Theorem 9.4:

$$\left(\phi\left|\int_0^t e^{i\tau A}e^{\alpha\tau}\psi d\tau\right.\right) = \int_0^t \left(\phi\left|e^{i\tau A}e^{\alpha\tau}\psi\right.\right)d\tau = \int_{\mathbb{R}}\int_0^t e^{i\tau(\lambda-i\alpha)}d\tau d\mu_{\phi,\psi}$$

$$= \int_{\mathbb{R}} i(\lambda-i\alpha)^{-1}(1-e^{i\tau(\lambda-i\alpha)})d\mu_{\phi,\psi} = \left(\phi\left|i(A-i\alpha I)^{-1}(I-e^{itA+\alpha I})\,\psi\right.\right)\,,$$

whence the claim, as ϕ is arbitrary. ■

10.1.2 The abstract Klein–Gordon/d'Alembert equation (with source and dissipative term)

The second equation we study is the *Klein–Gordon* equation, where again we will assume to have a source, plus a *dissipative* term proportional to the time derivative by a positive coefficient. We shall not return to it at a later stage, so the study begins and ends here. Yet it has to be remembered that the equation has great importance in quantum field theory. Assuming a certain parameter (the mass, in physics) vanishes and in absence of dissipation, the equation goes under the name of *D'Alembert* equation and describes small deformations of (any kind of) waves in linear media. Under the lens of standard PDE theory, the Klein–Gordon equation (with dissipative term and source, as well) is *hyperbolic* in nature. Its structure (ignoring the important physical meaning of the coefficients) is the following: given $J\subset\mathbb{R}$ an open interval and $\Omega\subset\mathbb{R}^n$ an open set the equation reads

$$\frac{\partial^2}{\partial t^2}u_t(x) + 2\gamma\frac{\partial}{\partial t}u_t(x) + (A_0u_t)(x) = S(t,x) \tag{10.14}$$

where, on some domain $D(A_0)\subset C^2(\Omega)$,

$$A_0 = -\Delta_x + m^2 : D(A_0)\to L^2(\Omega,dx) \tag{10.15}$$

$m>0$, $\gamma\geq 0$ constants, $V:\Omega\to\mathbb{R}$ and $S:J\times\Omega\to\mathbb{C}$ given functions, for instance continuous, and Δ_x the Laplace operator on $\mathbb{R}^n$. D'Alembert's equation arises from setting $m=0$, $\gamma=0$ in (10.14)–(10.15). One can consider equations where m and γ are functions, too.

A map $u=u(t,x)$ is a **classical solution** to (10.14) if it is defined on $(t,x)\in J\times\Omega$, twice differentiable with continuity in every variable and it solves the equation on its domain

We can make the same comments of the previous section about A_0. Supposing $A:=\overline{A_0}$ is self-adjoint, we can reinterpret equation (10.14), where now A_0 is replaced by any self-adjoint operator, here positive definite, and the t-derivative is in the topology of the Hilbert space.

Fix an open interval $J \subset \mathbb{R}$ with $J \ni 0$. Let $A : D(A) \to \mathsf{H}$ be self-adjoint on the Hilbert space H, $J \ni t \mapsto S_t \in \mathsf{H}$ a given map in $C(J;\mathsf{H})$, $\gamma > 0$ a constant. The abstract **Klein–Gordon equation** with source and dissipative term reads:

$$\frac{d^2}{dt^2}u_t + 2\gamma\frac{d}{dt}u_t + Au_t = S_t \tag{10.16}$$

where $u \in C^2(J;D(A))$ is the unknown function. Derivatives are defined with respect to H. The *source* is the function $S = S_t$ and the *dissipative term* is the one multiplied by $\gamma \geq 0$. If $S_t = 0$ for any $t \in J$, equation (10.16) is *homogeneous*.

The **Cauchy problem for the Klein–Gordon equation** with dissipative term, resp. with source or homogeneous, is the problem that seeks a solution $u \in C^2(J;D(A))$ to (10.16) with source or homogeneous, subject to **initial conditions**:

$$u_0 = v \in D(A)\,, \quad \frac{du_t}{dt}|_{t=0} = v' \in \mathsf{H}\,. \tag{10.17}$$

Remark 10.10. (1) If A_0 is of type (10.15) and essentially self-adjoint, we may take a classical solution $u = u(t,x)$ to (10.14), for which $u(t,\cdot) \in D(A_0)$ for any $t \in J$. Under assumptions of the kind of Proposition 10.3 for the first and second time derivatives, the solution also satisfies the abstract equation (10.16), as $D(A) = D(\overline{A_0}) \supset D(A_0)$. Therefore, under not-so-strong assumptions, classical solutions are solutions in the abstract sense.
(2) The abstract approach presented allows for operators A *different* from self-adjoint extensions of Laplacians. The abstract equation befits important situations in physics, like waves created by small deformations of elastic media with inner tensions at rest (a violin's sound board): A is a self-adjoint extension of the *squared* Laplacian Δ^2, which is a *fourth-order* differential operator. Allowing for dissipative term and source, the classical equation governing the deformation $u = u(t,x)$ is:

$$a\frac{\partial^2 u}{\partial t^2} + b\Delta_x^2 u + c\frac{\partial u}{\partial t} = S(t,\mathbf{x})$$

for $a,b > 0$, $c \geq 0$. ■

Our first result establishes uniqueness for the abstract Klein–Gordon with any given initial condition, provided A, beside being positive, does not have zero as an eigenvalue. These assumptions are automatic for operators like (10.15), and working on reasonable domains such as $\mathscr{D}(\mathbb{R}^n)$. Note that closing the operator might trigger a new null eigenvalue.

Proposition 10.11. *If $u = u_t$ solves the homogeneous equation (10.16), with $\gamma \geq 0$, $A \geq 0$, $Ker(A) = \{0\}$, then the* **energy estimate**

$$\frac{dE[u_t]}{dt} \leq -4\gamma\left|\left|\frac{du_t}{dt}\right|\right|^2 \tag{10.18}$$

holds, where the "energy of the solution at time t" is:

$$E[u_t] := \left|\left|\frac{du_t}{dt}\right|\right|^2 + (u_t|Au_t)\,. \qquad (10.19)$$

Hence, if a solution $J \ni t \mapsto u_t$ to (10.16) exists ($S_t \neq 0$ in general), it is uniquely determined, for $t \in [0,+\infty) \cap J$, by u_0 and $du_t/dt|_{t=0}$. If $\gamma = 0$ the solution is unique everywhere on J.

Proof. By continuity of the inner product:

$$\frac{d}{dt}E_t[u_t] = \left(\frac{d^2u_t}{dt^2}\left|\frac{du_t}{dt}\right.\right) + \left(\frac{du_t}{dt}\left|\frac{d^2u_t}{dt^2}\right.\right) + \frac{d}{dt}\left(u_t\,|Au_t\,\right)\,.$$

The last derivative is the limit as $h \to 0$ of

$$\frac{1}{h}\left((u_{t+h}\,|Au_{t+h}\,) - (u_t\,|Au_t\,)\right)$$

$$= \frac{1}{h}\left((u_{t+h}\,|Au_{t+h}\,) - (u_{t+h}\,|Au_t\,)\right) - \frac{1}{h}\left((u_{t+h}\,|Au_t\,) - (u_t\,|Au_t\,)\right)\,.$$

The last term, by inner product continuity, tends to

$$\left(\frac{du_t}{dt}\right|\left.Au_t\right)\,.$$

Further, by (10.16):

$$\frac{1}{h}\left((u_{t+h}\,|Au_{t+h}\,) - (u_{t+h}\,|Au_t\,)\right) = \frac{1}{h}\left((Au_{t+h}\,|u_{t+h}\,) - (Au_{t+h}\,|u_t\,)\right)$$
$$= \left(Au_{t+h}\left|\frac{u_{t+h}-u_t}{h}\right.\right) = -\left(\left(\frac{d^2}{dt^2}+2\gamma\frac{d}{dt}\right)u_{t+h}\left|\frac{u_{t+h}-u_t}{h}\right.\right)\,.$$

As $t \mapsto u_t$ is in $C^2(J;D(A))$ and the inner product is continuous,

$$\frac{1}{h}\left((u_{t+h}\,|Au_{t+h}\,) - (u_{t+h}\,|Au_t\,)\right) \to -\left(\left(\frac{d^2}{dt^2}+2\gamma\frac{d}{dt}\right)u_t\left|\frac{du_t}{dt}\right.\right)$$

as $h \to 0$. Therefore we obtain (10.18):

$$\frac{d}{dt}E[u_t] = \left(\frac{d^2u_t}{dt^2}\left|\frac{du_t}{dt}\right.\right) + \left(\frac{du_t}{dt}\left|\frac{d^2u_t}{dt^2}\right.\right) - \left(\left(\frac{d^2}{dt^2}+2\gamma\frac{d}{dt}\right)u_t\left|\frac{du_t}{dt}\right.\right)$$
$$-\left(\frac{du_t}{dt}\left|\left(\frac{d^2}{dt^2}+2\gamma\frac{d}{dt}\right)u_t\right.\right) = -4\gamma\left(\frac{du_t}{dt}\left|\frac{du_t}{dt}\right.\right) \leq 0\,.$$

Consider now two solutions to equation (10.16) with source, and suppose they have the same initial data. The difference of the solutions, $u = u_t$, solves the homogeneous equation, hence also (10.18). By construction $u_0 = 0$, $du_t/dt|_{t=0} = 0$, and the function on the right in (10.18) is continuous. Therefore, for any $t \geq 0$:

$$E[u_t] \leq E[u_0] = 0\,,$$

where we used $u_0 = 0$ and $du_t/dt|_{t=0} = 0$. As $E[u_t] \geq 0$ by definition (10.19), we conclude that

$$E[u_t] = 0 \quad \text{if } t \geq 0\,.$$

Definition (10.19) implies $(u_t|Au_t) = 0$, so by theorem (9.4) $(\sqrt{A}u_t|\sqrt{A}u_t) = 0$, i.e. $u_t \in Ker(\sqrt{A})$, if $t \geq 0$ (recall $D(A) \subset D(\sqrt{A})$ for any self-adjoint operator $A \geq 0$, by definition of $D(f(A))$). If we had $\sqrt{A}u_t = 0$, then $\sqrt{A}\sqrt{A}u_t = 0$ i.e. $Au_t = 0$, which is impossible unless $u_t = 0$. So $u_t = 0$ when $t \geq 0$, and the two solutions coincide for $t \geq 0$. If $\gamma = 0$ the argument works for $t < 0$ as well, by flipping the sign of t to $-t$. □

We are interested in having global existence on J. We will establish a result in the homogeneous case with "small" dissipative term, when $\sigma(A)$ is bounded from below by a positive constant and restricting the initial condition v'.

Proposition 10.12. *Let $\gamma \geq 0$ be given, and assume $A - \gamma^2 I \geq \varepsilon I$ for some $\varepsilon > 0$. Given initial conditions (10.17) with $v \in D(A)$, $v' \in D(\sqrt{A})$, the homogeneous Cauchy problem (10.16)-(10.17) admits a solution, for $t \in J$:*

$$u_t = \frac{e^{-\gamma t}}{2}\left(e^{it\sqrt{A-\gamma^2 I}} + e^{-it\sqrt{A-\gamma^2 I}}\right)v$$
$$-i\frac{e^{-\gamma t}}{2}\left(e^{it\sqrt{A-\gamma^2 I}} - e^{-it\sqrt{A-\gamma^2 I}}\right)(A-\gamma^2 I)^{-\frac{1}{2}}(v'+\gamma v) \tag{10.20}$$

where the exponential and the root are spectrally meant.

Proof. A direct computation shows the right-hand side of (10.20) solves (10.16): for this we need Theorem 9.29, the fact that $(A-\gamma^2 I)^{-\frac{1}{2}}$ is bounded and defined on the whole Hilbert space, and $D(A) = D(A-\gamma^2 I) \subset D(\sqrt{A-\gamma^2 I}) = D(\sqrt{A})$ by the assumptions made. By Proposition 10.11 the solution found is unique, because $A \geq 0$ and $Ker(A) = \{0\}$ from the lower bound $\gamma^2 + \varepsilon > 0$ of $\sigma(A)$.

That u_t is C^1 (as it should) descends from a computation of the derivative, which needs Stone's theorem, and the boundedness of $(A-\gamma^2 I)^{-1/2}$ (it has bounded spectrum). Thus

$$\frac{du_t}{dt} = -\gamma u_t + i\frac{e^{-\gamma t}}{2}\left(e^{it\sqrt{A-\gamma^2 I}} - e^{-it\sqrt{A-\gamma^2 I}}\right)\sqrt{A-\gamma^2 I}v$$
$$i\frac{e^{-\gamma t}}{2}\left(e^{it\sqrt{A-\gamma^2 I}} + e^{-it\sqrt{A-\gamma^2 I}}\right)(v'+\gamma v)\,,$$

is continuous since: u_t is continuous as differentiable, and the rest of du_t/dt is the action of strongly continuous one-parameter groups on given vectors (plus an extra continuous factor $e^{-\gamma t}$).

That u_t is in $C^2(J;\mathsf{H})$ goes as follows: write d^2u_t/dt^2 as combination of u_t, du_t/dt, Au_t and S_t using the differential equation, and recall u_t, du_t/dt and S_t are continuous together with:

$$Au_t = \frac{e^{-\gamma t}}{2}\left(e^{it\sqrt{A-\gamma^2 I}} + e^{-it\sqrt{A-\gamma^2 I}}\right)Av$$
$$-i\frac{e^{-\gamma t}}{2}\left(e^{it\sqrt{A-\gamma^2 I}} - e^{-it\sqrt{A-\gamma^2 I}}\right)A(A-\gamma^2 I)^{-\frac{1}{2}}(v'+\gamma v)\,.$$

As before, in fact, the above is the action of strongly continuous one-parameter groups on fixed vectors. Eventually, from the expression of u_t and du_t/dt we see the initial conditions are satisfied. □

Remarks 10.13. A more suggestive way to write (10.20), legitimate by recalling the meaning of function of an operator A, is:

$$u_t = e^{-\gamma t} \cos\left(t\sqrt{A-\gamma^2 I}\right) v + e^{-\gamma t} \sin\left(t\sqrt{A-\gamma^2 I}\right)(A-\gamma^2 I)^{-1/2}(v' + \gamma v)\,. \tag{10.21}$$

■

Example 10.14. A frequent situation in classical applications is that in which the self-adjoint operator A satisfies $A \geq \varepsilon I$ for some $\varepsilon > 0$ and has *compact resolvent* (by Corollary 8.6 it suffices for this to happen at one point of the resolvent set). The resolvent's spectrum computed at $\varepsilon/2$ (so that to have a self-adjoint operator) is made by eigenvalues only, possibly with 0 as point of the continuous spectrum, and every eigenspace is finite-dimensional by Theorem 4.17. Proposition 9.14 implies $\sigma(A) = \sigma_p(A)$, since $\sigma(A) = \overline{\{\mu^{-1} + \varepsilon/2 \,|\, \mu \in \sigma(R_{\varepsilon/2}(A))\}}$ and every eigenvector of the resolvent $R_{\varepsilon/2}(A)$ is an eigenvector for A.

For example, this is the case when $-A$ is the closure of the Laplacian on the relatively compact open set $\Omega \subset \mathbb{R}^n$, with $D(\Delta)$ containing maps $\psi \in C^2(\Omega)$ whose derivatives up to order two are finite on $\partial\Omega$, and that ψ vanish at the boundary. Then the Laplacian is essentially self-adjoint and the closure's resolvent is compact. If $c > 0$ is constant (the travelling speed of waves in the medium), the equation

$$\frac{d^2 u_t}{dt^2} - c^2 \overline{\Delta} u_t = 0$$

presides over the evolution of the vertical deformation $u_t(x)$ of a flat horizontal elastic membrane represented by the region $\Omega \subset \mathbb{R}^2$, assumed with fixed rim.

Let $A := -c^2\overline{\Delta}$, and call $\{\phi_\lambda\}_{\lambda \in \sigma_p(A)}$ an eigenvector basis for A. Decompose the initial conditions v, v':

$$v = \sum_{\lambda \in \sigma_p(A)} c_\lambda \phi_\lambda\,, \quad v' = \sum_{\lambda \in \sigma_p(A)} c'_\lambda \phi\,.$$

Using (10.21), with $\gamma = 0$, produces the explicit solution:

$$u_t = \sum_{\lambda \in \sigma_p(A)} \left(c_\lambda \cos\left(\sqrt{\lambda} t\right) + c'_\lambda \frac{\sin\left(\sqrt{\lambda} t\right)}{\sqrt{\lambda}} \right) \phi_\lambda\,. \tag{10.22}$$

The solution clearly oscillates by the system's **natural frequecies** (or **eigenfrequencies**: the numbers $\sqrt{\lambda}$, for $\lambda \in \sigma_p(A)$. ■

It should be clear that the solution set to the Klein–Gordon equation *with source* and dissipative term (10.16) – if not empty – consists of maps

$$J \ni t \mapsto u_t^{(0)} + s_t\,,$$

where: s is an arbitrary, fixed, solution to (10.4), while $u^{(0)}$ varies in the vector space of solutions to the *homogeneous* equation (possibly with dissipative term). This solution exists if the source is regular enough. In fact the following analogue to Theorem 10.8 holds.

Theorem 10.15. *Let $A : D(A) \to \mathsf{H}$ be a self-adjoint operator on H, $\gamma \geq 0$ a fixed number, $J \subset \mathbb{R}$ an open interval with $0 \in J$. If*

(i) $A - \gamma^2 I \geq \varepsilon I$ *for some* $\varepsilon > 0$;
(ii) $v \in D(A), v' \in D(\sqrt{A})$;
(iii) $J \ni t \mapsto S_t$ *is continuous in* H;
(iv) $S_t \in D(A)$ *for any* $t \in J$;
(v) $J \ni t \mapsto AS_t$ *is continuous in* H,

then there exists a solution $J \ni t \mapsto u_t \in C^2(J; D(A))$ to the Cauchy problem

$$\begin{cases} \dfrac{d^2u_t}{dt^2} + 2\gamma\dfrac{du_t}{dt} + Au_t = S_t\,, \\ u_0 = v\,, \quad \dfrac{du_t}{dt}|_{t=0} = v'\,, \end{cases} \tag{10.23}$$

of the form:

$$u_t = e^{-\gamma t}\cos\left(t\sqrt{A-\gamma^2 I}\right)v + e^{-\gamma t}\sin\left(t\sqrt{A-\gamma^2 I}\right)(A-\gamma^2 I)^{-1/2}(v'+\gamma v)$$

$$+e^{-\gamma t - it\sqrt{A-\gamma^2 I}}\int_0^t d\tau e^{2i\tau\sqrt{A-\gamma^2 I}}\int_0^\tau dx e^{\gamma x - ix\sqrt{A-\gamma^2 I}}S_x\,. \tag{10.24}$$

The latter is unique on $[0,+\infty) \cap J$, and also on J if $\gamma = 0$.
Integrals in (10.24) are defined using (10.10) repeatedly.

Sketch of proof. Uniqueness was proven earlier, so we have to show

$$u'_t := e^{-\gamma t - it\sqrt{A-\gamma^2 I}}\int_0^t d\tau e^{2i\tau\sqrt{A-\gamma^2 I}}\int_0^\tau dx e^{\gamma x - ix\sqrt{A-\gamma^2 I}}S_x$$

is in $C^2(J; D(A))$ and solves the differential equation with zero initial data. The initial conditions are satisfied by direct computation. The rest is proved applying Theorem 10.8 twice and bearing in mind the following. Since $D(A) = D(A - \gamma^2 I) \subset D(\sqrt{A-\gamma^2 I}) = D(\sqrt{A})$, by Theorem 9.4

$$\frac{d^2u_t}{dt^2} + 2\gamma\frac{du_t}{dt} + Au_t = \left[\frac{d}{dt} - \left(-\gamma I + i\sqrt{A-\gamma^2 I}\right)\right]\left[\frac{d}{dt} - \left(-\gamma I - i\sqrt{A-\gamma^2 I}\right)\right]u_t$$

if $u \in C^2(J; D(A))$. Then the PDE reads :

$$\left[\frac{d}{dt} - \left(-\gamma I + i\sqrt{A-\gamma^2 I}\right)\right]\left[\frac{d}{dt} - \left(-\gamma I - i\sqrt{A-\gamma^2 I}\right)\right]u_t = S_t\,.$$

Theorem 10.8 generalises easily to an operator $A+iaI$ with $a \in \mathbb{R}$, A self-adjoint. The equation thus becomes

$$\left[\frac{d}{dt} - \left(-\gamma I - i\sqrt{A-\gamma^2 I}\right)\right] u_t = e^{-t\gamma I + it\sqrt{A-\gamma^2 I}} \int_0^t e^{\tau\gamma I - i\tau\sqrt{A-\gamma^2 I}} S_\tau d\tau + u_t^{(0)}\,, \tag{10.25}$$

where $u^{(0)}$ denotes the generic homogeneous solution to:

$$\left[\frac{d}{dt} - \left(-\gamma I + i\sqrt{A-\gamma^2 I}\right)\right] u_t^{(0)} = 0\,.$$

Fixing $u^{(0)} = 0$ and iterating for the remaining term on the left in (10.25)

$$\left[\frac{d}{dt} - \left(-\gamma I - i\sqrt{A-\gamma^2 I}\right)\right],$$

produces the solution in the needed form. What is still missing is to check the assumptions granting we can invoke Theorem 10.8: this is left as exercise. □

Examples 10.16. (1) In the hypotheses of Theorem 10.15 let us consider a physical system described by the Klein–Gordon equation with dissipative term, and periodic source

$$S_t = e^{i\omega t}\psi$$

where $\omega \in \mathbb{R}$ is a given constant and $\psi \in D(A)$. Under Theorem 10.15, but explicitly with $\gamma > 0$, we want to study the solution $u = u_t$ of the Cauchy problem with initial conditions v, v' "for large times", meaning $t \gg 1$. A direct computation (see Exercise 10.1) following from (10.24), *if* $\gamma > 0$, yields:

$$u_t = e^{-\gamma t}\left[\cos\left(t\sqrt{A-\gamma^2 I}\right)v + \sin\left(t\sqrt{A-\gamma^2 I}\right)(A-\gamma^2 I)^{-1/2}(v'+\gamma v)\right]$$

$$+e^{-\gamma t}C_{\omega,t}\psi + e^{i\omega t}(A-\omega^2 I + 2i\gamma\omega I)^{-1}\psi$$

for $C_{\omega,t} \in \mathfrak{B}(\mathsf{H})$, $||C_{\omega,t}|| \leq K_\omega$, some constant $K_\omega \geq 0$ and any $t \geq 0$. Assuming $\gamma > 0$ the resolvent of A at $\omega^2 - 2i\gamma$, i.e. $(A-\omega^2 I + 2i\gamma I\omega)^{-1}$, is well defined and in $\mathfrak{B}(\mathsf{H})$, as $\sigma(A) \subset (0,+\infty)$ by assumption. For large $t > 0$ only the last summand in u_t above survives. The term $e^{-\gamma t}C_{\omega,t}\psi$ tends to zero in the norm of H and the part of solution depending on the initial conditions also goes to zero, because:

$$\left|\left|\cos\left(t\sqrt{A-\gamma^2 I}\right)\right|\right| \leq 1 \text{ and } \left|\left|\sin\left(t\sqrt{A-\gamma^2 I}\right)(A-\gamma^2 I)^{-1/2}\right|\right| \leq K\,,$$

for some constant $K \geq 0$. Thus at large times the solution oscillates at the same frequency of the source, and the information provided by the initial data gets lost:

$$||u_t - u_t^{(\infty,\psi,\omega)}|| \to 0 \quad \text{as } t \to +\infty,$$

where we call

$$u_t^{(\infty,\psi,\omega)} := e^{i\omega t}(A-\omega^2 I + 2i\gamma\omega I)^{-1}\psi$$

the **long-time solution**.

(2) Referring to example (1), we explain the phenomenon called *resonance*: if the source oscillates at a frequency ω that is (up to sign) in the spectrum of A, then the smaller the damping term γ is, the larger the long-time solution $u_t^{(\infty,\psi,\omega)}$ can be rendered by choosing a suitable ψ. In fact let $P^{(A)}$ be the PVM of A and $I^\delta_\omega := [\omega^2-\delta, \omega^2+\delta]$, for $\delta > 0$ finite. If the source is given by the unit vector $\psi \in P^{(A)}_{I^\delta_\omega}(\mathsf{H})$, with $\delta > 0$ small enough we have:

$$\left|\left|e^{i\omega}(A-\omega^2 I + 2i\gamma I\omega)^{-1}\psi\right|\right|^2 = \int_{I^\delta_\omega} \frac{d\mu_\psi(\lambda)}{|\lambda-\omega^2|^2+4\gamma^2\omega^2} \geq \inf_{\lambda\in I^\delta_\omega} \frac{1}{|\lambda-\omega^2|^2+4\gamma^2\omega^2}$$

and so:

$$\left|\left|u_t^{(\infty,\psi,\omega)}\right|\right| \geq \frac{1}{\sqrt{\delta^2+4\gamma^2\omega^2}}\,.$$

This is all the more evident if the resolvent of A is compact (see Example 10.14), in which case $\sigma(A) = \sigma_p(A)$. If so, picking $\omega \in \sigma_p(A)$ and ψ a corresponding unit eigenvector, the previous estimate strengthens to:

$$\left|\left|u_t^{(\infty,\psi,\omega)}\right|\right| \geq \frac{1}{2\gamma|\omega|}\,.$$

Continuing with a compact resolvent for A (self-adjoint with strictly positive spectrum), so $\sigma(A) = \sigma_p(A)$, let us take:

$$S_t = \sum_{j\in J} e^{i\omega_j t}\psi_j\,,$$

$\omega_j \in \mathbb{R}$ and $\psi_j \neq 0$, J finite. By linearity the long-time solution will be the superposition:

$$u_t^{(\infty)} = \sum_{j\in J} e^{i\omega_j t}(A-\omega^2 I + 2i\gamma I\omega)^{-1}\psi_j\,.$$

We can decompose every ψ_j using an eigenvector basis $\{\phi_\lambda\}_{\lambda\in\sigma_p(A)}$ for A:

$$\psi_j = \sum_{\lambda\in\sigma_p(A)} c_{\lambda,j}\phi_\lambda\,.$$

As $(A-\omega^2 I + 2i\gamma I\omega)^{-1}$ is continuous and J finite:

$$u_t^{(\infty)} = \sum_{j\in J}\sum_{\lambda\in\sigma_p(A)} \frac{c_{\lambda,j}e^{i\omega_j t}}{\lambda-\omega_j^2+2i\gamma\omega_j}\phi_\lambda\,. \tag{10.26}$$

In contrast to solution (10.22), which arises in absence of source and dissipation, the long-time solution, besides having lost the initial conditions, no longer oscillates by the natural frequecies $\sqrt{\lambda}$ of the system described by A as in (10.22); rather, the oscillations are forced by the frequencies of the source ω_j. However, the system's eigenfrequencies leave traces in the denominator on the right of (10.26), thus generating resonance. That is why the sound of a violin also depends on the harmonic frequencies of the strings despite being produced by the sound board, whose resonance frequencies are all distinct. ∎

10.1.3 The abstract heat equation

In the standard theory of PDEs the heat equation is *parabolic*. Coefficients apart, whose meaning – albeit relevant – we ignore as usual, the heat equation over a given open set $\Omega \subset \mathbb{R}^n$ reads:

$$\frac{\partial}{\partial t} u_t(x) + (A_0 u_t)(x) = S(t,x)\,. \tag{10.27}$$

Above

$$A_0 := -\Delta_x : D(A_0) \to L^2(\Omega, dx) \tag{10.28}$$

for some domain $D(A_0) \subset C^2(\Omega)$, with Δ_x being the Laplace operator on $\mathbb{R}^n$.

A map $u = u(t,x)$ is a **classical solution** to (10.27) if defined for $(t,x) \in [0,b) \times \Omega$, with $b \in (0,+\infty]$ given, continuous on $[0,b) \times \Omega$, differentiable with continuity in t and twice in $x_1,\ldots,x_n$ on $(0,b) \times \Omega$, and clearly if it solves (10.27) on $(0,b) \times \Omega$.

Assuming, as before, $A := \overline{A_0}$ is self-adjoint leads to a generalised interpretation of (10.27), where A_0 is replaced by any self-adjoint operator and t-differentiation is meant in the Hilbert space.

Fix $b \in (0,+\infty]$. If $A : D(A) \to \mathsf{H}$ is self-adjoint on the Hilbert space H and $[0,b) \ni t \mapsto S_t \in \mathsf{H}$ is continuous and fixed in $C((0,b);\mathsf{H})$, the abstract **heat equation** with source is

$$\frac{d}{dt} u_t + A u_t = S_t \tag{10.29}$$

with $u : [0,b) \to D(A)$ continuous and $u \in C^1((0,b);D(A))$ unknown. Continuity and derivative are meant in H. The *source* is the map $S = S_t$. As usual, if $S_t = 0$ for any $t \in [0,b)$ the equation (10.29) is *homogeneous*.

The **Cauchy problem for the heat equation**, resp. with source or homogeneous, seeks a C^1 map $u : [0,b) \to D(A)$ solving (10.29), with source or homogeneous, together with the **initial condition**:

$$u_0 = v \in D(A). \tag{10.30}$$

Remarks 10.17. If A_0 is of the form (10.28) and essentially self-adjoint, we may consider a classical solution $u = u(t,x)$ to (10.27), for which $u(t,\cdot) \in D_0(A_0)$ for any $t \in [0,b)$. Under assumptions of the kind of Proposition 10.3, this solution will solve the abstract equation (10.29) too, since $D(A) = D(\overline{A_0}) \supset D(A_0)$. Therefore with not-so-restrictive hypotheses, classical solutions are abstract solutions. ■

The next result guarantees the abstract heat equation has a unique solution for any initial condition, provided the operator A is positive.

Proposition 10.18. *If $A \geq 0$, and $u = u_t$ solves the homogeneous equation (10.29), then*

$$||u_t|| \leq ||u_0|| \quad \textit{for any } t \in [0,b). \tag{10.31}$$

Hence, for $A \geq 0$, the Cauchy problem (10.29)-(10.30), with $S_t \neq 0$ in general, has at most one solution.

Proof. By (10.1) and (10.29), for $S_t = 0$:

$$\frac{d}{dt}||u_t||^2 = \frac{d}{dt}(u_t|u_t) = -(Au_t|u_t) - (u_t|Au_t) \leq 0$$

as A is self-adjoint. Hence $||u_t|| \leq ||u_{t_0}||$ if $t \geq t_0 \in (0,b)$. By continuity the estimate holds at $t_0 = 0$. Uniqueness is proved as follows. If u, u' solve the Cauchy problem (with same S_t), then $J \ni t \mapsto u_t - u'_t$ solves the Cauchy problem (10.29) with $S_t = 0$ and initial condition $u_0 = 0$. Then $0 \leq ||u_t - u'_t|| \leq ||0|| = 0$ for any $t \in J$ and hence $u_t - u'_t$ for any $t \geq 0$. □

Now we go for an existence result.

Proposition 10.19. *If $A \geq 0$, a (the) solution to Cauchy problem (10.29)-(10.30) reads:*

$$u_t = e^{-tA}v\,, \quad t \in [0,b)\,,$$

where the exponential is meant in spectral sense.

Proof. Take $\psi \in \mathsf{H}$ and $t, t' \in [0,b)$ and observe, as $\sigma(A) \subset [0,+\infty)$, Theorem 9.4 implies $e^{-tA} \in \mathfrak{B}(\mathsf{H})$ if $t \geq 0$, $e^{-0A} = I$ and also:

$$||e^{-tA}\psi - e^{-t'A}\psi||^2 = \int_{\sigma(A)} |e^{-\lambda t} - e^{-\lambda t'}|^2 d\mu_\psi(\lambda) = \int_{[0,+\infty)} |e^{-\lambda t} - e^{-\lambda t'}|^2 d\mu_\psi(\lambda)\,.$$

Since μ_ψ is finite, and bounding the integrand with a constant, uniformly in t, t' in a given neighbourhood, dominated convergence forces $||e^{-tA}\psi - e^{-t'A}\psi||^2 \to 0$ as $t \to t'$. Consequently $u_t := e^{-tA}v$ is continuous on $[0,+\infty)$, and in particular $u_0 = v$. Suppose $v \in D(A)$, so

$$\int_{[0,+\infty)} |\lambda|^2 d\mu_v(\lambda) < +\infty.$$

Assume as well

$$\int_{[0,+\infty)} \lambda^2 e^{-2\lambda t} d\mu_v(\lambda) < +\infty\,,$$

justified by $\lambda \geq 0$ and $t \geq 0$, so $0 \leq e^{-\lambda t} \leq 1$. Then for $t \in (0,b)$:

$$\int_{[0,+\infty)} \left| \frac{e^{-\lambda(t+h)} - e^{-\lambda t}}{h} - \lambda \right|^2 d\mu_v(\lambda) \to 0 \quad \text{as } h \to 0,$$

because the integrand tends pointwise to 0 as $h \to 0$ and is bounded, uniformly around $h = 0$, by the map $[0,+\infty)\lambda \mapsto C\lambda^2 e^{-2\lambda t}$ (integrable if $v \in D(A)$). Thus we proved

$$\left|\left| \frac{1}{h}(e^{-(t+h)A} - e^{-tA})v - Av \right|\right|^2 \to 0 \quad \text{as } h \to 0,$$

if $t \in (0,+\infty)$. That means $u_t := e^{-tA}v$ solves (10.29) for $t \in (0,b)$. Recalling Theorem 9.4(f):

$$\left|\left| \frac{du_t}{dt}(t) - \frac{du_t}{dt}(t') \right|\right|^2 = \int_{[0,+\infty)} |\lambda|^2 |e^{-t\lambda} - e^{-t'\lambda}|^2 d\mu_v(\lambda)\,;$$

so with a similar argument involving Lebesgue's dominated convergence and using $\lambda e^{-t\lambda} \leq 1/t$, we immediately find $\frac{du_t}{dt}(t)$ is continuous in $(0,+\infty)$. Consequently $u \in C^1((0,b);D(A))$, as needed. □

Remark 10.20. (1) If $v \notin D(A)$, we may still define $u_t := e^{-tA}v$, since the domain of the unitary operator e^{-tA} is H. The map $[0,b) \ni t \mapsto u_t$ *does not* solve the homogeneous heat equation. Since $(e^{-tA})^* = e^{-tA}$, though:

$$\frac{d}{dt}(z|u_t) + (Az|u_t) = 0 \quad \text{for any } z \in D(A),\, t \in (0,b). \tag{10.32}$$

The map $[0,b) \ni t \mapsto u_t$ is called a **weak solution** to the homogeneous heat equation.
(2) In case $A \geq 0$ is self-adjoint, the set of operators $T_t := e^{-tA}, t \geq 0$, is a **strongly continuous semigroup of operators** (see Remark 9.34) generated by the self-adjoint A. Put otherwise the functions $[0,+\infty) \ni t \mapsto T_t$, beside strongly continuous (cf. above), satisfy $T_0 = I$ and $T_tT_s = T_{t+s}$, $t,s \in [0,+\infty)$.

From $\lambda^n e^{-\lambda t} \leq C_n t^{-n}$, with $C_n := n^n e^{-n}$, $n \geq 0$, $t,\lambda > 0$, we obtain, for a self-adjoint $A \geq 0$ on H:

$$||A^n T_t \psi||^2 = \int_{[0,+\infty)} |\lambda^n e^{-\lambda t}|^2 d\mu_\psi(\lambda) \leq C_n^2 t^{-2n}$$

for any unit $\psi \in \mathsf{H}$ (bearing in mind Theorem 9.4(c)). Therefore:

$$||A^n T_t|| \leq C_n t^{-n}\,.$$

In particular:

$$Ran(T_t\psi) \subset D(A^n) \quad \text{for any } n = 0,1,2,\ldots,\ \psi \in \mathsf{H} \text{ and } t > 0.$$

When A is the closure of $-\Delta$ on $\mathscr{S}(\mathbb{R}^n)$, say, then

$$\psi_t := e^{-t\overline{\Delta}}\psi \in D(\overline{\Delta}^n) \quad \text{for any } n = 0,1,2,\ldots,\ \psi \in L^2(\mathbb{R}^n,dx) \text{ and } t > 0.$$

Using the *Fourier-Placherel transform* (see Section 3.6), we obtain easily that ψ_t admits *weak derivatives* (Definition 5.24) of any order, and the latter belong to $L^2(\mathbb{R}^n,dx)$. Well-known results of Sobolev (cf. [Rud91], always with $t > 0$) imply ψ_t is in $C^\infty(\mathbb{R}^n)$ up to a zero-measure set; on the other hand $\psi_t \to \psi$ as $t \to 0^+$ in $L^2(\mathbb{R}^n,dx)$. In this sense semigroups generated by elliptic operators like $-\Delta$ are employed to *regularise* functions.
(3) It should be clear, once again, that the solutions to the heat equation *with source* (10.29) – if any at all – have the form

$$J \ni t \mapsto u_t^{(0)} + s_t$$

where: s is any fixed solution to (10.29) and $u^{(0)}$ roams the vector space of *homogeneous* solutions. ■

10.2 Hilbert tensor products

We shall see in Chapter 13 that composite quantum systems are described on *tensor products* of the Hilbert spaces of the component subsystems. We will explain in the sequel what exactly a tensor product in the category of Hilbert spaces is, and assume the reader has a familiarity with (standard) tensor products of finite-dimensional vector spaces [Lan10] concerning the general motivations and the notations used. For the infinite-dimensional case we follow the approach of [ReSi80].

10.2.1 Tensor product of Hilbert spaces and spectral properties

Consider n (complex) Hilbert spaces $(\mathsf{H}_i, (\cdot|\cdot)_i)$, $i = 1,2,\cdots,n$, and choose a vector v_i in each H_i. Mimicking the finite-dimensional situation one can define the *tensor product* of the v_i, $v_1 \otimes \cdots \otimes v_n$, as the *multilinear functional*:

$$v_1 \otimes \cdots \otimes v_n : \mathsf{H}'_1 \times \cdots \times \mathsf{H}'_n \ni (f_1, \cdots, f_n) \mapsto f_1(v_1) \cdots f_n(v_n) \in \mathbb{C}\,,$$

where H'_i is the topological dual to H_i and the dot $\cdot$, on the right, is the product of two complex numbers. *Equivalently*, by Riesz's theorem, we may define the action of $v_1 \otimes \cdots \otimes v_n$ on n-tuples in $\mathsf{H}_1 \times \cdots \times \mathsf{H}_n$ rather than in $\mathsf{H}'_1 \times \cdots \times \mathsf{H}'_n$. This keeps track of the anti-isomorphism (built from the inner product) that identifies dual Hilbert spaces. In this manner $v_1 \otimes \cdots \otimes v_n$ acts on n-tuples $(u_1, \cdots, u_n) \in \mathsf{H}_1 \times \cdots \times \mathsf{H}_n$ via the scalar products, and induces an *antilinear* functional in each variable. This latter, more viable, definition will be our choice.

Definition 10.21. *Consider n (complex) Hilbert spaces* $(\mathsf{H}_i, (\cdot|\cdot)_i)$*,* $i = 1,2,\cdots,n$*, and pick one* v_i *in each* H_i*. The* **tensor product** $v_1 \otimes \cdots \otimes v_n$ *of* $v_1, \ldots, v_n$*, is the mapping:*

$$v_1 \otimes \cdots \otimes v_n : \mathsf{H}_1 \times \cdots \times \mathsf{H}_n \ni (u_1, \cdots, u_n) \mapsto (u_1|v_1)_1 \cdots (u_n|v_n)_n \in \mathbb{C}\,. \tag{10.33}$$

With $T_{i=1}^n \mathsf{H}_i$ *we denote the collecion of maps* $\{v_1 \otimes \cdots \otimes v_n \mid v_i \in \mathsf{H}_i\,,\, i = 1,2,\cdots,n\}$ *while* $\widetilde{\bigotimes_{i=1}^n} \mathsf{H}_i$ *is the* $\mathbb{C}$*-vector space spanned by finitely many tensor products* $v_1 \otimes \cdots \otimes v_n \in T_{i=1}^n \mathsf{H}_i$*.*

Remarks 10.22. From this definition it is evident that the mapping $v_1 \otimes \cdots \otimes v_n : \mathsf{H}_1 \times \cdots \times \mathsf{H}_n \to \mathbb{C}$ is *conjugate*-multilinear, that is antilinear in each variable $u_i \in \mathsf{H}_i$ separately, as we see from (10.33), since the inner product is conjugate-linear on the left. This notwithstanding, $(v_1, \ldots v_n) \mapsto v_1 \otimes \cdots \otimes v_n$ is multi-linear, as one proves immediately. ■

Let us show how we may define an inner product $(\cdot|\cdot)$ on $\widetilde{\bigotimes_{i=1}^n} \mathsf{H}_i$. Consider the map $S : T_{i=1}^n \mathsf{H}_i \times T_{i=1}^n \mathsf{H}_i \to \mathbb{C}$,

$$S(v_1 \otimes \cdots \otimes v_n, v'_1 \otimes \cdots \otimes v'_n) := (v_1|v'_1) \cdots (v_n|v'_n)\,.$$

The following result holds.

Proposition 10.23. *The mapping* $S : T_{i=1}^n \mathsf{H}_i \times T_{i=1}^n \mathsf{H}_i \to \mathbb{C}$ *extends by linearity in the right slot and antilinearity in the left, to a unique Hermitian inner product on the complex space* $\widetilde{\bigotimes}_{i=1}^n \mathsf{H}_i$*:*

$$(\Psi|\Phi) := \sum_i \sum_j \overline{\alpha_i}\beta_j S(v_{1i} \otimes \cdots \otimes v_{ni}, u_{1j} \otimes \cdots \otimes u_{nj})$$

for $\Psi = \sum_i \alpha_i v_{1i} \otimes \cdots \otimes v_{ni}$ *and* $\Phi = \sum_j \beta_j u_{1j} \otimes \cdots \otimes u_{nj}$ *(both sums being finite).*

Proof. Just to preserve readability let us reduce to the case $n = 2$. If $n > 2$ the proof is conceptually identical, only more tedious to write. Take $\Psi, \Phi \in \mathsf{H}_1 \widetilde{\otimes} \mathsf{H}_2$. First we have to show that given *distinct* decompositions of the *same* vectors

$$\Psi = \sum_j \alpha_j v_j \otimes v'_j = \sum_h \beta_h u_h \otimes u'_h\,, \quad \Phi = \sum_k \gamma_k w_k \otimes w'_k = \sum_s \delta_s z_s \otimes z'_s\,,$$

forces

$$\sum_j \sum_k \overline{\alpha_j}\gamma_k S(v_j \otimes v'_j, w_k \otimes w'_k) = \sum_j \sum_s \overline{\alpha_j}\delta_s S(v_j \otimes v'_j, z_s \otimes z'_s) \tag{10.34}$$

and:

$$\sum_h \sum_k \overline{\beta_h}\gamma_k S(u_h \otimes u'_h, w_k \otimes w'_k) = \sum_h \sum_s \overline{\beta_h}\delta_s S(u_h \otimes u'_h, z_s \otimes z'_s)\,. \tag{10.35}$$

This would prove that the (anti)linear extension of S to $\mathsf{H}_1 \widetilde{\otimes} \mathsf{H}_2$ is *well defined*, for it does not depend on the particular representatives S acts on. So let us prove this fact just for the right variable (10.34), because for (10.35) the argument is similar. The left-hand side in (10.34) may be written:

$$\sum_j \sum_k \overline{\alpha_j}\gamma_k S(v_j \otimes v'_j, w_k \otimes w'_k) = \sum_j \left(\sum_k \gamma_k w_k \otimes w'_k\right)(\alpha_j v_j, v'_j) = \sum_j \Phi(\alpha_j v_j, v'_j)$$

and, analogously, the right side of (10.34) reads

$$\sum_j \sum_s \overline{\alpha_j}\delta_s S(v_j \otimes v'_j, z_s \otimes z'_s) = \sum_j \left(\sum_s \delta_s z_s \otimes z'_s\right)(\alpha_j v_j, v'_j) = \sum_j \Phi(\alpha_j v_j, v'_j)\,,$$

where we used $\Phi = \sum_k \gamma_k w_k \otimes w'_k = \sum_s \delta_s z_s \otimes z'_s$. Therefore S extends uniquely to a map, linear in the second argument and antilinear in the first, $(\cdot|\cdot) : \mathsf{H}_1 \widetilde{\otimes} \mathsf{H}_2 \to \mathbb{C}$. By definition of S:

$$(\Psi|\Phi) = \overline{(\Phi|\Psi)}\,.$$

To prove $(\cdot|\cdot)$ is indeed a Hermiitian scalar product we just have to show positive definiteness. That is easy. If $\Psi = \sum_{j=1}^n \alpha_j v_j \otimes v'_j$, where $n < +\infty$ by assumption, consider the (finite!) basis $u_1, \cdots, u_m$ $(m \le n)$ on the span of $v_1, \cdots, v_n$, and a similar basis $u'_1, \cdots, u'_l$, $(l \le n)$ on the span of $v'_1, \cdots, v'_n$. Using the bilinearity of $\otimes$, we can write

$\Psi = \sum_{j=1}^{m}\sum_{k=1}^{l} b_{jk} u_j \otimes u'_k$, for suitable coefficients b_{jk}. By definition di S and since the bases are orthonormal, we obtain

$$(\Psi|\Psi) = \left(\sum_{j=1}^{m}\sum_{k=1}^{l} b_{jk} u_j \otimes u'_k \middle| \sum_{i=1}^{m}\sum_{s=1}^{l} b_{is} u_i \otimes u'_s \right) = \sum_{j=1}^{n}\sum_{k=1}^{l} |b_{jk}|^2 .$$

Now it is patent that $(\cdot|\cdot)$ is positive definite. □

Definition 10.24. *Consider n (complex) Hilbert spaces* $(\mathsf{H}_i, (\cdot|\cdot)_i)$, $i = 1, 2, \cdots, n$. *The* **(Hilbert) tensor product** *of the* H_i, $\bigotimes_{i=1}^{n} \mathsf{H}_i$, *also written* $\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$, *is the completion of* $\widetilde{\bigotimes_{i=1}^{n}} \mathsf{H}_i$ *with respect to the inner product* $(\cdot|\cdot)$ *of Proposition 10.23.*

It is immediate to verify that the definition reduces to the elementary (algebraic) one if the spaces H_i are finite-dimensional. Moreover, the next results hold.

Proposition 10.25. *Take n (complex) Hilbert spaces* $(\mathsf{H}_i, (\cdot|\cdot)_i)$ *with bases* $N_i \subset \mathsf{H}_i$, $i = 1, 2, \cdots, n$. *Then*

$$N := \{z_1 \otimes \cdots \otimes z_n \mid z_i \in N_i\,,\, i = 1, 2, \cdots, n\}$$

is a basis for $\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$. *In particular,* $\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$ *is separable if every* H_i *is.*

Proof. By construction N is an orthonormal system (by definition of tensor inner product). We have to prove $< N >$ is dense in $\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$. As linear combinations of $v_1 \otimes \cdots \otimes v_n$ are dense in $\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$, it is enough to show any $v_1 \otimes \cdots \otimes v_n$ can be approximated arbitrarily well by combinations of $z_1 \otimes \cdots \otimes z_n$ in N. To simplify the notation we will reduce again to $n = 2$, since the general case $n > 2$ goes in the same way. For suitable coefficients γ_z and $\beta_{z'}$, we have $v_1 = \sum_{z \in N_1} \gamma_z z$ and $v_2 = \sum_{z' \in N_2} \beta_{z'} z'$, so (Theorem 3.26 and Definition 3.19):

$$||v_1||^2 = \sup \left\{ \sum_{z \in F_1} |\gamma_z|^2 \,\middle|\, F_1 \text{ finite subset of } N_1 \right\} \tag{10.36}$$

and

$$||v_2||^2 = \sup \left\{ \sum_{z' \in F_2} |\beta_{z'}|^2 \,\middle|\, F_2 \text{ finite subset of } N_2 \right\} . \tag{10.37}$$

If $F \subset N_1 \times N_2$ is finite, a direct computation using the orthonormality of $z \otimes z'$ and the definition of inner product on $\mathsf{H}_1 \otimes \mathsf{H}_2$ gives

$$\left|\left| v_1 \otimes v_2 - \sum_{(z,z') \in F} \gamma_z \beta_{z'} z \otimes z' \right|\right|^2 = ||v_1||^2 ||v_2||^2 - \sum_{(z,z') \in F} |\gamma_z|^2 |\beta_{z'}|^2 .$$

Having (10.36) and (10.37) we can make the right-hand side as small as we like by increasing F. This ends the proof. □

Proposition 10.26. *Let $(\mathsf{H}_i, (\cdot|\cdot)_i)$ be (complex) Hilbert spaces, $D_i \subset \mathsf{H}_i$ dense subspaces, $i = 1, 2, \cdots, n$. The subspace $D_1 \otimes \cdots \otimes D_n \subset \mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$, spanned by tensor products $v_1 \otimes \cdots \otimes v_n$, $v_i \in D_i$, $i = 1, \ldots, n$, is dense in $\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_n$.*

Proof. As is by now customary, we prove the claim for $n = 2$. Finite combinations of tensor products $u \otimes v$ are dense in $\mathsf{H}_1 \otimes \mathsf{H}_2$, so it is enough to prove the following: if $\psi := u \otimes v \in \mathsf{H}_1 \otimes \mathsf{H}_2$, there exists a sequence in $D_1 \otimes D_2$ converging to ψ. By construction there exist sequences $\{u_n\}_{n\in\mathbb{N}} \subset D_1$ and $\{v_n\}_{n\in\mathbb{N}} \subset D_2$ respectively tending to u and v. Then

$$||u_n \otimes v_n - u \otimes v|| \leq ||u_n \otimes v_n - u_n \otimes v|| + ||u_n \otimes v - u \otimes v|| .$$

But $||u_n \otimes v_n - u_n \otimes v||^2 = ||u_n \otimes (v_n - v)||^2 = ||u_n||_1^2 ||v_n - v||_2^2 \to 0$ as $n \to +\infty$, for u_n convergent implies $\{||u_n||_1\}_{n\in\mathbb{N}}$ bounded. Similarly $||u_n \otimes v - u \otimes v||^2 = ||(u_n - u) \otimes v||^2 = ||u_n - u||_1^2 ||v||_2^2 \to 0$ as $n \to +\infty$. □

Examples 10.27. (1) We will exemplify Hilbert tensor products by showing that for separable L^2 spaces (the spaces of wavefunctions in QM), the Hilbert tensor product may be understood alternatively using product measures.

Consider a pair of *separable* Hilbert spaces $L^2(\mathsf{X}_i, \mu_i)$, $i = 1, 2$, and assume both measures are *σ-finite*, so that the product $\mu_1 \otimes \mu_2$ is well defined on $\mathsf{X}_1 \times \mathsf{X}_2$.

Proposition 10.28. *Let $L^2(\mathsf{X}_i, \mu_i)$ be separable Hilbert spaces, $i = 1, 2$, with σ-finite measures. Then*

$$L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2) \quad \textit{and} \quad L^2(\mathsf{X}_1, \mu_1) \otimes L^2(\mathsf{X}_2, \mu_2)$$

are canonically isomorphic Hilbert spaces.
The unitary identification is the unique linear bounded extension of:

$$U_0 : L^2(\mathsf{X}_1, \mu_1) \otimes L^2(\mathsf{X}_2, \mu_2) \ni \psi \otimes \phi \mapsto \psi \cdot \phi \in L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2)$$

where $(\psi \cdot \phi)(x, y) := \psi(x)\phi(y)$, $x \in \mathsf{X}_1$, $y \in \mathsf{X}_2$.

Proof. First, if $N_1 := \{\psi_n\}_{n\in\mathbb{N}}$ and $N_2 := \{\phi_n\}_{n\in\mathbb{N}}$ are bases of $L^2(\mathsf{X}_1, \mu_1)$ and $L^2(\mathsf{X}_2, \mu_2)$ respectively, then $N := \{\psi_n \cdot \phi_m\}_{(n,m)\in\mathbb{N}\times\mathbb{N}}$ is a basis in $L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2)$: N is clearly an orthonormal system by elementary properties of the product measure, and if $f \in L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2)$ is such, for any $\psi_n \cdot \phi_m$, that

$$\int_{\mathsf{X}\times\mathsf{X}_2} \overline{f(x,y)} \psi_n(x)\phi_m(y) d\mu_1(x) \otimes d\mu_2(y) = 0 ,$$

by Fubini–Tonelli we have:

$$\int_{\mathsf{X}_2} \left(\int_{\mathsf{X}_1} \overline{f(x,y)} \psi_n(x) d\mu_1(x) \right) \phi_m(y) d\mu_2(y) = 0 .$$

As the ϕ_m form a basis

$$\int_{\mathsf{X}_1} \overline{f(x,y)} \psi_n(x) d\mu_1(x) = 0 ,$$

except for a set $S_m \subset \mathsf{X}_2$ of zero measure in μ_2. Then for $y \notin S := U_{m\in\mathbb{N}} S_m$ (of zero measure as *countable* union of zero-measure sets):

$$\int_{\mathsf{X}_1} \overline{f(x,y)}\, \psi_n(x) d\mu_1(x) = 0$$

for any $\psi_n \in N_1$, which implies $f(x,y) = 0$ except for $x \in B$, B having zero measure for μ_1. Overall $f(x,y) = 0$, with the exception of the points in $S \times B$, of zero measure for $\mu_1 \otimes \mu_2$ by elementary properties of product measure. Viewing f as in $L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2)$, we then have $f = 0$. Consequently N is a basis, being an orthonormal system with trivial orthogonal complement.

Consider the unique bounded linear function U mapping the basis element $\psi_n \otimes \phi_m$ of $L^2(\mathsf{X}_1, \mu_1) \otimes L^2(\mathsf{X}_2, \mu_2)$ to the basis element $\psi_n \cdot \phi_m$ of $L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2)$. By construction U is unitary. Moreover, U sends $\psi \otimes \phi \in L^2(\mathsf{X}_1, \mu_1) \otimes L^2(\mathsf{X}_2, \mu_2)$ to the corresponding $\psi \cdot \phi \in L^2(\mathsf{X}_1 \times \mathsf{X}_2, \mu_1 \otimes \mu_2)$ (just note $\psi \otimes \phi$ and $\psi \cdot \phi$ have the same components in the respective bases), so U is a linear unitary extension of U_0. Any other linear bounded extension U' of U_0 must reduce to U on bases $\psi_n \otimes \phi_m$, $\psi_n \cdot \phi_m$, and as such it concides with U by linearity and continuity. □

The result clearly generalises to n-fold products of L^2 spaces with separable σ-finite measures.

(2) If $(\mathsf{H}_k, (\cdot|\cdot)_k)$, $k = 1, 2, \ldots, n \le +\infty$, are Hilbert spaces, in general distinct, their **Hilbert sum** $\bigoplus_{k=1}^n \mathsf{H}_k$ is the following Hilbert space. Start from the vector space of n-tuples (*sequences* if $n = +\infty$) $(\psi_1, \psi_2, \ldots) \in \times_{k=1}^n \mathsf{H}_k$ such that $\psi_k = 0$ for $k \ge k_0$, k_0 arbitrarily large but finite and depending on the n-tuple (if $n < +\infty$ this is meaningless, because we can take $k_0 > n$). The linear operations are the usual ones, and the inner product is

$$((\psi_1, \psi_2 \ldots)|(\phi_1, \phi_2 \ldots)) := \sum_{k=1}^n (\psi_k|\phi_k)_k$$

(the sum is finite even in case $n = +\infty$). Then $\bigoplus_{k=1}^n \mathsf{H}_k$ is defined as the Hilbert completion of the above. This agrees with Definition 7.34 when H_k are taken as subspaces of $\bigoplus_{k=1}^n \mathsf{H}_k$, as is easy to see.

Here is another important result about Hilbert tensor products, that deals with the case where all summands H_k of a Hilbert sum, $n < +\infty$, coincide.

Proposition 10.29. *If* H *is a Hilbert space and* $0 < n \in \mathbb{N}$ *is fixed, the Hilbert space* $\mathsf{H} \otimes \mathbb{C}^n$ *is naturally isomorphic to* $\bigoplus_{k=1}^n \mathsf{H}$.
The unitary identification is the unique linear bounded extension of

$$V_0 : \mathsf{H} \otimes \mathbb{C}^n \ni \psi \otimes (v_1, \ldots, v_n) \mapsto (v_1 \psi, \ldots, v_n \psi) \in \bigoplus_{k=1}^n \mathsf{H}\,.$$

Proof. The proof is similar to the one in example (1). Fix a basis $N \subset \mathsf{H}$, so by construction the vectors

$$(\psi, 0, \ldots, 0), (0, \psi, 0, \ldots, 0), \cdots, (0, \ldots, 0, \psi)$$

form a basis of $\bigoplus_{k=1}^{n} \mathsf{H}$ as ψ varies in N. Take the unique linear bounded transformation mapping $\psi \otimes e_i$ to $(0,\dots,\psi,\dots,0)$, where: $\psi \in N$, e_i is the ith canonical vector in $\mathbb{C}^n$, and the only non-zero entry in the n-tuple, ψ, is in the ith place. This is easily unitary, and restricts to V_0 on $\psi \otimes (v_1,\dots,v_n)$. Uniqueness is proved in analogy to example (1). $\square$

(3) The **Fock space** $\mathscr{F}(\mathsf{H})$ generated by H is the infinite Hilbert sum (cf. example (2))

$$\mathscr{F}(\mathsf{H}) := \bigoplus_{n=0}^{+\infty} \mathsf{H}^{n\otimes}$$

where $\mathsf{H}^{0\otimes} := \mathbb{C}$, $\mathsf{H}^{n\otimes} := \underbrace{\mathsf{H} \otimes \cdots \otimes \mathsf{H}}_{n \text{ times}}$. Notice $\mathscr{F}(\mathsf{H})$ is separable if H is. $\blacksquare$

Remarks 10.30. This discussion begs the question whether it makes sense to define a tensor product of *infinitely many* Hilbert spaces. The answer is yes (see [BrRo02, vol.1]). The definition, however, depends on free choices. Consider a collection $\{\mathsf{H}_\alpha\}_{\alpha\in\Lambda}$ of Hilbert spaces (over $\mathbb{C}$) of any cardinality. Fix vectors $U := \{\psi_\alpha\}_{\alpha\in\Lambda}$, $\psi_\alpha \in \mathsf{H}_\alpha$ and $||\psi_\alpha||_\alpha = 1$. We can construct the Hilbert tensor product $\bigotimes_{\alpha\in\Lambda}^{(U)} \mathsf{H}_\alpha$ of as many Hilbert spaces as we like in this way.
(1) Take the subspace in $\times_{\alpha\in\Lambda} \mathsf{H}_\alpha$ of elements $(x_\alpha)_{\alpha\in\Lambda}$ for which only finitely many x_α are distinct from the corresponding ψ_α. Define conjugate-linear maps in each argument $\otimes_\alpha \phi_\alpha : \times_{\alpha\in\Lambda} \mathsf{H}_\alpha \to \mathbb{C}$ by $\otimes_\beta \phi_\beta((x_\alpha)_\alpha) := \Pi_{\alpha\in\Lambda}(x_\alpha|\phi_\alpha)_\alpha$, where, again, only finitely many ϕ_α (depending on $\otimes_\alpha \phi_\alpha$) do not belong in U. Consider the finite span $\widetilde{\otimes^{(U)}}_{\alpha\in\Lambda} \mathsf{H}_\alpha$ of the functionals $\otimes_{\alpha\in\Lambda} \phi_\alpha$.
(2) Define $\otimes_{\alpha\in\Lambda}^{(U)} \mathsf{H}_\alpha$ to be the completion of $\widetilde{\otimes^{(U)}}_{\alpha\in\Lambda} \mathsf{H}_\alpha$ in the norm generated by the only inner product such that $(\otimes_\alpha \phi_\alpha | \otimes_\alpha \phi'_\alpha) := \Pi_\alpha(\phi_\alpha|\phi'_\alpha)_\alpha$.

If Λ is finite it is not hard to see the definition reduces to the previous one and does not depend on the choice of U. The latter ceases to hold, in general, if Λ is infinite. $\blacksquare$

10.2.2 Tensor product of operators (typically unbounded) and spectral properties

As final mathematical topic we discuss the *Hilbert tensor product of operators*. If A and B are operators with domains $D(A)$ and $D(B)$ in the respective Hilbert spaces H_1 and H_2, we will denote by $D(A) \otimes D(B) \subset \mathsf{H}_1 \otimes \mathsf{H}_2$ the subspace of finite combinations of $\psi \otimes \phi$, with $\psi \in D(A)$, $\phi \in D(B)$. Let us try to define an operator:

$$A \otimes B : D(A) \otimes D(B) \to \mathsf{H}_1 \otimes \mathsf{H}_2$$

by extending linearly $\psi \otimes \phi \mapsto (A\psi) \otimes (B\phi)$. The question is whether it is well defined. So suppose, for $\Psi \in D(A) \otimes D(B)$, to have two (finite!) decompositions

$$\Psi = \sum_k c_k \psi_k \otimes \phi_k = \sum_j c'_k \psi'_j \otimes \phi'_j .$$

We have to check that

$$\sum_k c_k(A\psi_k)\otimes(B\phi_k) = \sum_j c'_j(A\psi'_j)\otimes(B\phi'_j)\,.$$

Take a basis (finite!) of vectors f_r for the joint span of the ψ_k and the ψ'_j, and a similar basis g_s for the span of ϕ_k and ϕ'_j. In particular,

$$\psi_i\otimes\phi_i = \sum_{r,s}\alpha^{(i)}_{rs} f_r\otimes g_s\,,\quad \psi'_j\otimes\phi'_j = \sum_{r,s}\beta^{(j)}_{rs} f_r\otimes g_s\,.$$

Having started with a single Ψ decomposed in different ways, necessarily

$$\sum_i c_i\alpha^{(i)}_{rs} = \sum_j c'_j\beta^{(j)}_{rs}\,.$$

From these identities we obtain

$$A\otimes B\left(\sum_i c_i\psi_i\otimes\phi_i\right) = \sum_{rs}(\sum_i c_i\alpha^{(i)}_{rs})((Af_r)\otimes(Bg_s))$$

$$= \sum_{rs}(\sum_j c'_j\beta^{(j)}_{rs})((Af_r)\otimes(Bg_s)) = A\otimes B\left(\sum_j c'_j\psi'_j\otimes\phi'_j\right)\,,$$

making $A\otimes B$ well defined indeed. The procedure extends trivially to N operators $A_k : D(A_k)\to\mathsf{H}_k$, with domains $D(A_k)$ contained in the H_k.

Definition 10.31. *If $A_k : D(A_k)\to\mathsf{H}_k$, $k = 1,2,\ldots,N$, are operators with domain $D(A_k)\subset\mathsf{H}_k$ Hilbert space, the* **(Hilbert) tensor product** *of the A_k is the unique operator*

$$A_1\otimes\cdots\otimes A_N : D(A_k)\otimes\cdots\otimes D(A_N)\to\mathsf{H}_1\otimes\cdots\otimes\mathsf{H}_N$$

extending

$$A_1\otimes\cdots\otimes A_N(v_1\otimes\cdots\otimes v_N) = (A_1v_1)\otimes\cdots\otimes(A_Nv_N)\quad \textit{for } v_k\in D(A_k),\ k=1,\ldots,N.$$

In view of the applications the next elementary result is useful.

Proposition 10.32. *Let $A_k : D(A_k)\to\mathsf{H}_k$, $k = 1,\ldots,N$, be operators on Hilbert spaces H_k. Then for any k:*

(a) *If $\overline{D(A_k)} = \mathsf{H}_k$ and A_k is closable, the operators*

$$I\,,\quad A_1\otimes I\otimes\cdots\otimes I\,,\quad \ldots,\ I\otimes\cdots\otimes A_k\otimes\cdots\otimes A_h\otimes\cdots\otimes I\,,\quad \ldots, A_1\otimes\cdots\otimes A_n$$

and their finite combinations, defined on $D(A_k)\otimes\cdots\otimes D(A_N)$, are all closable.
(b) *If $D(A_k) = \mathsf{H}_k$ and $A_k\in\mathfrak{B}(\mathsf{H}_k)$:*

(i) $A_1\otimes\cdots\otimes A_N\in\mathfrak{B}(\mathsf{H}_1\otimes\cdots\otimes\mathsf{H}_N)$;
(ii) $||A_1\otimes\cdots\otimes A_N|| = ||A_1||\cdots||A_N||$.

Proof. (a) Let us study $n = 2$, the rest being completely analogous. Note $D(A_1) \otimes D(A_2)$ is dense by construction (use Proposition 10.26), so the operators in (a) have adjoints. The generic $\Psi \in D(A_1^*) \otimes D(A_2^*)$, by definition, satisfies $(\Psi|A_1 \otimes A_2 \Phi) = (A_1^* \otimes A_2^* \Psi|\Phi)$ for any $\Phi \in D(A_1) \otimes D(A_2)$. By definition of adjoint

$$D(A_1^*) \otimes D(A_2^*) \subset D((A_1 \otimes A_2)^*) .$$

Apply Theorem 5.5(b), for A_1, A_2 densely defined and closable, to the effect that the adjoints are densely defined and so $D((A_1 \otimes A_2)^*)$ is, too. Thus $A_1 \otimes A_2$ is closable. For linear combinations the argument is the same. Claim (b) is proved in the exercise section. □

At this point we wish to consider *polynomials of operators* $A_1 \otimes \cdots \otimes A_N$, when A_k are self-adjoint [ReSi80]. In the ensuing statement, the arguments A_k of the polynomial Q should more precisely understood as $I \otimes \cdots \otimes I \otimes A_k \otimes I \otimes \cdots \otimes I$, but we will simplify the otherwise cumbersome notation.

Theorem 10.33. *Let $A_k : D(A_k) \to \mathsf{H}_k$, $D(A_k) \subset \mathsf{H}_k$, $k = 1,2,\ldots,N$, be self-adjoint operators and $Q(a_1,\ldots,a_n)$ a real polynomial of degree n_k in the kth variable. Let $D_k \subset D(A_k)$ be a domain where $A_k^{n_k}$ is essentially self-adjoint.*

(a) *$Q(A_1,\ldots,A_1)$ is essentially self-adjoint on $\bigotimes_{k=1}^N D(A_k^{n_k})$ and on $\bigotimes_{k=1}^N D_k$.*
(b) *If every H_k is separable, the spectrum of $\overline{Q(A_1,\ldots,A_N)}$ is the closure of the range of Q in the product of the spectra of the A_k:*

$$\sigma\left(\overline{Q(A_1,\ldots A_N)}\right) = \overline{Q(\sigma(A_1),\ldots,\sigma(A_N))} .$$

Proof. (a) The operator $Q(A_1,\ldots,A_n)$ is well defined on $D := \otimes_{k=1}^N D(A_k^{n_k})$ (in particular by Theorem 9.4(d)) and symmetric, by a direct computation involving the definition of tensor product and the fact that Q has real coefficients and every A_k^m, $m \le n_k$ is symmetric on $D(A_k^{n_k})$. More can be said: $Q(A_1,\ldots,A_n)$ is essentially self-adjoint on D, by Nelson's Theorem 5.47, for we can exhibit a set of analytic vectors for $Q(A_1,\ldots,A_n)$ whose span is dense in the overall Hilbert space. Keeping in mind Example 10.27(1), a set of analytic vectors is given by tensor products $\psi_{\alpha_L}^{(L,1)} \otimes \cdots \otimes \psi_{\alpha_L}^{(L,N)}$, $L = 1,2,\ldots$, where $\{\psi_{\alpha_L}^{(L,k)}\}_{\alpha_L \in G_L} \subset D(A_k)$ is a basis for the closed subspace $P^{(A_k)}([-L,L) \cap \sigma(A_k))$, and $P^{(A_k)}$ is the spectral measure of A_k. Above, when passing from $[-L,L]$ to $[-L-1,L) \cup [-L,L) \cup [L,L+1)$, we must care to keep the same basis for the subsapce associated to $[-L,L)$. That those vectors are analytic for $Q(A_1,\ldots,A_n)$ is easy, just replicating Proposition 9.21(f). To prove $Q(A_1,\ldots,A_n)$ is essentially self-adjoint on $D^{(e)} := \otimes_{k=1}^N D_k$ it suffices to prove the inclusion $\overline{Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}} \supset Q(A_1,\ldots,A_n)\restriction_D$ (by construction, in fact, $Q(A_1,\ldots,A_n)\restriction_{D^{(e)}} \subset Q(A_1,\ldots,A_n)\restriction_D$ so $\overline{Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}} \subset \overline{Q(A_1,\ldots,A_n)\restriction_D}$; if, further, $\overline{Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}} \supset Q(A_1,\ldots,A_n)\restriction_D$, then $\overline{Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}} = \overline{Q(A_1,\ldots,A_n)\restriction_D}$ and the right side is self-adjoint, whence $Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}$ is essentially self-adjoint being symmetric with self-adjoint closure).

To prove $\overline{Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}} \supset Q(A_1,\ldots,A_n)\restriction_D$ assume $\otimes_{k=1}^N \phi_k \in D$; then $\phi_k \in D(A_k^{n_k})$, and being D_k the domain of essential self-adjointness of $A_k^{n_k}$, there exists a sequence $\{\phi_k^l\}_{l\in\mathbb{N}}$ with $\phi_k^l \to \phi_k$ and $A_k^{n_k}\phi^l \to A_k^{n_k}\phi_k$, as $l \to +\infty$. A simple estimate involving the spectral decomposition of A_k tells $A_k^m\phi^l \to A_k^m\phi_k$, $l \to +\infty$, when $1 \le m \le n_k$. Therefore $\otimes_{k=1}^N \phi_k^l \to \otimes_{k=1}^N \phi_k$ and $Q(A_1,\ldots,A_N)(\otimes_{k=1}^N \phi_k^l) \to Q(A_1,\ldots,A_N)(\otimes_{k=1}^N \phi_k)$ as $l \to +\infty$. The results generalises to finite combinations of $\otimes_{k=1}^N \phi_k$, which implies $\overline{Q(A_1,\ldots,A_n)\restriction_{D^{(e)}}} \supset Q(A_1,\ldots,A_n)\restriction_D$.

Let us prove (b). Using Theorem 9.15(c) and remembering the separability of each H, we represent every A_k via a multiplication operator by F_k on the Hilbert space $L^2(M_k,\mu_k)$ identified with H_k. We recall $\otimes_k^N L^2(M_k,\mu_k)$ is isomorphic to $L^2(\times_{k=1}^N M_k,\mu)$, $\mu = \otimes_{k=1}^N \mu_k$, as we saw in Example 10.27(1). In that correspondence $Q(A_1,\ldots,A_N)$ on D is mapped to the multiplication by $Q(F_1,\ldots,F_N)$, and D corresponds to the span, inside $L^2(\times_{k=1}^N M_k,\mu)$, of finite products $\phi_1(m_1)\cdots\phi_N(m_N)$ such that $F_k^{n_k}\cdot\phi_k \in L^2(M_k,\mu_k)$.

Suppose $\lambda \in \overline{Q(\sigma(A_1),\ldots,\sigma(A_N))}$. If $I \ni \lambda$ is an open interval, $Q^{-1}(I) \supset \times_{k=1}^N I_k$ for some open interval $I_k \subset \mathbb{R}$, so that $I_k \cap \sigma(A_k) \neq \varnothing$ for any $k = 1,2,\ldots,N$. Notice $\sigma(A_k) = \text{ess ran}(F_k)$, by Exercise 9.5. Therefore $\mu_k(F_k^{-1}(I_k)) \neq 0$, and so $\mu[Q(F_1,\ldots,F_N)^{-1}(I)] \neq 0$. This means $\lambda \in \text{ess ran}Q(F_1,\ldots,F_N) = \sigma(\overline{Q(A_1,\ldots,P_N)})$ by Exercise 9.5. Conversely, if $\lambda \notin \overline{Q(\sigma(A_1),\ldots,\sigma(A_N))}$ the function $(\lambda - Q(F_1,\ldots,F_N)) : \times_{k=1}^N M_k \to \mathbb{R}$ is bounded, hence $\lambda \in \rho(\overline{Q(A_1,\ldots,A_N)})$, i.e. $\lambda \notin \sigma(\overline{Q(A_1,\ldots,A_N)})$. □

10.2.3 An example: the orbital angular momentum

The observables corresponding to the three Cartesian components of the orbital angular momentum of a particle, in QM, are the unique self-adjoint extensions of the operators:

$$\mathscr{L}_1 := X_2\restriction_{\mathscr{S}(\mathbb{R}^3)} P_3\restriction_{\mathscr{S}(\mathbb{R}^3)} - X_3\restriction_{\mathscr{S}(\mathbb{R}^3)} P_2\restriction_{\mathscr{S}(\mathbb{R}^3)}\,,$$

$$\mathscr{L}_2 := X_3\restriction_{\mathscr{S}(\mathbb{R}^3)} P_1\restriction_{\mathscr{S}(\mathbb{R}^3)} - X_1\restriction_{\mathscr{S}(\mathbb{R}^3)} P_3\restriction_{\mathscr{S}(\mathbb{R}^3)}$$

$$\mathscr{L}_3 := X_1\restriction_{\mathscr{S}(\mathbb{R}^3)} P_2\restriction_{\mathscr{S}(\mathbb{R}^3)} - X_2\restriction_{\mathscr{S}(\mathbb{R}^3)} P_1\restriction_{\mathscr{S}(\mathbb{R}^3)}\,, \tag{10.38}$$

on the Hilbert space $L^2(\mathbb{R}^3,dx)$ (dx being the Lebesgue measure on $\mathbb{R}^3$). Notation-wise, recall X_i and P_i are the operators position and momentum of Section 5.3, while $\mathscr{S}(\mathbb{R}^3)$ is the Schwartz space of smooth complex functions that vanish at infinity faster than any inverse power of $r := \sqrt{x_1^2+x_2^2+x_3^2}$ together with all their derivatives (cf. Section 3.6). In the sequel we will take $D(\mathscr{L}_1) = D(\mathscr{L}_2) = D(\mathscr{L}_3) = \mathscr{S}(\mathbb{R}^3)$ as domain, since $\mathscr{S}(\mathbb{R}^3)$ is invariant under the X_i, P_i (hence under $\mathscr{L}_i$). We will show the **orbital angular momentum operators** $\mathscr{L}_i$ are essentially self-adjoint on the aforementioned domain, and we will find spectrum and a spectral expression for them. In this part we will only see to the mathematical features, reserving any physical consideration for Chapters 11, 12.

We will focus on the operator $\mathscr{L}_3$, because what we will find applies to the other two by rotating coordinates. Explicitly:

$$\mathscr{L}_3 = -i\hbar\left(x_1\frac{\partial}{\partial x_2} - x_2\frac{\partial}{\partial x_1}\right),$$

where x_1, x_2 are viewed as multiplicative operators by the corresponding functions. A fourth operator used in the sequel is the **total angular momentum** (squared):

$$\mathscr{L}^2 := \mathscr{L}_1^2 + \mathscr{L}_2^2 + \mathscr{L}_3^2,$$

defined on $\mathscr{S}(\mathbb{R}^3)$. This, too, is essentially self-adjoint on $\mathscr{S}(\mathbb{R}^3)$. We will compute its spectrum and make the spectral expansion of $L^2 := \overline{\mathscr{L}^2}$ explicit.

In order to find said spectral expansions and spectra, it is better to write the operators in spherical coordinates r, θ, ϕ, where $x_1 = r\sin\theta\cos\phi$, $x_2 = r\sin\theta\sin\phi$, $x_3 = r\cos\theta$, so $r \in (0+\infty)$, $\theta \in (0,\pi)$, $\phi \in (-\pi,\pi)$. In this manner a simple computation produces

$$\mathscr{L}_3 = -i\hbar\frac{\partial}{\partial\phi}, \quad \mathscr{L}^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right], \tag{10.39}$$

where operators acts on functions in $\mathscr{S}(\mathbb{R}^3)$ whose argument has undergone coordinate change. From (10.39) it is evident that the operators do not depend on the radial coordinate r, a fact of the utmost importance. Keeping that in mind, note that $\mathbb{R}^3 = \mathbb{S}^2 \times [0,+\infty)$, where (up to zero-measure sets) the unit sphere $\mathbb{S}^2$ is the domain of θ, ϕ, whilst $[0,+\infty)$ is where r varies; furthermore, Lebesgue's measure on $\mathbb{R}^3$ can be seen as the product

$$dx = d\omega(\theta,\phi) \otimes r^2 dr,$$

where

$$d\omega(\theta,\phi) = \sin\theta d\theta d\phi$$

is the standard measure on $\mathbb{S}^2$ identified with the rectangle $(0,\pi)\times(-\pi,\pi)$ by the spherical angles (θ,ϕ) (the complement to $(0,\pi)\times(-\pi,\pi)$ in $\mathbb{S}^2$ has null $d\omega$-measure, so it does not interfere). Thus we obtain the decomposition:

$$L^2(\mathbb{R}^3, dx) = L^2(\mathbb{S}^2 \times [0,+\infty), d\omega(\theta,\phi) \otimes r^2 dr).$$

By Example 10.27(1):

$$L^2(\mathbb{R}, dx) = L^2(\mathbb{S}^2, d\omega) \otimes L^2((0,+\infty), r^2 dr). \tag{10.40}$$

At this point we introduce operators *on the Hilbert space* $L^2(\mathbb{S}^2, d\omega)$:

$$_{\mathbb{S}^2}\mathscr{L}_3 = -i\hbar\frac{\partial}{\partial\phi}, \quad _{\mathbb{S}^2}\mathscr{L}^2 = -\hbar^2\left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right], \tag{10.41}$$

with domain $C^\infty(\mathbb{S}^2)$. As the sphere is a C^∞ manifold,[1] the space $C^\infty(\mathbb{S}^2)$ of smooth maps on $\mathbb{S}^2$ is dense in $L^2(\mathbb{S}^2,d\omega)$ (exercise). These operators are Hermitian, hence symmetric, as a simple direct computation reveals. Way before QM's formulation, it was known from the study of the Laplace equation (and classical electrodynamics) that there is a distinguished basis of $L^2(\mathbb{S}^2,d\omega)$ whose elements are called **spherical harmonics** [NiOu82]:

$$Y_l^m(\theta,\phi) := \frac{(-1)^l}{2^l l!}\sqrt{\frac{(2l+1)}{4\pi}\frac{(l+m)!}{(l-m)!}}e^{im\phi}\frac{1}{\sin^m\phi}\frac{d^{l-m}}{d(\cos\theta)^{l-m}}(1-\cos^2\theta)^l\,, \tag{10.42}$$

where:

$$l=0,1,2,\ldots \quad m\in\mathbb{N}\,,|m|\leq 2l+1\,. \tag{10.43}$$

The maps $Y^l_m \in C^\infty(\mathbb{S}^2)$ are notoriously eigenfunctions of the differential operators ${}_{\mathbb{S}^2}\mathscr{L}_3$ and ${}_{\mathbb{S}^2}\mathscr{L}^2$ given in (10.41):

$${}_{\mathbb{S}^2}\mathscr{L}_3 Y^l_m = \hbar\, m Y^l_m\,, \quad {}_{\mathbb{S}^2}\mathscr{L}^2 Y^l_m = \hbar^2\, l(l+1) Y^l_m\,. \tag{10.44}$$

Note how the first is obvious by definition of Y^l_m. In particular the vectors Y^l_m are analytic for the symmetric operators ${}_{\mathbb{S}^2}\mathscr{L}^2$, ${}_{\mathbb{S}^2}\mathscr{L}_3$ defined on $C^\infty(\mathbb{S}^2)$. As Y^l_m form a basis of $L^2(\mathbb{S}^2,d\omega)$, by Nelson's criterion (Theorem 5.47) they warrant essential self-adjointness to the operators ${}_{\mathbb{S}^2}\mathscr{L}^2$, ${}_{\mathbb{S}^2}\mathscr{L}_3$ on $C^\infty(\mathbb{S}^2)$. Following the recipe of Section 9.1.4 concerning the Hamiltonian operator of the one-dimensional harmonic oscillator, we obtain analogue spectral decompositions:

$$\overline{{}_{\mathbb{S}^2}\mathscr{L}^2} = \text{s-}\sum_{l\in\mathbb{N},\, m\in\mathbb{Z},|m|\leq 2l+1} \hbar^2\, l(l+1) Y^l_m (Y^l_m|\ \ ) \quad \text{and} \quad \overline{{}_{\mathbb{S}^2}L_3} = \text{s-}\sum_{l\in\mathbb{N},\, m\in\mathbb{Z},|m|\leq 2l+1} \hbar\, m Y^l_m (Y^l_m|\ \ )\,. \tag{10.45}$$

In this context the spectra read

$$\sigma(\overline{{}_{\mathbb{S}^2}\mathscr{L}^2}) = \sigma_p(\overline{{}_{\mathbb{S}^2}\mathscr{L}^2}) = \left\{\hbar^2\, l(l+1)\ \middle|\ l=0,1,2\ldots\right\}\,, \tag{10.46}$$

and

$$\sigma(\overline{{}_{\mathbb{S}^2}\mathscr{L}_3}) = \sigma_p(\overline{{}_{\mathbb{S}^2}\mathscr{L}_3}) = \{\hbar m \,|\, |m|\leq 2l+1\,, m\in\mathbb{Z},\, l=0,1,2\ldots\}\,. \tag{10.47}$$

Now let us go back to $L^2(\mathbb{R}^3,dx)$. As the space $\mathscr{D}(0,+\infty)$ of smooth maps with compact support in $(0,+\infty)$ is dense in the separable Hilbert space $L^2((0+\infty),r^2dr)$, by Proposition 3.31(b), there will be a basis $\{\psi_n\}_{n\in\mathbb{N}}$ of maps in $\mathscr{D}(0,+\infty)$. Passing to Cartesian coordinates it is easy to see that

$$f_{l,m,n}(x) = Y^l_m(\theta,\phi)\psi_n(r) \tag{10.48}$$

[1] See Appendix B: the idea is to cover $\mathbb{S}^2$ with local charts in θ,ϕ, by rotating the Cartesian axes. Three local frame systems suffice to cover $\mathbb{S}^2$. The functions of $C^\infty(\mathbb{S}^2)$, by definition, go from $\mathbb{S}^2$ to $\mathbb{C}$ and are C^∞ when restricted to any local chart of the sphere.

are $C^\infty(\mathbb{R}^3)$ (the only singularity could be at $x=0$, but around that point the maps vanish by construction). By definition $f_{l,m,n}$ have compact support, so they live in $\mathscr{S}(\mathbb{R}^3)$. By the definitions and domains given,

$$ {}_{\mathbb{S}^2}\mathscr{L}_3 \otimes I\restriction_{\mathscr{D}(0,+\infty)} \subset \mathscr{L}_3 \quad \text{and} \quad {}_{\mathbb{S}^2}\mathscr{L}^2 \otimes I\restriction_{\mathscr{D}(0,+\infty)} \subset \mathscr{L}^2 , $$

so $\{Y^l_m \otimes \psi_n \mid n,l \in \mathbb{N}, |m| \le 2l+1\ , m \in \mathbb{Z}\} \subset \mathscr{S}(\mathbb{R}^3)$ is a basis of $L^2(\mathbb{R}^3,dx) = L^2(\mathbb{S}^2,d\omega)\otimes L^2((0,+\infty),r^2dr)$ by Example 10.27(2). Thinking $\mathscr{L}_3$ and $\mathscr{L}^2$ as acting on $\mathscr{S}(\mathbb{R}^3)$,

$$ \mathscr{L}_3 Y^l_m \otimes \psi_n = \hbar m Y^l_m \otimes \psi_n\,, \quad \mathscr{L}^2 Y^l_m \otimes \psi_n = \hbar^2\, l(l+1) Y^l_m \otimes \psi_n\,. \tag{10.49} $$

Again, $\mathscr{L}^2$, $\mathscr{L}_3$ are essentially self-adjoint on that domain, and their unique self-adjoint extensions $L^2 := \overline{\mathscr{L}^2}$, $L_z := \overline{\mathscr{L}_z}$ decompose spectrally as

$$ L^2 = \text{s-}\sum_{l,n\in\mathbb{N},\, m\in\mathbb{Z}\,, |m|\le 2l+1} \hbar^2\, l(l+1) Y^l_m \otimes \psi_n (Y^l_m \otimes \psi_n|\ ) \ \text{ and } \ L_z = \text{s-}\sum_{l,n\in\mathbb{N},\, m\in\mathbb{Z}\,, |m|\le 2l+1} \hbar m Y^l_m \otimes \psi_n (Y^l_m \otimes \psi_n|\ )\,. \tag{10.50} $$

The spectra of L^2, L_3 remain those of (10.46), (10.47). Note the spectral measures of L^2, L_3 commute.

The same conclusions can be reached using Theorem 10.33 appropriately.

Remarks 10.34. To finish consider two von Neumann algebras $\mathfrak{R}_1$, $\mathfrak{R}_2$ on Hilbert spaces H_1, H_2. There is a corresponding **tensor product of von Neumann algebras** $\mathfrak{R}_1 \otimes \mathfrak{R}_2$. This is the von Neumann algebra on $\mathsf{H}_1 \otimes \mathsf{H}_2$ given by the strong completion in $\mathfrak{B}(\mathsf{H}_1 \otimes \mathsf{H}_2)$ of the *-algebra of finite combinations of products $A_1 \otimes A_2$, with $A_1 \in \mathfrak{R}_1, A_2 \in \mathfrak{R}_2$. The generalisation to finite products is straightforward, while the tensor product of infinitely many von Neumann algebras is more complicated to define [BrRo02].

The important **commutator theorem of tensor products** [BrRo02] asserts that $(\mathfrak{R}_1 \otimes \mathfrak{R}_2)' = \mathfrak{R}_1' \otimes \mathfrak{R}_2'$.

In reference to the largest von Neumann algebras in H_1, H_2 we then have $(\mathfrak{B}(\mathsf{H}_1) \otimes \mathfrak{B}(\mathsf{H}_2))' = \{c_1 I_1 \otimes c_2 I_2\}_{c_1,c_2\in\mathbb{C}} = \{cI\}_{c\in\mathbb{C}}$; since $(\mathfrak{B}(\mathsf{H}_1) \otimes \mathfrak{B}(\mathsf{H}_2))'' = \mathfrak{B}(\mathsf{H}_1) \otimes \mathfrak{B}(\mathsf{H}_2)$ by the double-commutant theorem, we recover the known fact $\mathfrak{B}(\mathsf{H}_1) \otimes \mathfrak{B}(\mathsf{H}_2) = \mathfrak{B}(\mathsf{H}_1 \otimes \mathsf{H}_2)$. ■

10.3 Polar decomposition theorem for unbounded operators

Consider a densely-defined closed operator $A : D(A) \to \mathsf{H}$ on the Hilbert space H. Using the fact that A^*A is self-adjoint and positive (as we will see) and by the spectral theorem for unbounded operators it is possible to define the positive self-adjoint operator $|A| := \sqrt{A^*A}$. Letting $U = A|A|^{-1}$ at least on $Ran(|A|)$, and then extending trivially (as zero) to the complement of $Ran(|A|)$, we immediately find the decomposition

$$ A = U|A|\,. $$

Formally, and without caring too much about domains, $U\upharpoonright_{Ran(|A|)}$ is an isometry. Heuristically, this is a generalisation of Theorem 3.71, that we proved for bounded operators defined on the entire Hilbert space. The hands-on approach is flawed in that is does not say where the polar decomposition should be valid (the domains of A and $|A|$ could be different, *a priori*) and any attempt to formalise the argument soon becomes punishing. That is why we will follow an indirect route based on a more general theorem.

The generalised polar decomposition we will eventually prove plays a crucial part in rigorous quantum fields theory, especially in relationship to the Tomita-Takesaki *modular theory* and in defining *KMS thermal states* [BrRo02].

10.3.1 Properties of operators A^*A, square roots of unbounded positive self-adjoint operators

We proceed in steps, proving first that if A is closed and densely defined, A^*A is self-adjoint and $D(A^*A)$ is a core for A. Then we will show a result that in some sense generalises the polar decomposition theorem, thus specifiying properly the domains involved. At last we will prove the existence and uniqueness of positive self-adjoint square roots of unbounded self-adjoint operators.

Theorem 10.35. *Consider a closed, densely-defined operator $A : D(A) \to \mathsf{H}$ on the Hilbert space H. Then:*

(a) *A^*A, defined on the natural domain $D(A^*A)$ (Definition 5.1), is self-adjoint.*
(b) *The dense subspace $D(A^*A)$ is a core for A:*

$$\overline{A\upharpoonright_{D(A^*A)}} = A\,. \tag{10.51}$$

Proof. For (a), call $I : \mathsf{H} \to \mathsf{H}$ the identity and introduce $I+A^*A$ on its natural domain (coinciding with $D(A^*A)$, by Definition 5.1). We claim there is a positive operator $P \in \mathfrak{B}(\mathsf{H})$ with $Ran(P) = D(I+A^*A)$ and

$$(I+A^*A)P = I\,, \quad P(I+A^*A) = I\upharpoonright_{D(I+A^*A)}\;. \tag{10.52}$$

By Proposition 3.54(f) $P \in \mathfrak{B}(\mathsf{H})$ is self-adjoint as positive. By uniqueness of the inverse the operator $I+A^*A$ coincides with the inverse of P, obtained by spectral decomposition:

$$P^{-1} = \int_{\sigma(P)} \lambda^{-1} dP^{(P)}(\lambda)\,.$$

This is self-adjoint by Theorem 9.4. Thus $A^*A = (I+A^*A) - I$ is self-adjoint on $D(I+A^*A) = D(A^*A)$, consequently dense.

Now we have to exhibit $P \in \mathfrak{B}(\mathsf{H})$ positive, with $Ran(P) = D(I+A^*A)$ and satisfying (10.52). If $f \in D(I+A^*A) = D(A^*A)$ then $Af \in D(A^*)$ by definition of $D(A^*A)$. Hence

$$(f|f) + (Af|Af) = (f|f) + (f|A^*Af) = (f|(I+A^*A)f)\,.$$

We proved $(I+A^*A) \geq 0$, and by Cauchy-Schwarz also $||f||^2 \leq ||f||\,||(I+A^*A)f||$, so $I+A^*A : D(A^*A) \to \mathsf{H}$ is injective. Consider the operator A, closed and densely defined. The identity of Theorem 5.10(d) says that for any $h \in \mathsf{H}$ there are unique $Ph \in D(A)$ and $Qh \in D(A^*)$ such that

$$(0,h) = (-APh, Ph) + (Qh, A^*Qh) \tag{10.53}$$

in $\mathsf{H} \oplus \mathsf{H}$. By construction P, Q are defined on all of H, and the two vectors on the right, seen in $\mathsf{H} \oplus \mathsf{H}$, are orthogonal. By definition of norm on $\mathsf{H} \oplus \mathsf{H}$, the identity also tells:

$$||h||^2 \geq ||Ph||^2 + ||Qh||^2 ,$$

for any $h \in \mathsf{H}$, so $P, Q \in \mathfrak{B}(\mathsf{H})$ because $||P||, ||Q|| \leq 1$. Considering the single components in (10.53), we have

$$Q = AP \quad \text{and} \quad h = Ph + A^*Qh = Ph + A^*APh = (I + A^*A)Ph ,$$

for all $h \in \mathsf{H}$. Hence $(I + A^*A)P = I$ and $P : \mathsf{H} \to D(I + A^*A)$ must be one-to-one, but also onto since we saw $(I + A^*A)$ is injective. The inverse of a bijection is unique, so

$$P(I + A^*A) = I\restriction_{D(I+A^*A)} .$$

Up to now we have proved $P \in \mathfrak{B}(\mathsf{H})$ has range covering $D(I + A^*A)$, and that (10.52) holds. Let us see that $P \geq 0$. If $h \in \mathsf{H}$, then $h = (I + A^*A)f$ for some $f \in D(A^*A)$, so:

$$(Ph|h) = (P(I + A^*A)f|(I + A^*A)f) = (f|(I + A^*A)f) \geq 0 .$$

To finish we prove (b). As A is closed, its graph is closed in $\mathsf{H} \oplus \mathsf{H}$, so a Hilbert space itself. Suppose $(f, Af) \in G(A)$ is orthogonal to $G(A\restriction_{D(A^*A)})$. Then for any $x \in D(A^*A)$:

$$0 = ((f, Af)|(x, Ax)) = (f|x) + (Af|Ax) = (f|x) + (f|A^*Ax) = (f|(I + A^*A)x)) .$$

But $Ran(I + A^*A) = \mathsf{H}$, so $f = 0$ and the orthogonal complement to $G(A\restriction_{D(A^*A)})$ in the Hilbert space $G(A)$ is trivial. Therefore $\overline{G(A\restriction_{D(A^*A)})} = G(A)$. □

Together with the uniqueness for positive roots of (unbounded) positive self-adjoint operators, the next theorem contains, as subcase, the polar decomposition theorem for closed and densely-defined operators. Recall that for a pair P, Q with the same domain D, $P \leq Q$ means $(f|Pf) \leq (f|Qf)$ for any $f \in D$.

Theorem 10.36. *Let $A : D(A) \to \mathsf{H}$, $B : D(B) \to \mathsf{H}$ be closed and densely defined on the Hilbert space H.*

(a) *If*

$$D(A^*A) \supset D(B^*B) \quad \text{and} \quad A^*A\restriction_{D(B^*B)} \leq B^*B , \tag{10.54}$$

then $D(A) \supset D(B)$ and there exists a bounded operator $C : \mathsf{H} \to \mathsf{H}$ uniquely determined by:

$$A\restriction_{D(B)} = CB , \quad Ker(C) \supset Ker(B^*) . \tag{10.55}$$

Furthermore, $||C|| \leq 1$ *and* $C\restriction_{(Ran(B))^\perp} = 0$.
(b) *If*

$$D(A^*A) \supset D(B^*B) \quad and \quad A^*A\restriction_{D(B^*B)} = B^*B\,, \tag{10.56}$$

then $C\restriction_{\overline{Ran(B)}}$ *is an isometry and* $Ker(C) = Ker(B^*)$.
(c) *If*

$$D(A^*A) = D(B^*B) \quad and \quad A^*A = B^*B\,, \tag{10.57}$$

then $D(A) = D(B)$.

Proof. (a) Begin by the uniqueness of C. If C and C' are bounded and $A = CB$, $A = C'B$, then $C - C'$ is the null operator on $Ran(B)$. By continuity $C\restriction_{\overline{Ran(B)}} = C'\restriction_{\overline{Ran(B)}}$. From the splitting $\mathsf{H} = \overline{Ran(B)} \oplus (\overline{Ran(B)})^\perp$, where $(\overline{Ran(B)})^\perp = Ker(B^*)$, having $KerC \supset Ker(B^*)$, $KerC' \supset Ker(B^*)$ implies $C\restriction_{(\overline{Ran(B)})^\perp} = C'\restriction_{(\overline{Ran(B)})^\perp}$. Hence $C = C'$. Let us prove there is a $C \in \mathfrak{B}(\mathsf{H})$ such that $A = CB$ on $D(B)$ (hence $D(B) \subset D(A)$), $Ker(C) \supset Ker(B^*)$, $||C|| \leq 1$ and $C\restriction_{(Ran(B))^\perp} = 0$.

Call A', B' the restrictions of A, B to $D(A^*A)$, $D(B^*B)$ respectively. By the previous theorem these are cores for A, B, so $\overline{Ran(A')} = \overline{Ran(A)}$ and $\overline{Ran(B')} = \overline{Ran(B)}$. Note $Ker(A) = Ker(A')$, $Ker(B) = Ker(B')$, as $D(A^*A) \subset D(A)$ and $D(B^*B) \subset D(B)$.

Let us define an operator such that $A' = CB'$, to begin with. This determines a linear operator C on $Ran(B')$:

$$A'f = CB'f\,, \quad \text{for any } f \in D(B^*B) \subset D(A^*A)\,.$$

For it to be well defined, we need $B'f = B'g$ to imply $A'f = A'g$, i.e. that $B'h = 0$ implies $A'h = 0$ for any $h \in D(B^*B) \subset D(A^*A)$. But the latter is true, in fact: $B'h = 0$ implies $(B'h|B'h) = 0$, so $0 = (B'h|B'h) = (h|B^*Bh) \geq (h|A^*Ah) = (A'h|A'h) = ||A'h||^2 \geq 0$, because $h \in D(B^*B) \subset D(A^*A)$. Hence $A'h = 0$. The claim is C is bounded on $Ran(B')$ with $||C|| \leq 1$. Since $A^*A \leq B^*B$, using $D(A^*A) \subset D(A)$ and $D(B^*B) \subset D(B)$, we have

$$||C(B'f)||^2 = (CB'f|CB'f) = (A'f|A'f) = (f|A^*Af) \leq (f|B^*Bf)$$

$$= (B'f|B'f) = ||B'f||^2\,, \tag{10.58}$$

if $f \in D(B^*B) \subset D(A^*A)$. Therefore C extends uniquely to $\overline{Ran(B')} = \overline{Ran(B)}$, preserving $||C|| \leq 1$. To fully define $C : \mathsf{H} \to \mathsf{H}$ it suffices to know it on the complement $(\overline{Ran(B)})^\perp = Ker(B^*)$. Let C be null there. Thus $C : \mathsf{H} \to \mathsf{H}$ is bounded, $||C|| \leq 1$ and $Ker(C) \supset Ker(B^*)$. By construction, for any $f \in D(B^*B) \subset D(A^*A)$:

$$Af = CBf\,.$$

Since $D(B^*B)$ is a core for B, if $g \in D(B)$ there is a sequence $\{f_n\}_{n\in\mathbb{N}} \subset D(B^*B) \subset D(A^*A)$ such that $f_n \to g$ and $Bf_n \to Bg$. By continuity

$$\lim_{n\to+\infty} Af_n = \lim_{n\to+\infty} CBf_n = C \lim_{n\to+\infty} Bf_n = CBg\,.$$

But A is closed, so $g \in D(A)$ and $\lim_{n\to+\infty} Af_n = Ag$. Hence $A' = CB'$ actually extends to $A = CB$ on $D(B) \subset D(A)$.
(b) Assuming $A^*A = B^*B$ on $D(B^*B) \subset D(A^*A)$, equation (10.58) is replaced by:

$$||C(B'f)||^2 = (CB'f|CB'f) = (A'f|A'f) = (f|A^*Af) = (f|B^*Bf) = (B'f|B'f)$$
$$= ||B'f||^2$$

if $f \in D(B^*B) \subset D(A^*A)$. Therefore C is an isometry on $Ran(B)$ and by continuity on $\overline{Ran(B)}$ as well. There remains to prove $Ker(C) \subset Ker(B^*)$, for the other inclusion is valid in the general case (a). If $s \in Ker(C)$, from $\mathsf{H} = \overline{Ran(B)} \oplus Ker(B^*)$ we have $s = r + n$, $r \in \overline{Ran(B)}$, $n \in Ker(B^*)$. Since $Ker(B^*) \subset Ker(C)$, we obtain $0 = Cs = C(r+n) = Cr + Cn = Cr + 0 = Cr$. On the other hand C is isometric on $\overline{Ran(B)}$, so $0 = ||Cr|| = ||r||$ and $r = 0$. Therefore $s \in Ker(C)$ implies $s = n \in Ker(B^*)$, ending the proof of $Ker(C) \subset Ker(B^*)$.
(c) We show that $D(A) = D(B)$ if $D(A^*A) = D(B^*B)$, $A^*A = B^*B$. From the proof of the more general (a), $D(B) \subset D(A)$. In the present case A and B can be swapped, so $D(A) = D(B)$. □

And now the last ingredient, generalising part of Theorem 3.66.

Theorem 10.37. *Let $A : D(A) \to \mathsf{H}$ be self-adjoint on the Hilbert space H.*
(a) *$A \geq 0$ iff $\sigma(A) \subset [0+\infty)$.*
(b) *If $A \geq 0$, there is a unique self-adjoint operator $B \geq 0$ such that $B^2 = A$, where the left-hand side is defined on its natural domain $D(B^2)$, coinciding with $D(A)$. Using the integral in the spectral measure of A it turns out that $B = \sqrt{A}$.*

Proof. (a) If $\sigma(A) \subset [0+\infty)$, Theorem 9.4(g), referred to the PVM $P^{(A)}$ of A, implies $A \geq 0$. *Vice versa*, suppose $A \geq 0$ and, by contradiction, that there is λ with $0 > \lambda \in \sigma(A)$. If λ were in $\sigma_p(A)$ there would be an eigenvector $\psi \in \mathsf{H} \setminus \{0\}$ for λ, and $0 \leq (\psi|A\psi) = ||\psi||^2\lambda < 0$, impossible. Instead, if $\lambda \in \sigma_c(A)$, by Theorem 9.10 $P^{(A)}_{(a,b)} \neq 0$ for any open interval $(a,b) \ni \lambda$. So we could choose $(a,b) = (2\lambda, \lambda/2)$, getting, for $0 \neq \psi \in P^{(A)}_{(a,b)}(\mathsf{H})$,

$$0 \leq (\psi|A\psi) = \int_{\mathbb{R}} x d\mu_\psi(x) = \int_{(2\lambda,\lambda/2)} x d\mu_\psi(x) \leq \int_{(2\lambda,\lambda/2)} \frac{\lambda}{2} d\mu_\psi(x) = \frac{\lambda}{2}||\psi||^2 < 0,$$

using Theorem 9.4 and that μ_ψ vanishes outside (a,b). This is absurd.
(b) A positive self-adjoint root of A is just

$$B = \sqrt{A} := \int_{\sigma(A)} \sqrt{x} dP^{(A)}(x).$$

The operator is well defined, since $\sigma(A) \subset [0,+\infty)$, it is self-adjoint by Theorem 9.4(b) and $B^2 = A$, where B^2 is defined on its natural domain $D(B^2) = D(A)$ by Theorem 9.4(c–d). At last $B \geq 0$ by Theorem 9.4(g). Let us pass to uniqueness.

Assume $B \geq 0$ is self-adjoint and $B = \int_{[0,+\infty)} x dP^{(B)}(x)$. If $B^2 = A$ with $A \geq 0$, by (9.47) we obtain

$$\int_{[0,+\infty)} x dP^{(A)}(x) = \int_{[0,+\infty)} x^2 dP^{(B)}(x) = \int_{[0,+\infty)} x dP^{(B)}(f^{-1}(x)),$$

where $f(x) = x^2$, $x \geq 0$, so $f^{-1} : [0,+\infty) \to [0+\infty)$ is a well-defined map. The spectral measure of A is unique, so in particular $P^{(B)}(f^{-1}(E')) = P^{(A)}(E')$ for any Borel set $E' \subset [0,+\infty)$. If $E \subset [0,+\infty)$ is a Borel set, $f(E) \subset [0,+\infty)$. Setting $E' = f(E)$ gives $P^{(B)}(E) = P^{(A)}(f(E))$ for any Borel set $E \subset [0,+\infty)$ (and $P^{(B)}(E) = 0$ if $E \subset (-\infty,0)$). Therefore A determines B completely, for it determines its unique PVM. □

10.3.2 Polar decomposition theorem for closed and densely-defined operators

We can finally prove the polar decomposition for closed, densely-defined operators. The idea of the proof is to start, rather than from A, from A^*A. If the polar decomposition is to hold, one expects $A^*A = |A|\,|A|$, with $|A| := \sqrt{A^*A}$ defined spectrally, remembering A^*A is self-adjoint. Now Theorem 10.36(c) yields the required polar decomposition of A. The powerfulness of this approach becomes apparent when considering the properties of the domains involved: usually hard to study by a more direct method, they are now automatic, by Theorem 10.36.

Theorem 10.38. *Let $A : D(A) \to \mathsf{H}$ be closed and defined densely on the Hilbert space H.*

(a) *There exists a pair P, U in H such that:*

(1) *the polar decomposition*

$$A = UP \tag{10.59}$$

holds;

(2) *P is positive, self-adjoint and $D(P) = D(A)$;*

(3) *$U \in \mathfrak{B}(\mathsf{H})$ is isometric on $Ran(P)$;*

(4) *$Ker(U) \supset Ker(P)$.*

(b) *Moreover:*

(i) *$P = |A| := \sqrt{A^*A}$;*

(ii) *$Ker(U) = Ker(P) = Ker(A) = (Ran(P))^{\perp}$ and $\overline{Ran(P)} = \overline{Ran(A^*)}$;*

(iii) *$Ran(U) = \overline{Ran(A)}$.*

Remarks 10.39. Note the U of (10.59) is a *partial isometry* (Definition 3.61) with *initial space*

$$[Ker(U)]^{\perp} = \overline{Ran(P)} = [Ker(A)]^{\perp} = \overline{Ran(A^*)}$$

and *final space*

$$Ran(U) = \overline{Ran(A)}.$$

■

Proof of Theorem 10.38. (a)–(b) We prove uniqueness by finding P explicitly, assuming (10.59) plus (2), (3), (4). First we show $D(A^*A) = D(PP)$. By definition of adjoint, as $U \in \mathfrak{B}(\mathsf{H})$, (10.59) implies $A^* = P^*U^* = PU^*$. Hence $f \in D(A^*A)$ if and only if $f \in D(PU^*UP)$. Splitting H into $\overline{Ran(P)} \oplus Ker(P^*) = \overline{Ran(P)} \oplus Ker(P)$, and since U is isometric on $\overline{Ran(P)}$ and zero on $Ker(P)$, we get $(U^*U)\restriction_{\overline{Ran(P)}} = I\restriction_{\overline{Ran(P)}}$. Hence the claim $f \in D(A^*A)$ iff $f \in D(PU^*UP)$ is equivalent to $f \in D(A^*A)$ iff $f \in D(PP)$. So we have proved $D(A^*A) = D(PP)$. Let us use it towards uniqueness. If $g \in D(A^*A) \subset D(A)$ (i.e $g \in D(PP) \subset D(P)$), recalling U is isometric on $Ran(P)$, then

$$(f|A^*Ag) = (Af|Ag) = (UPf|UPg) = (Pf|Pg) = (f|PPg) \quad \text{for } f \in D(A) = D(P).$$

Being $D(A) = D(P)$ dense, we conclude $A^*A = PP$. Therefore P is a positive self-adjoint root of A^*A, hence unique by Theorem 10.37, and $P = \sqrt{A^*A} =: |A|$. Now we can apply Theorem 10.36(c) with $B = P$ (closed and densely defined, being self-adjoint), to find that U coincides with C. That $Ker(U) = Ker(P) = (Ran(P))^\perp$ follows from part (b) of that theorem, because $B = P = P^* = B^*$ in our case and $(Ran(P^*))^\perp = Ker(P)$. The claim $Ker(A) = Ker(P)$ goes as follows:

$$0 = ||Af||^2 = (Af|Af) = (f|A^*Af) = (f|PPf) = (Pf|Pf) = ||Pf||^2 ,$$

where we used the fact that $Af = 0$ implies $f \in D(A^*A)$ by definition of the latter. Now $\overline{Ran(P)} = \overline{Ran(A^*)}$ follows immediately from the previous properties: $\overline{Ran(P)} = Ker(P)^\perp = Ker(A)^\perp = \overline{Ran(A^*)}$. In the final equality we used Theorem 5.10(c), as A is closed and the domain of A^* dense by the aforementioned part (b). Let us prove $Ran(U) = \overline{Ran(A)}$. By the decomposition $UP = A$ we have $Ran(U) = U(Ran(P)) = Ran(A)$, so $\overline{Ran(U)} = \overline{U(Ran(P))} = \overline{Ran(A)}$. Then it sufficies to show $Ran(U)$ is closed. Let $y \in \overline{Ran(U)} \setminus \{0\}$. There exists $\{x_n\}_{n\in\mathbb{N}} \subset (KerU)^\perp$ with $Ux_n \to y$, $n \to +\infty$. Since $||U(x_n - x_m)|| = ||x_n - x_m||$, $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence. Define $x = \lim_{n\to+\infty} x_n$, so $Ux = y$, $y \in Ran(U)$ and then $Ran(U)$ contains its limit points, i.e. it is closed. □

Corollary 10.40. *In the hypotheses of Theorem 10.38 the operator $U : \mathsf{H} \to \mathsf{H}$ is unitary precisely when A is one-to-one and $Ran(A)$ is dense in* H.
In particular U is unitary and $U = A|A|^{-1}$ in case A is bijective.

Proof. If U is unitary, in particular it is one-to-one and onto, so (ii) and (iii) in (b) imply A is injective and $Ran(A)$ dense in H. Conversely, if A is injective, by (a)–(b) in the theorem above and by continuity of U, we have U isometric on $\overline{Ran(|A|)} = Ker(|A|)^\perp = Ker(A)^\perp = \mathsf{H}$. As $Ran(U) = \overline{Ran(A)}$ in case $Ran(A)$ is dense, we conclude $U : \mathsf{H} \to \mathsf{H}$ is isometric and onto, hence unitary as claimed. If A is further bijective, it is injective and $Ran(A)$ is trivially dense, so U is unitary as seen before. From $A = U|A|$, then, A and U being bijective, we obtain $|A|$ is bijective, so $U = A|A|^{-1}$, ending the proof. □

10.4 The theorems of Kato-Rellich and Kato

The last results we will state and prove are those of Kato-Rellich and Kato. They are extremely useful to study self-adjointness and lower boundedness for operators of Quantum Mechanics (especially the so-called *Hamiltonian operators*), in the framework of *perturbation theory*. The former theorem provides sufficient conditions for an operator of the form $T+V$, a *perturbation* of T, to be self-adjoint, and have lower-bounded spectrum when T has. The latter considers specific situations, where T is the Laplacian on $\mathbb{R}^3$ or $\mathbb{R}^n$. A general treatise, with applications to quantum physics, is [ReSi80], from which several proofs of this section are taken.

10.4.1 The Kato-Rellich theorem

A preliminary definition is in order.

Definition 10.41. *Let $T : D(T) \to \mathsf{H}$ and $V : D(V) \to \mathsf{H}$ be densely-defined operators on the Hilbert space H, with $D(T) \subset D(V)$. If there are $a, b \in [0, +\infty)$ such that*

$$||V\varphi|| \leq a||T\varphi|| + b||\varphi|| \quad \textit{for any } \varphi \in D(T)\,, \tag{10.60}$$

*V is called T-**bounded**. The greatest lower bound of the numbers a satisfying (10.60) for some b is called the* **relative bound** *of V with respect to T. If the relative bound is zero, V is said* **infinitesimally small** *with respect to T.*

Remark 10.42. (1) If T and V are closable, by Definition 5.20) it suffices to verify (10.60) on a core of T.
(2) Equation (10.60) is equivalent to:

$$||V\varphi||^2 \leq a_1^2||T\varphi||^2 + b_1^2||\varphi||^2 \quad \text{for any } \varphi \in D(T)\,, \tag{10.61}$$

In fact, (10.61) implies (10.60) by putting $a = a_1$, $b = b_1$. For the converse, take $a_1^2 = (1+\delta)a^2$, $b_1^2 = (1-\delta^{-1})b^2$, any $\delta > 0$: then (10.60) implies (10.61). ∎

Let us pass to the *Kato-Rellich Theorem*. For a self-adjoint operator $A : D(A) \to \mathsf{H}$, we know $\sigma(A) \subset [M, +\infty)$ if and only if $(\psi|A\psi) \geq M(\psi|\psi)$ for any $\psi \in D(A)$, by Theorem 10.37(a). Therefore statement (c) below may be equivalently expressed in terms of lower-bounded quadratic forms.

Theorem 10.43 (Kato-Relich). *Let $T : D(T) \to \mathsf{H}$ and $V : D(V) \to \mathsf{H}$ be densely-defined operators on the Hilbert space H such that:*

(i) *T is self-adjoint;*
(ii) *V is symmetric;*
(iii) *V is T-bounded with* **relative bound** $a < 1$.

Then:

(a) *$T+V$ is self-adjoint on $D(T)$.*
(b) *$T+V$ is essentially self-adjoint on every core of T.*

(c) *If* $\sigma(T) \subset [M, +\infty)$ *then* $\sigma(T+V) \subset [M', +\infty)$ *where:*

$$M' = M - \max\left\{ \frac{b}{(1-a)}, a|M| + b \right\}, \quad \textit{with a, b satisfying (10.60).}$$

Proof. For (a) we try to apply Theorem 5.18, showing that if we choose $D(T)$ as domain for the symmetric $T+V$, we obtain $Ran(T+V \pm iI) = \mathsf{H}$. Actually we will prove there exists $\nu > 0$ such that

$$Ran(T+V \pm i\nu I) = \mathsf{H},$$

giving the previous by linearity. If $\varphi \in D(T)$, T self-adjoint implies $Ran(T+i\mu I) = \mathsf{H}$ and

$$||(T+i\mu I)\varphi||^2 = ||T\varphi||^2 + \mu^2 ||\varphi||^2 .$$

Setting $\varphi = (T+i\mu I)^{-1}\psi$, gives

$$||T(T+i\mu I)^{-1}|| \leq 1 \quad \text{and} \quad ||(T+i\mu I)^{-1}|| \leq \mu^{-1} .$$

Applying (10.60) with $\varphi = (T+i\mu I)^{-1}\psi$ produces

$$||V(T+i\mu I)^{-1}\psi|| \leq a||T(T+i\mu I)^{-1}\psi|| + b||(T+i\mu I)^{-1}\psi|| \leq \left(a + \frac{b}{\mu}\right) ||\psi|| .$$

If $\mu = \nu$ is large enough, the bounded operator

$$U := V(T+i\nu I)^{-1},$$

defined on H, satisfies $||U|| < 1$, as $a < 1$. This implies $-1 \notin \sigma(U)$ by (iii) in Theorem 8.4(c). By Theorem 8.4(a) (U is closed as bounded), we have $Ran(I+U) = \mathsf{H}$. At the same time, since T is self-adjoint, $Ran(T+i\nu I) = \mathsf{H}$ by Theorem 5.18. Hence

$$(I+U)(T+i\nu I)\varphi = (T+V+i\nu I)\varphi , \quad \varphi \in D(T)$$

implies, as claimed, $Ran(T+V+i\nu I) = \mathsf{H}$. The proof of $Ran(T+V-i\nu I) = \mathsf{H}$ is completely similar, so (a) is over.

Let us pass to (b). Equation (10.60) implies, if $\mathscr{D} \subset D(T)$ is a core for T:

$$D(T) = D\left(\overline{T\restriction_{\mathscr{D}}}\right) \subset D\left(\overline{(T+V)\restriction_{\mathscr{D}}}\right) .$$

On the other hand, by construction and because $T+V$ is self-adjoint on $D(T)$ hence closed:

$$D\left(\overline{(T+V)\restriction_{\mathscr{D}}}\right) \subset D\left(\overline{(T+V)}\right) = D(T+V) = D(T) .$$

Putting all inclusions together produces $D\left(\overline{(T+V)\restriction_{\mathscr{D}}}\right) = D(T+V)$ so $\overline{(T+V)\restriction_{\mathscr{D}}} = T+V$, as $T+V$ is closed. $(T+V)\restriction_{\mathscr{D}}$ is then essentially self-adjoint by Proposition 5.21.

Now (c). By assumption, the spectral theorem implies $\sigma(T) \geq M$ (with obvious notation). Choosing $s > -M$ (with $s \in \mathbb{R}$) gives $\sigma(T+sI) > 0$, so $0 \notin \sigma(T+sI)$. But $T+sI$ is self-adjoint, so it is closed, and by Theorem 8.4(a) $Ran(T+sI) = \mathsf{H}$. The same estimates used before prove $||V(T+sI)^{-1}|| < 1$ if

$$-s < M' := M - \max\left\{\frac{b}{(1-a)}, a|M|+b\right\}.$$

Consequently, for these s:

$$Ran(T+V+sI) = \mathsf{H} \quad \text{and} \quad (T+V+sI)^{-1} = (T+sI)^{-1}(I+U)^{-1},$$

implying $-s \in \rho(T+V)$, and then $-s \notin \sigma(T+V)$. $T+V$ self-adjoint has real spectrum, whence $\sigma(T+V) \geq M'$. □

10.4.2 An example: the operator $-\Delta + V$ and Kato's theorem

Condition (10.60) arises naturally in certain contexts, and is of great use, in physical applications, to study the Schrödinger equation, where the *Laplace operator* Δ is perturbed by a potential V. To discuss this application of the Kato-Relich theorem we begin with a proposition and a lemma.

Proposition 10.44. *Let*

$$\Delta := \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2} \tag{10.62}$$

be the Laplace operator on $\mathbb{R}^n$ thought as operator on $L^2(\mathbb{R}^n, dx)$.

(a) *If $\widehat{\mathscr{F}} : L^2(\mathbb{R}^n, dx) \to L^2(\mathbb{R}^n, dk)$ is the Fourier-Plancherel unitary operator (cf. Section 3.6), Δ is essentially self-adjoint on $\mathscr{S}(\mathbb{R}^n)$, on $\mathscr{D}(\mathbb{R}^n)$ and on $\widehat{\mathscr{F}}(\mathscr{D}(\mathbb{R}^n))$ with the same (unique) self-adjoint extension $\overline{\Delta}$.*
(b) *If $k^2 = k_1^2 + k_2^2 + \ldots + k_n^2$ then*

$$\left(\widehat{\mathscr{F}}\overline{\Delta}\widehat{\mathscr{F}}^{-1} f\right)(k) = -k^2 f(k) \tag{10.63}$$

on the natural domain

$$D(\widehat{\mathscr{F}}\overline{\Delta}\widehat{\mathscr{F}}^{-1}) = \left\{ f \in L^2(\mathbb{R}^n, dk) \,\middle|\, \int_{\mathbb{R}^n} k^4 |f(k)|^2 dk < +\infty \right\}.$$

(c) *The operator $\overline{-\Delta} = -\overline{\Delta}$ is bounded from below:*

$$\sigma(\overline{-\Delta}) \subset [0, +\infty), \quad \text{or equivalently } (\psi|\overline{-\Delta}\psi) \geq 0 \text{ for any } \psi \in D(\overline{-\Delta}). \tag{10.64}$$

Proof. Most of (a) and (b) was proven in Exercises 5.11, 5.12. What we still do not have is that Δ is essentially self-adjoint on $\widehat{\mathscr{F}}(\mathscr{D}(\mathbb{R}^n))$ and has a common self-adjoint extension on $\mathscr{D}(\mathbb{R}^n)$ and $\mathscr{S}(\mathbb{R}^n)$ alike. To this end notice $\widehat{\mathscr{F}}(\mathscr{D}(\mathbb{R}^n)), \mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n)$, so the three extensions coincide because there is one self-adjoint extension to any essentially self-adjoint operator. That Δ is essentially self-adjoint on $\widehat{\mathscr{F}}(\mathscr{D}(\mathbb{R}^n))$, given $\widehat{\mathscr{F}}$ is unitary and (10.63), is equivalent to the essential self-adjointness of the symmetric multiplication by $-k^2$ on $\mathscr{D}(\mathbb{R}^n)$. In turn the latter, in view of Nelson's Theorem 5.47, follows from the observation that every $\varphi = \varphi(k)$ in $\mathscr{D}(\mathbb{R}^n)$ is analytic for the multiplication by $-k^2$, since $||-(k^2)^n\varphi|| \leq ||\varphi||(\sup_{k\in supp\varphi}|k|^2)^n$. Statement (c) descends from (b) and from Theorem 10.37(a). □

Now a fundamental, classical result.

Lemma 10.45. *Fix $n = 1,2,3$ and consider $f \in D(\overline{\Delta})$. Then f coincides almost everywhere with a continuous bounded map, and for any $a > 0$ there exists $b > 0$ independent from f such that:*

$$||f||_\infty \leq a||\overline{\Delta}f|| + b||f||\,. \tag{10.65}$$

Proof. Let us prove (10.66) for $n = 3$, the other cases being similar. Call $\hat{f} := \widehat{\mathscr{F}}f$. By Proposition 3.81(a) and Plancherel's theorem (Theorem 3.84), the claim is true if we manage to prove $\hat{f} \in L^1(\mathbb{R}^3, dk)$, and for any given $a > 0$ there is $b \in \mathbb{R}$ such that:

$$||\hat{f}||_1 \leq a||k^2\hat{f}||_2 + b||\hat{f}||_2\,. \tag{10.66}$$

If $f \in D(\overline{\Delta})$, by Proposition 10.44, $\hat{f} \in D(\widehat{\mathscr{F}}\overline{\Delta}\widehat{\mathscr{F}}^{-1})$, so also $(1+k^2)\hat{f} \in L^2(\mathbb{R}^3, dk)$. Since $(k_1,k_2,k_3) \mapsto 1/(1+k^2)$ belongs to that same space, $\hat{f} \in L^1(\mathbb{R}^3, dk)$ by the Hölder inequality. Moreover:

$$||\hat{f}||_1 \leq c||(1+k^2)\hat{f}||_2 \leq c(||k^2\hat{f}||_2 + ||\hat{f}||_2) \tag{10.67}$$

where $c := \sqrt{\int(1+k^2)^{-1}dk}$. If $r > 0$ define $\hat{f}_r(k) := r^3\hat{f}(rk)$. Then $||\hat{f}_r||_1 = ||\hat{f}||_1$, $||\hat{f}_r||_2 = r^{3/2}||\hat{f}||_2$ and $||k^2\hat{f}_r||_2 = r^{-1/2}||k^2\hat{f}||_2$. Using (10.67) for $\hat{f}_r$ the three previous identities give

$$||\hat{f}||_1 \leq cr^{-1/2}||k^2\hat{f}||_2 + cr^{3/2}||\hat{f}||_2 \quad \text{for any } r > 0.$$

(10.66) holds for $a = cr^{-1/2}$. □

Remarks 10.46. The lemma can be generalised (see [ReSi80, vol. II]) by this statement based on Young's inequality: *Consider $f \in L^2(\mathbb{R}^n, dx)$ with $f \in D(\overline{\Delta})$. If $n \geq 4$ and $2 \leq q < 2n/(n-4)$ then $f \in L^q(\mathbb{R}^n, dk)$, and for any $a > 0$ there exists $b \in \mathbb{R}$ not depending on f (but on q, n, a) such that $||f||_q \leq a||\overline{\Delta}f|| + b||f||$.* ■

We can eventually apply the Kato-Rellich theorem to a very interesting case for Quantum Mechanics, and prove a result due to Kato. Later we will see a more general statement, known in the litarature as Kato's theorem.

Theorem 10.47 (Essential self-adjointness of $-\Delta + V$). *Fix $n = 1,2,3$ and take $V = V_2 + V_\infty$, with $V_2 \in L^2(\mathbb{R}^n, dx)$, $V_\infty \in L^\infty(\mathbb{R}^n, dx)$ real functions.*

(a) *$-\Delta + V$ is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^n)$ and on $\mathscr{S}(\mathbb{R}^n)$.*
(b) *The only self-adjoint extension $\overline{-\Delta + V}$ of the operators of (a) coincides with the (self-adjoint) operator $\overline{-\Delta} + V$ defined on $D(\overline{\Delta})$.*
(c) *$\sigma(\overline{-\Delta + V})$ is bounded from below.*

Proof. As V is real it gives a multiplicative operator on the domain

$$D(V) := \{\varphi \in L^2(\mathbb{R}^n, dx) \mid V\varphi \in L^2(\mathbb{R}^n, dx)\} .$$

Using the definition it is easy to see the operator is self-adjoint. By construction, moreover,

$$||V\varphi||_2 \leq ||V_2||_2||\varphi||_\infty + ||V_\infty||_\infty||\varphi||_2 < +\infty \tag{10.68}$$

for $\varphi \in \mathscr{D}(\mathbb{R}^n)$ or $\varphi \in \mathscr{S}(\mathbb{R}^n)$. Hence $\mathscr{D}(\mathbb{R}^n) \subset \mathscr{S}(\mathbb{R}^n) \subset D(V)$. What is more, since $\mathscr{S}(\mathbb{R}^n) \subset D(\overline{\Delta})$ (by Proposition 10.44), using (10.65) in Lemma 10.45 ($n \leq 3$) we find, for any $a > 0$, a number $b > 0$ such that:

$$||V\varphi||_2 \leq a||V_2||_2\,|| - \Delta\varphi||_2 + (b + ||V_\infty||_\infty)||\varphi||_2 \quad \text{for any } \varphi \in \mathscr{S}(\mathbb{R}^n) .$$

That is say, given $a' > 0$ there is $b' > 0$ with

$$||V\varphi||_2 \leq a'|| - \Delta\varphi||_2 + b'||\varphi||_2 \quad \text{for any } \varphi \in \mathscr{S}(\mathbb{R}^n) \tag{10.69}$$

so in particular for any $\varphi \in \mathscr{D}(\mathbb{R}^n)$. Consequently

$$||V\varphi - V\varphi'||_2 \leq a'\,||(-\Delta\varphi) - (-\Delta\varphi')||_2 + b'||\varphi - \varphi'||_2$$

with φ, φ' in $\mathscr{S}(\mathbb{R}^n)$. V is closed, as self-adjoint, and $\mathscr{S}(\mathbb{R}^n)$ is a core for the self-adjoint (hence closed) $\overline{-\Delta}$ (by Proposition 10.44), so the inequality proves $D(V) \supset D(\overline{-\Delta})$. Exploiting the closure of operators we conclude that (10.69) holds on the entire domain of $\overline{-\Delta}$:

$$||V\varphi||_2 \leq a'||\overline{-\Delta}\varphi||_2 + b'||\varphi||_2 \quad \text{for any } \varphi \in D(\overline{-\Delta}).$$

If we choose $a' < 1$, $T := \overline{-\Delta}$ satisfies the assumptions of Theorem 10.43, with V as we have it now. By Kato-Rellich, using that $\mathscr{S}(\mathbb{R}^n)$ and $\mathscr{D}(\mathbb{R}^n)$ are cores for $\overline{-\Delta}$ by Proposition 10.44, we conclude. □

Remarks 10.48. Remembering Remark 10.46, this theorem generalises to $n > 3$ with these modifications: $V = V_p + V_\infty$ with $V_p \in L^p(\mathbb{R}^n, dx)$, $V_\infty \in L^\infty(\mathbb{R}^n, dx)$, where $p > 2$ for $n = 4$, $p = n/2$ for $n \geq 5$. The proof is analogous. ■

For the classical result known as *Kato's theorem*, we shall interpret $f \in L^p(\mathbb{R}^n, dx) + L^q(\mathbb{R}^n, dx)$ to mean f is the sum of a function in $L^p(\mathbb{R}^n, dx)$ and one in $L^q(\mathbb{R}^n, dx)$.

Theorem 10.49 (Kato). *Fix $n = 1, 2, 3$ and denote by $(\mathbf{y}_1, \ldots, \mathbf{y}_N)$ the elements in $\mathbb{R}^{nN}$, where $\mathbf{y}_k \in \mathbb{R}^n$ for any $k = 1, \ldots, N$. If Δ is the Laplacian (10.62) on $\mathbb{R}^{nN}$, consider the differential operator $-\Delta + V$, V being the multiplicative operator given by:*

$$V(\mathbf{y}_1, \ldots, \mathbf{y}_N) := \sum_{k=1}^{N} V_k(\mathbf{y}_k) + \sum_{i,j=1\, i<j}^{N} V_{ij}(\mathbf{y}_i - \mathbf{y}_j) , \tag{10.70}$$

where

$$\{V_k\}_{k=1,\ldots,N} \subset L^2(\mathbb{R}^n, dx) + L^\infty(\mathbb{R}^n, dx)\,,\ \{V_{ij}\}_{i<j\, i,j=1,\ldots,N} \subset L^2(\mathbb{R}^n, dx) + L^\infty(\mathbb{R}^n, dx)$$

are real functions. Then:

(a) *$-\Delta + V$ is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^{nN})$ and $\mathscr{S}(\mathbb{R}^{nN})$.*
(b) *The only self-adjoint extension $\overline{-\Delta + V}$ of the operators in (a) coincides with the (self-adjoint) operator $\overline{-\Delta} + V$ defined on $D(\overline{-\Delta})$.*
(c) *$\sigma(\overline{-\Delta + V})$ is lower bounded.*

Proof. We prove for $n = 3$, for the other cases are identical. Consider the potential $V_{12}(\mathbf{y}_1 - \mathbf{y}_2)$ and call Δ_1 the Laplacian corresponding to the coordinates of $\mathbf{y}_1$. Take $\varphi \in \mathscr{S}(\mathbb{R}^{3N})$. Fix $\mathbf{y}_2, \ldots \mathbf{y}_N \in \mathbb{R}^{3(N-1)}$ and define $\mathbb{R}^3 \ni \mathbf{y}_1 \mapsto \varphi'(\mathbf{y}_1) := \varphi(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_N)$. φ' belongs in $\mathscr{D}(\mathbb{R}^{3N})$ or $\mathscr{S}(\mathbb{R}^{3N})$, according to whether $\varphi \in \mathscr{D}(\mathbb{R}^{3N})$ or $\varphi \in \mathscr{S}(\mathbb{R}^{3N})$ respectively. Similarly, let $\mathbb{R}^3 \ni \mathbf{y}_1 \mapsto V'_{12}(\mathbf{y}_1) := V_{12}(\mathbf{y}_1 - \mathbf{y}_2)$. As in the previous proof, by decomposing $V_{12} = (V_{12})_2 + (V_{12})_\infty$ we arrive at the estimate, for any $a > 0$ and any $\mathbf{y}_2, \ldots, \mathbf{y}_N$:

$$||V'_{12}\varphi'||_{L^2(\mathbb{R}^3)} \leq a||(V_{12})_2||_{L^2(\mathbb{R}^3)}\, ||-\Delta_1\varphi'||_{L^2(\mathbb{R}^3)} + (b + ||(V_{12})_\infty||_{L^\infty(\mathbb{R}^3)})||\varphi'||_2$$

where $b > 0$ depends on a, *not* on $\mathbf{y}_2, \ldots \mathbf{y}_N \in \mathbb{R}^{3(N-1)}$. Norms are in the spaces over the first copy of $\mathbb{R}^3$ in $\mathbb{R}^{3N}$. It is important to note, due to the invariance of $(\mathbf{y}_1, \mathbf{y}_2) \mapsto V_{12}(\mathbf{y}_1 - \mathbf{y}_2)$ under translations, that the norms $||(V_{12})_k||_{L^k(\mathbb{R}^3)}$ do not depend on the variable $\mathbf{y}_2$. From Remarks 10.42 this inequality is the same as

$$||V'_{12}\varphi'||^2_{L^2(\mathbb{R}^3)} \leq a'\, ||-\Delta_1\varphi'||^2_{L^2(\mathbb{R}^3)} + b'||\varphi'||^2_{L^2(\mathbb{R}^3)}$$

for certain $a', b' > 0$ with a' arbitrarily small because of $a||V_{12}||_2$. Integrating the inequality in the variables $\mathbf{y}_2, \ldots \mathbf{y}_N \in \mathbb{R}^{3(N-1)}$ produces, for any $a' > 0$, a corresponding $b' > 0$ such that

$$||V_{12}\varphi||^2_{L^2(\mathbb{R}^{3N})} \leq a'\, ||-\Delta_1\varphi||^2_{L^2(\mathbb{R}^{3N})} + b'||\varphi||^2_{L^2(\mathbb{R}^{3N})}\,. \tag{10.71}$$

Transforming with Fourier-Plancherel on $\mathbb{R}^{3N}$, we now have

$$||-\Delta_1\varphi||^2_{L^2(\mathbb{R}^{3N})} = \int_{\mathbb{R}^{3N}} \left|\sum_{r=1}^{3} k_r^2\right|^2 |(\widehat{\mathscr{F}\varphi})(k_1, \ldots, k_{3N})|^2 dk_1 \cdots dk_{3N}$$
$$\leq \int_{\mathbb{R}^{3N}} \left|\sum_{r=1}^{3N} k_r^2\right|^2 |(\widehat{\mathscr{F}\varphi})(k_1, \ldots, k_{3N})|^2 dk_1 \cdots dk_{3N} = ||-\Delta\varphi||^2_{L^2(\mathbb{R}^{3N})}\,.$$

Substituting in (10.71) we conclude that if $\varphi \in \mathscr{D}(\mathbb{R}^{3N})$ or $\mathscr{S}(\mathbb{R}^{3N})$, then for any $a > 0$ there is a $b_{12} > 0$ satisfying

$$||V_{12}\varphi||^2_{L^2(\mathbb{R}^{3N})} \leq a\, ||-\Delta\varphi||^2_{L^2(\mathbb{R}^{3N})} + b_{12}||\varphi||^2_{L^2(\mathbb{R}^{3N})}\,.$$

The same result holds for the other potentials V_{ij}, V_k: the proof goes along the same lines, and is even simpler. If $\varphi \in \mathscr{D}(\mathbb{R}^{3N})$ or $\mathscr{S}(\mathbb{R}^{3N})$, for any $a > 0$ there are corresponding $b_i > 0$ and $b_{ij} > 0$ ($i, j = 1, \ldots, N$, $j > i$) such that:

$$||V_i\varphi||^2_{L^2(\mathbb{R}^{3N})} \leq a||-\Delta\varphi||^2_{L^2(\mathbb{R}^{3N})} + b_i||\varphi||^2_{L^2(\mathbb{R}^{3N})}\,, \tag{10.72}$$

$$||V_{ij}\varphi||^2_{L^2(\mathbb{R}^{3N})} \leq a||-\Delta\varphi||^2_{L^2(\mathbb{R}^{3N})} + b_{ij}||\varphi||^2_{L^2(\mathbb{R}^{3N})}\,. \tag{10.73}$$

On any Hermitian inner product space the Cauchy-Schwartz inequality implies $\left|\left|\sum_{r=1}^M \psi_r\right|\right|^2 \leq \left(\sum_{r=1}^M ||\psi_r||\right)^2$. There are $N + N(N-1)/2 = N(N+1)/2$ potentials V_k and V_{ij}, so Cauchy-Schwartz and (10.72)–(10.73) force

$$\left|\left|\left(\sum_{k=1}^N V_k + \sum_{i,j=1\, i<j}^N V_{ij}\right)\varphi\right|\right|^2_{L^2(\mathbb{R}^{3N})}$$

$$\leq \left(\frac{N(N+1)}{2}\right)^2 a||-\Delta\varphi||^2_{L^2(\mathbb{R}^{3N})} + \left(\frac{N(N+1)}{2}\right)^2 b||\varphi||^2_{L^2(\mathbb{R}^{3N})}$$

where b is the maximum of the b_k, b_{ij}. From Remarks 10.42 the result has an equivalent formulation. For every $a' > 0$ there exists a $b' > 0$ such that

$$||V\varphi|| \leq a'||-\Delta\varphi|| + b'||\varphi|| \quad \text{for any } \varphi \in \mathscr{S}(\mathbb{R}^{3N}).$$

From this point onwards the proof picks up from equation (10.69) in the proof of Theorem 10.47, replacing $\mathbb{R}^n$ with $\mathbb{R}^{3N}$. □

In conclusion we mention, without full proof, another important result of Kato. The demands on V to have $-\Delta + V$ essentially self-adjoint on $\mathscr{D}(\mathbb{R}^n)$ are different (and weaker than Theorem 10.47 if $n = 3$). Recall that $f : \mathbb{R}^n \to \mathbb{C}$ is called **locally square-integrable** if $f \cdot g$ is in $L^2(\mathbb{R}^n, dx)$ for every $g \in \mathscr{D}(\mathbb{R}^n)$.

Theorem 10.50. *The operator $-\Delta + V_\Delta + V_C$ defined on $L^2(\mathbb{R}^n, dx)$ is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^n)$, and its unique self-adjoint extension $\overline{-\Delta + V_\Delta + V_C}$ is bounded from below, provided the following conditions hold.*

(i) *$V_\Delta : \mathbb{R}^n \to \mathbb{R}$ is measurable and induces a $(-\Delta)$-bounded multiplicative operator with relative bound $a < 1$ (cf. Definition 10.41);*

(ii) *$V_C : \mathbb{R}^n \to \mathbb{R}$ is locally square-integrable with $V_C \geq C$ almost everywhere, for some $C \in \mathbb{R}$.*

Part (i) on V_Δ holds if

$$V_\Delta \in L^p(\mathbb{R}^n, dx) + L^\infty(\mathbb{R}^n, dx)\,,$$

with $p = 2$ when $n \leq 3$, $p > 2$ when $n = 4$ and $p = n/2$ when $n \geq 5$.

Sketch of proof. The final statement was proved with Theorem 10.47 if $n \leq 3$. The argument is the same for $n > 4$ by the remark ensuing Lemma 10.45. If (i) holds

$-\Delta + V_\Delta$ is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^n)$ and $\overline{-\Delta + V_\Delta}$ is lower bounded by the Kato-Rellich theorem. If (ii) holds as well, $-\Delta + V_\Delta + (V_C - C)$ is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^n)$ by [ReSi80, vol.II, theorem X.29], for $V_c - C \geq 0$. Therefore $-\Delta + V_\Delta + V_C = (-\Delta + V_\Delta + (V_C - C)) + CI$ is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^n)$. Since $-\Delta + V_\Delta$ and V_C are both bounded from below on that domain, so are $-\Delta + V_\Delta + V_C$ and $\overline{-\Delta + V_\Delta + V_C}$. □

Examples 10.51. **(1)** A case in $\mathbb{R}^3$ that is interesting to physics is one where the Laplacian perturbation V is the *attractive Coulomb potential*:

$$V(x) = \frac{eQ}{|x|},$$

with $e < 0$, $Q > 0$ constants, $|x| := \sqrt{x_1^2 + x_2^2 + x_3^2}$. The hypotheses of Kato's Theorem 10.49 (or 10.47) are valid for the operator:

$$H_0 := -\frac{\hbar^2}{2m}\Delta + V(x)$$

(the constants $m, \hbar > 0$ are irrelevant to the previous theorem, since we may multiply the operator by $2m/\hbar^2$ and then apply it, without losing in generality). So H_0 is essentially self-adjoint if defined on $\mathscr{D}(\mathbb{R}^3)$ or $\mathscr{S}(\mathbb{R}^3)$. The self-adjoint extension $\overline{H_0}$, if $Q = -e$, corresponds to the *Hamiltonian* operator of an electron in the electric field of a proton (neglecting spin effects and viewing the proton as a classical object). This gives the simplest quantum description of the Hamiltonian operator of the hydrogen atom. Here $-e$ is the common absolute value of the charge of electron and proton, m is the electronic mass, $\hbar > 0$ is Planck's constant divided by 2π. The spectrum of the unique self-adjoint extension of this operator determines, in physics, the admissible values of the energy of the system. Despite V is not bounded from below, it is important that the spectrum of the operator considered is always bounded, and therefore also the energy values that are physically admissible have a lower bound. In Chapters 11, 12 and 13 we will examine better the meaning of the operators here briefly described.

(2) A second case of physical interest, always in $\mathbb{R}^3$, is given by the *Yukawa potential*:

$$V(x) = \frac{-e^{-\mu|x|}}{|x|},$$

where $\mu > 0$ is another positive number. Here, too, the operator $H_0 = -\frac{\hbar^2}{2m}\Delta + V(x)$ is essentially self-adjoint if defined on $\mathscr{D}(\mathbb{R}^3)$ or on $\mathscr{S}(\mathbb{R}^3)$, as we know from Kato's Theorem 10.49 (or 10.47). The Yukawa potential describes, roughly, interactions between a *pion* and a source of *strong force*, the latter thought of, in this manner of speaking, as caused by a macroscopic source.

(3) The third physically-relevant case is the Hamiltonian of a system of N particles that interact under an external Coulomb potential and the Coulomb potentials of all pairs (not necessarily attractive). Call $\mathbf{x}_i \in \mathbb{R}^3$ the position vectors, $m_i > 0$ the masses and $e_i \in \mathbb{R} \setminus \{0\}$ the charges ($i = 1, \ldots, N$). The full operator is

$$H_0 := \sum_{i=1}^{N} -\frac{\hbar^2}{2m_i}\Delta_i + \sum_{i=1}^{N} \frac{Q_i e_i}{|\mathbf{x}_i|} + \sum_{i<j}^{N} \frac{e_i e_j}{|\mathbf{x}_i - \mathbf{x}_j|} ,$$

where Δ_i is the Laplacian in the three coordinates of $\mathbf{x}_i$. In order to apply Kato's theorem we must eliminate all factors $\frac{\hbar^2}{2m_i}$ multiplying the Δ_i. For this we can just change coordinates to $\mathbf{y}_i := \frac{\sqrt{2m_i}}{\hbar}\mathbf{x}_i$. Thus the first sum above gives the Laplacian on $\mathbb{R}^{3N}$ in the collective $3N$ components of the $\mathbf{y}_i$. It is not hard to see that the perturbation $V(\mathbf{y}_1, \ldots, \mathbf{y}_N)$ satisfies Kato's Theorem 10.49, so H_0 is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^{3N})$ and its unique self-adjoint extension is bounded from below.

(4) Theorem 10.50 allows to say the following. Adding any real function V', locally integrable and bounded from below, to the Hamiltonian operators H_0 seen in the previous examples gives an essentially self-adjoint operator on the corresponding $\mathscr{D}(\mathbb{R}^n)$. An important instance is the harmonic potential (non-isotropic, in general) $V'(x) = kx_1^2 + k_2x^2 + k_3x_3^2$ with $k_1, k_2, k_3 \geq 0$. ■

Exercises

10.1. Referring to Example 10.16(1), assume $\gamma > 0$. Prove the solution to the Klein–Gordon equation with source $e^{i\omega}\psi$ and dissipative term has the form:

$$u_t = e^{-\gamma t}\left[\cos\left(t\sqrt{A-\gamma I}\right)v + \sin\left(t\sqrt{A-\gamma I}\right)(A-\gamma I)^{-1/2}(v' + \gamma v)\right]$$

$$+ e^{-\gamma t}C_t\psi + e^{i\omega t}(A^2 - \omega^2 I + 2i\gamma\omega I)^{-1}\psi ,$$

where $||C_t|| \leq 1$.

Hint. Apply the definition of $\int_a^b L_\tau \psi_\tau d\tau$ given in (10.10). Then pass to the spectral measures of A and use the Fubini–Tonelli theorem, carefully verifying the assumptions.

10.2. Suppose $V : \mathbb{R}^3 \to \mathbb{R}$ makes the symmetric operator H_1, given by the differential operator $-\Delta_x + V(x)$, essentially self-adjoint on $\mathscr{S}(\mathbb{R}^3)$, where $\Delta_x := \sum_{k=1}^{3} \frac{\partial^2}{\partial x_i^2}$ is the Laplacian. Prove that the symmetric operator H on $L^2(\mathbb{R}^3 \times \mathbb{R}^3, dx \otimes dy)$ defined by the differential operator $-\Delta_x + V(x) - \Delta_y + V(y)$ is essentially self-adjoint on the span of finite products of a map in x in $\mathscr{S}(\mathbb{R}^3)$ and a map in y in $\mathscr{S}(\mathbb{R}^3)$. Then show

$$\sigma(\overline{H}) = \overline{\sigma(H_1) + \sigma(H_1)} .$$

10.3. If $A_k \in \mathfrak{B}(\mathsf{H}_k)$, $k = 1, \ldots, N$, prove

$$A_1 \otimes \cdots \otimes A_k \in \mathfrak{B}(\mathsf{H}_1 \otimes \cdots \otimes \mathsf{H}_N) .$$

Solution. Consider $N=2$, the general case being similar. If $\psi=\{f_i\}_{i\in I}$ and $\{g_j\}_{j\in J}$ are bases of H_1 and H_2, take the finite sum $\psi := \sum_{ij} c_{ij} f_i \otimes g_j$. Then $||(A_1\otimes I)\psi||^2 = \sum_j ||\sum_i c_{ij} A_1 f_i||^2 \leq \sum_j ||A_1||^2 \sum_i |c_{ij}|^2 = ||A_1||^2 ||\psi||^2$. A density argument allows to conclude $||A_1 \otimes I|| \leq A$, and therefore $||A_1\otimes A_2|| \leq ||A_1\otimes I||\,||I\otimes A_2|| \leq ||A_1||\,||A_2||$.

10.4. Under the assumptions in (2), prove that

$$||A_1\otimes\cdots\otimes A_k|| = ||A_1||\cdots||A_N||\,.$$

Solution. Take $n=2$, the generalisation being similar. If $A_1=0$ or $A_2=0$ the claim is obvious, so assume $||A_1||,||A_2||>0$. When solving (2) we found $||A_1\otimes A_2|| \leq ||A_1||\,||A_2||$, so it is enough to obtain the opposite inequality. By definition of $||A_1||$ and $||A_2||$, for any $\varepsilon>0$ there are $\psi_1^{(\varepsilon)}\in\mathsf{H}_1$ and $\psi_2^{(\varepsilon)}\in\mathsf{H}_2$, $||\psi_1^{(\varepsilon)}||,||\psi_2^{(\varepsilon)}||=1$, such that $|\,||A_i\psi_i^{(\varepsilon)}|| - ||A_i||\,| < \varepsilon$. In particular, $||\psi_i^{(\varepsilon)}|| \geq ||A_i|| - \varepsilon$. With these choices

$$||(A_1\otimes A_2)(\psi_1^{(\varepsilon)}\otimes\psi_2^{(\varepsilon)})|| \\ = ||A_1\psi_1^{(\varepsilon)}||\,||A_2\psi_2^{(\varepsilon)}|| \geq ||A_1||\,||A_2|| - \varepsilon(||A_1||+||A_2||)+\varepsilon^2\,.$$

Since $||\psi_1^{(\varepsilon)}\otimes\psi_2^{(\varepsilon)}||=1$, and from

$$||A_1\otimes A_2|| = \sup_{||\psi||=1} ||A_1\otimes A_2\psi|| \geq ||(A_1\otimes A_2)(\psi_1^{(\varepsilon)}\otimes\psi_2^{(\varepsilon)})||\,,$$

for any $\varepsilon>0$ we have $||A_1\otimes A_2|| \geq ||A_1||\,||A_2|| - \varepsilon(||A_1||+||A_2||)+\varepsilon^2$, where $-\varepsilon(||A_1||+||A_2||)+\varepsilon^2<0$. That value tends to 0 as $\varepsilon\to 0^+$. Eventually, $||A_1\otimes A_2|| \geq ||A_1||\,||A_2||$ as required.

10.5. If $A_k\in\mathfrak{B}(\mathsf{H}_k)$, $k=1,\ldots,N$ prove

$$(A_1\otimes\cdots\otimes A_k)^* = A_1^*\otimes\cdots\otimes A_N^*\,.$$

Hint. Check $A_1^*\otimes\cdots\otimes A_N^*$ satisfies the properties of the adjoint to a bounded operator (Proposition 3.36).

10.6. If $P_k\in\mathfrak{B}(\mathsf{H}_k)$, $k=1,\ldots,N$ are orthogonal projections, show $P_1\otimes\cdots\otimes P_k$ is an orthogonal projection.

This page intentionally left blank

11

Mathematical formulation of non-relativistic Quantum Mechanics

Every science would be redundant if the essence of things and their phenomenic appearance coincided.

Karl Marx

In this chapter we will enucleate the axioms of QM for the elementary system made by a non-relativistic particle, without spin, and discuss a series of important results related to the *canonical commutation relations*.

In section one we will recall the first four axioms given in Chapter 7, comment further on superselection rules and present a few new technical results. Then we will add an axiom relative to the formalisation of the quantum theory of the spin-zero particle. We will introduce, in particular, the *canonical commutation relations* (CCRs) and prove these cannot be satisfied by bounded operators. We will show how *Heisenberg's uncertainty principle* is actually a theorem in the formulation.

The last section is dedicated to the famous *theorem of Stone–von Neumann*, refined by Mackey, that characterises continuous unitary representations of CCRs. To prove the theorem we will introduce *Weyl *-algebras* and discuss their main properties. After proving the theorems of Stone–von Neumann and Mackey, we will use the formalism to extend Heisenberg's relations under rather weak hypotheses on the states involved, and then generalise them to mixtures. We will reformulate the results of Stone–von Neumann and Mackey in terms of the *Heisenberg group*. A short description of the so-called *Dirac correspondence principle* will close the section.

11.1 Round-up and remarks on axioms A1, A2, A3, A4 and superselection rules

In Chapter 7 we saw the general axioms of QM. Let us summarise part of that chapter in the light of the spectral theory developed subsequently.

A1. *Given a quantum system S described in an (inertial) frame system $\mathscr{I}$, experimentally testable propositions on S at any given time correspond bijectively to (a subset, in presence of superselection rules, of) the lattice $\mathfrak{P}(\mathsf{H}_S)$ of orthogonal projectors of a (complex) separable Hilbert space H_S, called the* **Hilbert space associated to** *S. Moreover (using the same letter for propositions and corresponding projectors):*

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_11, © Springer-Verlag Italia 2013

(1) *The compatibility of two propositions (from measuring processes attributing simultaneous truth values to both) corresponds to the commutation of the orthogonal projectors.*
(2) *The logical implication of two compatible propositions $P \Rightarrow Q$ corresponds to the projectors' relation $P \leq Q$.*
(3) *I (identity operator) and 0 (null operator) correspond to the tautology and the contradiction.*
(4) *The negation of P, $\dashv P$, corresponds to the orthogonal projector $\neg P = I - P$.*
(5) *Only when P, Q are compatible the propositions $P \mathscr{O} Q$ and $P \mathscr{E} Q$ have a physical meaning and correspond to orthogonal projectors $P \vee Q$ and $P \wedge Q$ (respectively projecting on the closure of the union and the intersection of the projection spaces of P, Q).*
(6) *If $\{Q_n\}_{n\in\mathbb{N}}$ is a countable collection of pairwise-compatible propositions, the propositions corresponding to $\vee_{n\in\mathbb{N}}Q_n$ and $\wedge_{n\in\mathbb{N}}Q_n$ have a physical meaning.*

In the sequel we shall assume, loosely speaking, that all the elements in $\mathfrak{P}(\mathsf{H}_S)$ describe elementary propositions on S. We shall also explain what happens in presence of superselection rules.

Remark 11.1. (1) The Hilbert space H_S actually depends on the frame system $\mathscr{I}$ as well, as explained in Remark 7.21(4); another system will give an isomorphic Hilbert space. We will return to this in Chapter 13.
(2) In Chapter 7.3.3 we saw that assuming elementary propositions on S are described by all of $\mathfrak{P}(\mathsf{H}_S)$ is a highly non-trivial matter. One could ask less, like having elementary propositions described by the sublattice of orthogonal projectors of a *von Neumann algebra* $\mathfrak{R}_S \subset \mathfrak{B}(\mathsf{H}_S)$. Self-adjoint elements in $\mathfrak{R}_S$ identify with bounded observables on S, since $\mathfrak{R}_S$ must contain the orthogonal projectors of every observable's spectral measure, and as any von Neumann algebra is closed in the strong topology; the latter is the reference topology to state the spectral theorem and define bounded measurable functions of self-adjoint operators. Hence the sublattice $\mathfrak{P}_{\mathfrak{R}_S}(\mathsf{H}_S)$ of orthogonal projectors in $\mathfrak{R}_S$ contains all spectral measures of all observables, and by the double-commutant Theorem 3.46 we have $\mathfrak{R}_S = \mathfrak{P}_{\mathfrak{R}_S}(\mathsf{H}_S)''$. If superselection rules are present (see Remark 7.36(2)) it becomes necessary to assume $\mathfrak{R}'_S \cap \mathfrak{R}_S$ be non-trivial, for that intersection contains projectors onto coherent sectors that, by nature, commute with every elementary proposition. In this case $\mathfrak{R}_S$ is not a factor. ■

A2. *A* **state** *ρ at time t on a quantum system S, with associated Hilbert space H_S, is a positive, trace-class operator on H_S with trace one. The probability that the proposition $P \in \mathfrak{P}(\mathsf{H}_S)$ on ρ is true equals $tr(\rho P)$.*

If we suppose elementary propositions are described by the whole space $\mathfrak{P}(\mathsf{H}_S)$, states ρ are convex combinations (also infinite, if we consider spectral decompositions of states) of extreme states in the convex set $\mathfrak{S}(\mathsf{H}_S)$ of states. Extreme states are called **pure** and have the form $\rho = \psi(\psi|\ \ )$, with $\psi \in \mathsf{H}_S$ such that $||\psi|| = 1$. The space of pure states is denoted by $\mathfrak{S}_p(\mathsf{H}_S)$ and is in one-to-one correspondence with the (projective) space of rays of H_S, i.e. the quotient of H_S minus 0 by the equivalence relation $\phi \sim \phi' \Leftrightarrow \phi = a\phi'$ for some $a \in \mathbb{C} \setminus \{0\}$. States that are *not* pure are called **mixed states** or **mixtures**, and the corresponding trace-class operators are often called **statistical operators** or **density matrices** in the literature. The convex decomposition of a mixed state in terms of pure states, arising for instance from the spectral theorem, is called **incoherent superposition** of pure states. There are typically more than one convex decompositions of a mixed state into pure ones.

An important notion in physics, also historically speaking, is the **transition (or probability) amplitude** $(\psi|\phi)$ of the pure state determined by the unit vector ϕ on the pure state determined by the unit ψ. The square modulus of the transition amplitude represents the probability that the system in state ϕ passes to state ψ after a measurement. Note that we may swap states, by symmetry of Hermitian inner products, without changing the transition probability.

A3. *If the quantum system S is in state $\rho \in \mathfrak{S}(\mathsf{H}_S)$ at time t and proposition $P \in \mathfrak{P}(\mathsf{H}_S)$ is validated by a measurement taken at the same t, the system's immediate post-measurement state is*

$$\rho_P := \frac{P\rho P}{tr(\rho P)} .$$

In particular if ρ is pure and given by $\psi \in \mathsf{H}_S$, $||\psi|| = 1$, the post-measurement state is still pure, and given by the vector

$$\psi_P = \frac{P\psi}{||P\psi||} .$$

We emphasise that this axiom refers to ideal *first-kind measurements*, or *non-destructive* or *indirect*, as they are known; a lab's practice adopts several types of testing, that in general do not obey the axiom.

Remark 11.2. (1) In presence of superselection rules, when the Hilbert space H_S splits in *countably many, at most,* coherent sectors $\mathsf{H}_S = \oplus_k \mathsf{H}_{Sk}$, one assumes states are statistical operators of the form (in the strong topology) $\rho = \sum_k p_k \psi_k(\psi_k|\cdot)$, with $\psi_k \in \mathsf{H}_{Sk}$, $p_k \in [0,1]$, $\sum_k p_k = 1$. Hence pure states, i.e. extreme points of the convex set of states, are still given by unit vectors ψ_k; now, however, these vectors must belong to a coherent sector H_{Sk} (depending on the state), and can no longer be considered coherent superpositions. For example $\psi = a\psi_k + b\psi_{k'}$ does not determine a state if $\psi_k \in \mathsf{H}_{Sk}$, $\psi_{k'} \in \mathsf{H}_{Sk'}$ and $k \neq k'$. We shall denote the space of pure states under superselection rules by $\mathfrak{S}_p(\mathsf{H})_{adm}$. The elementary propositions of the theory are, as we said, described by projectors in $\mathfrak{P}_{\mathfrak{R}_S}(\mathsf{H}_S)$, which commute with every orthogonal projector P_k associated to the coherent decomposition ($\mathsf{H}_{Sk} = P_k(\mathsf{H}_{Sk})$) since $P_k \in \mathfrak{R}_S \cap \mathfrak{R}'_S$.

There is an alternative way to operate in presence of superselection rules. One can impose no restriction on the choice of vectors representing states by the following observation: an apparently pure state described by the coherent superposition $\psi = \sum_k \psi_k$, $\psi_k \in \mathsf{H}_k$ is a mixed state $\rho = \sum_k \psi_k(\psi_k|\cdot)$ in every respect (concerning

measurements of elementary propositions and observables, post-measurement states, or time evolution, of which more in the next chapter). In this case we should take care that it is no longer true that vectors represent necessarily pure states; moreover $\sum_k \psi_k$ and $\sum_k e^{i\alpha_k}\psi_k$ represent the same (mixed) state even if the $\alpha_k \in \mathbb{R}$ are non-zero. In contrast to the general case (see remark (4) after axiom **A3**, Chapter 7), the summands of the incoherent splitting $\sum_k \psi_k(\psi_k|\cdot)$ of ψ now are uniquely fixed. This allows to distinguish between "classical probabilities", the numbers $p_k = ||\psi_k||^2$, and "quantum probabilities", related to the probability amplitudes in each H_{Sk} (cf. [Giu00]).
(2) A more difficult question is to decide, if $\mathfrak{R}'_S \cap \mathfrak{R}_S$ is not trivial, whether superselection rules are present. The point is to understand if it is possible to determine a set, *at most countable*, of orthogonal projectors $P_k \in \mathfrak{R}'_S \cap \mathfrak{R}_S$ such that $P_k P_h = 0$ when $h \neq k$, $\sum_k P_k = I$ (in strong topology if over an infinite set) and such, for any k, that there is no orthogonal projector $Q \in \mathfrak{R}'_S \cap \mathfrak{R}_S$ satisfying $Q \leq P_k$ with $Q \neq 0$ and $Q \neq P_k$. If these conditions hold, coherent sectors in the Hilbert space can be identified with the projection spaces of the P_k, at least mathematically. (If the last condition failed for P_h, for instance, $\mathsf{H}_h = P_h(\mathsf{H})$ would not automatically be coherent, because we could still decompose it under Q, thus preventing superpositions in H_k itself.) If the centre of $\mathfrak{R}_S$ contained a bounded observable A with *continuous spectrum* (see axiom **A4**), then the desired decomposition $\sum_k P_k = I$ would not be achievable. This is because of the spectral measure of A, whose elements belong to the centre of $\mathfrak{R}_S$ (and because the spectral measure of any open non-empty set intersecting the continuous spectrum is non-zero). Still, the Hilbert space could be decomposed as *direct integral*, and one could talk about *continuous superselection rules* [Giu00], which have a different nature and will not be of our concern in this book. We will only say the structure of the space of states is much more intricate in that case, and the *algebriaic formulation* of Chapter 14 is required to explain it thoroughly.

Instead, in case the decomposition $\sum_k P_k = I$ exists with all requirements, one can try to interpret $\{P_k\}_{k\in K}$ as a collection generating the joint spectral measure of a family of observables with discrete spectrum, representing "charges" that have to be defined simultaneously in every pure state: the H_k are common eigenspaces for all charges. It has been conjectured more than once that quantities with classical behaviour are actually quantum observables of this kind [BGJKS00], i.e. superselection charges. If so the superselection rule is dynamical and arises from the interaction between the physical system and the ambient producing the phenomenon of *decoherence* [BGJKS00]. ■

A4. *Every observable A of the quantum system S is described by a projector-valued measure $P^{(A)}$ over $\mathbb{R}$, on the system's Hilbert space H_S, so that the projector $P^{(A)}(E)$ corresponds to the proposition "The outcome of measuring A falls in E", for any Borel set E in $\mathbb{R}$*

The spectral theorem for unbounded self-adjoint operators, proved in its maximal generality in Chapter 9 (Theorem 9.10), allows to associate to any observable a *unique* self-adjoint operator on the Hilbert space of the physical system. With this, if H_S is the Hilbert space of some system, the spectrum $\sigma(A) \subset \mathbb{R}$ of an observable

A, i.e. of a self-adjoint operator $A : D(A) \to \mathsf{H}_S$, contains all possible outcomes of a measurement of the observable A. Mathematically, $\sigma(A)$ coincides with the support of the PVM $P^{(A)}$ associated to the observable.

Remarks 11.3. With superselection rules, Remark 7.47(1) easily implies that every self-adjoint operator A representing an observable satisfies $AP_k = P_k A$ for any k, where P_k is the orthogonal projector on the kth coherent sector. (See also Remark 11.1(2).) ■

Important in physics is the notion of *compatible observables*:

Definition 11.4. *Let S be a quantum system described on the Hilbert space H_S. Two observables A, B of S are* **compatible** *if the spectral measures $P^{(A)}$, $P^{(B)}$ of the corresponding self-adjoint operators commute, i.e.*

$$P^{(A)}(E)P^{(B)}(E) = P^{(B)}(E)P^{(A)}(E)\,, \quad \textit{for any Borel set } E \subset \mathbb{R}.$$

Two observables that are not compatible are called **incompatible**.

In physics, compatibility means the observables can be measured at the same time (in agreement with axiom **A1** and with the meaning of the associated spectral measures). If we have a finite set of compatible observables $\mathbf{A} = \{A_1, A_2, \ldots, A_n\}$, a *joint spectral measure* can be constructed using the spectral measures of the self-adjoint operators representing the observables, by Theorem 9.16. Retaining those notations, if $f : \mathbb{R}^n \to \mathbb{C}$ is measurable, the self-adjoint operator

$$f(A_1, \ldots, A_n) := \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) dP^{(\mathbf{A})}(x_1, \ldots, x_n) \tag{11.1}$$

– with domain given by vectors $\psi \in \mathsf{H}$ for which $f \in L^2(\mathbb{R}^n; \mu_\psi)$, where $\mu_\psi(E) = (\psi|P^{(\mathbf{A})}(E)\psi)$, $E \in \mathscr{B}(\mathbb{R}^n)$, as usual – has the customary meaning of *observable* that is *function* of the observables $A_1, \ldots, A_n$.

Remark 11.5. **(1)** A necessary condition to have compatible observables is that the corresponding operators commute, paying attention to domains as prescribed by Theorem 9.35. For unbounded self-adjoint operators this condition is not sufficient, despite what certain physics books might say: Nelson [Nel59] proved that there are pairs of operators that commute on a dense subspace, invariant for both and on which both are essentially self-adjoint, yet the spectral measures of the self-adjoint extensions do not commute. A useful necessary *and sufficient* condition for A and B to be compatible is Theorem 9.35(c):

$$e^{itA}e^{isB} = e^{isB}e^{itA} \quad \text{for any } s, t \in \mathbb{R}\,.$$

An second necessary *and sufficient* condition is (d) in the same theorem.
(2) Take a quantum system S described on the Hilbert space H_S and a family $\mathbf{A} = \{A_1, \ldots, A_n\}$ of observables with pairwise-commuting spectral measures. Suppose $\mathbf{A}' = \mathbf{A}''$, i.e. every operator in $\mathfrak{B}(\mathsf{H}_S)$ that commutes with the spectral measures of

the A_k has the form $f(A_1,\ldots,A_n)$ for $f : supp(P^{(\mathbf{A})}) \to \mathbb{C}$ bounded measurable. In that case $\mathbf{A} = \{A_1,\ldots,A_n\}$ is called a **maximal observable**. This can be proved to be the same as asking the Hilbert space be isomorphic to an L^2 space on the joint spectrum of the A_k [BeCa81], or to the existence of a cyclic vector for the joint spectral measure. Dirac speculated that the set of observables of a quantum system always admits a maximal observable. Jauch gave the general version of Dirac's postulate in terms of von Neumann algebras, positing the existence of a finite set of pairwise-commuting observables $\mathbf{A}$ such that $\mathbf{A}' = \mathbf{A}''$. Dirac's original conjecture referred only to observables with point spectrum.

It can also be proved that the existence of a maximal observable amounts to the demand that the commutant $\mathfrak{R}'_S$ be Abelian [Giu00], where $\mathfrak{R}_S$ is the usual von Neumann algebra of bounded observables in the system. ■

Proposition 7.52 enables to associate to any pair "observable–state", A,ρ, a probability measure on $\mathbb{R}$, $\mu_\rho^{(A)} : E \mapsto tr(\rho P^{(A)}(E))$ with $E \in \mathscr{B}(\mathbb{R})$ (coinciding with $\mu_\psi^{(A)} = (\psi|P^{(A)}(E)\psi)$, cf. Theorem 8.50(c), if ρ is pure and determined by the unit $\psi \in \mathsf{H}$). By construction $supp(\mu_\rho^{(A)}) \subset \sigma(A)$. Since by definition $\mu_\rho^{(A)}(E)$ coincides with the probability that the measurement of A belongs to E, in the state ρ, it makes sense to define the *mean value* and the *standard deviation* of A in the state ρ.

Definition 11.6. *Let A be an observable of the physical system S described on the Hilbert space H_S, let $\rho \in \mathfrak{S}(\mathsf{H}_S)$ be a state of S and $\mu_\rho^{(A)}$ the probability measure associated to ρ and A as above. The* **mean value** *of A^n, $n = 1,2,\ldots$, equals*

$$\langle A^n\rangle_\rho := \int_{\mathbb{R}} \lambda^n \, d\mu_\rho^{(A)}(\lambda)\,, \quad \text{when } \mathbb{R} \ni \lambda \mapsto \lambda^n \text{ is in } L^1(\mathbb{R},\mu_\rho^{(A)}). \tag{11.2}$$

The **standard deviation in state** *ρ of A equals*

$$\Delta A_\rho := \sqrt{\int_{\mathbb{R}} \left(\lambda - \langle A\rangle_\rho\right)^2 d\mu_\rho^{(A)}(\lambda)}\,, \quad \text{when } \mathbb{R} \ni \lambda \mapsto \lambda^2 \text{ is in } L^1(\mathbb{R},\mu_\rho^{(A)})\,. \tag{11.3}$$

If the requirement in (11.2) (or (11.3)) does not hold, the mean value of A^n (resp. standard deviation) does not exists for ρ.

Remarks 11.7. If the map λ^n belongs to $L^1(\mathbb{R},\mu_\rho^{(A)})$ then also λ^k does, for any $k = 1,2,\ldots,n-1$, because $\mu_\rho^{(A)}$ is finite. Therefore if $\langle A^n\rangle_\rho$ exists, so does $\langle A^k\rangle_\rho$ (and ΔA_ρ if $n \geq 2$) for any $k = 1,2,\ldots,n-1$. ■

The properties of Examples 7.53 generalise as follows.

Proposition 11.8. *Let A be an observable for a system described on the separable Hilbert space H_S, and take $\rho_\psi \in \mathfrak{S}_p(\mathsf{H}_S)$ associated to $\psi \in \mathsf{H}_S$ with $||\psi|| = 1$ and $\rho \in \mathfrak{S}(\mathsf{H}_S)$.*

(a) (i) *$\langle A\rangle_{\rho_\psi}$ exists $\Leftrightarrow$ $\psi \in D(|A|^{1/2})$, and ΔA_{ρ_ψ} exists $\Leftrightarrow$ $\psi \in D(A)$;*
(ii) *if $\psi \in D(A)$ then $\langle A\rangle_{\rho_\psi}$ exists, and:*

$$\langle A\rangle_{\rho_\psi} = (\psi|A\psi)\,; \tag{11.4}$$

(iii) *if $\psi \in D(A^2)$ then $\langle A\rangle_{\rho_\psi}$ and ΔA_{ρ_ψ} exist, equation (11.4) holds, and:*

$$\Delta A^2_{\rho_\psi} = \left(\psi\left|(A - \langle A\rangle_\psi I)^2\,\psi\right.\right) = (\psi|A^2\psi) - (\psi|A\psi)^2\,. \tag{11.5}$$

(b) (i) *$\langle A\rangle_\rho$ exists $\Leftrightarrow$ $Ran(\rho^{1/2}) \subset D(|A|^{1/2})$ and $|A|^{1/2}\rho^{1/2} \in \mathfrak{B}_2(\mathsf{H})$;*
(ii) *ΔA_ρ exists (equivalently $\langle A^2\rangle_\rho$ exists) iff $Ran(\rho^{1/2}) \subset D(A)$ and $A\rho^{1/2} \in \mathfrak{B}_2(\mathsf{H})$;*
(iii) *if $\langle A^2\rangle_\rho$ exists, then $A\rho \in \mathfrak{B}_1(\mathsf{H}_S)$ and:*

$$\langle A\rangle_\rho = tr(A\rho)\,; \tag{11.6}$$

(iv) *if $\langle A^4\rangle_\rho$ exists, then $A\rho \in \mathfrak{B}_1(\mathsf{H}_S)$, equation (11.6) holds, $(A - \langle A\rangle_\rho I)^2\rho \in \mathfrak{B}_1(\mathsf{H}_S)$ and:*

$$\Delta A^2_\rho = tr\left((A - \langle A\rangle_\rho I)^2\rho\right) = tr(A^2\rho) - tr(A\rho)^2\,. \tag{11.7}$$

Proof. (a) We have $tr(\rho_\psi P^{(A)}(E)) = (\psi|P^{(A)}(E)\psi) = \mu^{(A)}_\psi(E)$. Therefore asking $\mathbb{R} \ni \lambda \mapsto \lambda$ and $\mathbb{R} \ni \lambda \mapsto \lambda^2$ in $L^1(\mathbb{R}, \mu^{(A)}_{\rho_\psi})$ is respectively equivalent to $\psi \in D(|A|^{1/2})$ and $\psi \in D(A)$, by Definition 9.11 and Theorem 9.4(f). By definition, and using Theorem 9.4(e-f) for the standard deviation:

$$\langle A\rangle_{\rho_\psi} = \int_{\mathbb{R}} \lambda d\mu^{(A)}_\psi(\lambda), \tag{11.8}$$

$$\Delta A^2_{\rho_\psi} = \left(A\psi - \left(\int_{\mathbb{R}} \lambda d\mu^{(A)}_\psi(\lambda)\right)\psi \middle| A\psi - \left(\int_{\mathbb{R}} \lambda d\mu^{(A)}_\psi(\lambda)\right)\psi\right). \tag{11.9}$$

Using Theorem 9.4(g) these imply (11.4) and (11.5) if $\psi \in D(A)(\subset D(|A|^{1/2}))$ and $\psi \in D(A^2)(\subset D(A))$, respectively.
(b) Let $\{\psi_n\}_{n\in\mathbb{N}}$ be a basis of H_S (separable). Then $\mu^{(A)}_\rho(E) = tr(\rho P^{(A)}(E)) = tr(\rho^{1/2}P^{(A)}(E)\rho^{1/2}) = \sum_{n=0}^{+\infty}(\rho^{1/2}\psi_n|P^{(A)}(E)\rho^{1/2}\psi_n) = \sum_{n=0}^{+\infty}\mu^{(A)}_{\rho^{1/2}\psi}(E)$, for any Borel set $E \in \mathscr{B}(\mathbb{R})$, where we used $\rho^{1/2} \in \mathfrak{B}_2(\mathsf{H}_S)$ (as $\mathfrak{B}_1(\mathsf{H}) \ni \rho \geq 0$) and Proposition 4.36(c). If $f : \mathbb{R} \to \mathbb{C}$ is measurable, then,

$$\int_{\mathbb{R}} |f(\lambda)|d\mu^{(A)}_\rho(\lambda) = \sum_{n=0}^{+\infty}\int_{\mathbb{R}} |f(\lambda)|d\mu^{(A)}_{\rho^{1/2}\psi_n}(\lambda) \leq +\infty\,. \tag{11.10}$$

Moreover, if the right-hand side in (11.10) (so also the left side) is finite, then:

$$\int_{\mathbb{R}} f(\lambda)d\mu^{(A)}_\rho(\lambda) = \sum_{n=0}^{+\infty}\int_{\mathbb{R}} f(\lambda)d\mu^{(A)}_{\rho^{1/2}\psi_n}(\lambda) \in \mathbb{C}\,. \tag{11.11}$$

In fact, $\mu_\rho^{(A)}(E) = \sum_{n=0}^{+\infty} \mu^{(A)}_{\rho^{1/2}\psi}(E)$ implies (11.10) is trivially true if $|f| = s$ is a simple non-negative map. For any Borel-measurable $g \geq 0$ there is a simple sequence $0 \leq s_0 \leq s_1 \leq \cdots \leq s_n \to g$ (Proposition 7.49). By monotone convergence on the single integrals and on the counting measure of $\mathbb{N}$ we have (11.10), with $|f|$ replaced by an arbitrary $g \geq 0$. If f is real-valued, and in (11.10) we have $< +\infty$, decomposing f in its positive and negative parts $f = f_+ - f_-$, $0 \leq f_+, f_- \leq |f|$, gives (11.11) by linearity. If f is complex-valued the argument is similar, we just work with real and imaginary parts separately. $\langle f(A)\rangle_\rho$ exists precisely when the left-hand side in (11.10) is finite. In turn, this is the same as saying every summand on the right is finite and the sum is finite. The generic term is finite if and only if $\rho^{1/2}\psi_n \in D(|f(A)|^{1/2})$ by definition of $D(g(A))$ (Definition 9.11). Since ψ_n is an arbitrary unit vector in H, $Ran(\rho^{1/2}) \subset D(|f(A)|^{1/2})$. Every integral on the right in (11.10) can be written (see Theorem 9.4(f)) $|||f(A)|^{1/2}\rho^{1/2}\psi_n||^2$, where $|f(A)|^{1/2}\rho^{1/2} \in \mathfrak{B}(\mathsf{H}_S)$ by Proposition 5.6. By Definition 4.22 we conclude that the left-hand side (11.10) is finite iff $|f(A)|^{1/2}\rho^{1/2} \in \mathfrak{B}_2(\mathsf{H}_S)$. Choose $f(\lambda) = \lambda$, so (i) in (b) holds, then choose $f(\lambda) = \lambda^2$ to obtain (ii) in (b), because λ^2 integrable in $\mu_\rho^{(A)}$ implies λ integrable, plus $D(A) = D(|A|)$. To prove (iii) assume $\langle A^2\rangle_\rho$ exists (so also $\langle A\rangle_\rho$ exists) and notice $Ran(\rho) \subset D(A)$ from $Ran(\rho^{1/2}) \subset D(A)$, since $\rho = \rho^{1/2}\rho^{1/2}$ implies $Ran(\rho^{1/2}) \supset ran(\rho)$. Applying (11.11) to $f(\lambda) = \lambda$ and recalling Theorem 9.4(g) we find:

$$\langle A\rangle_\rho = \sum_{n\in\mathbb{N}} (\rho^{1/2}\psi_n|A\rho^{1/2}\psi_n) = \sum_{n\in\mathbb{N}} (\psi_n|\rho^{1/2}A\rho^{1/2}\psi_n) = tr(\rho^{1/2}A\rho^{1/2}),$$

where we used $\rho^{1/2}, A\rho^{1/2} \in \mathfrak{B}_2(\mathsf{H}_2)$, so their product (in any order) is of trace class. Since the trace is invariant under cyclic permutations (Proposition 4.36(c)) we have $\langle A\rangle_\rho = tr(A\rho^{1/2}\rho^{1/2}) = tr(A\rho)$, concluding (iii). The proof of (iv) is similar: replace A with A^2 and observe that if $\langle A^4\rangle_\rho$ exists, so does $\langle A^2\rangle_\rho$, and (iii) holds. The second identity in (11.7) now follows from the first with obvious algebraic manipulations. □

Remarks 11.9. The right-hand sides of (11.6) and (11.7) ((11.4) and (11.5) for pure states) are *not* the definitions of mean value and standard deviation; these are given, in general, by (11.2) and (11.3), independently from Proposition 11.8. ■

11.2 Axiom A5: non-relativistic elementary systems

To go further into the mathematical formulation of QM we must establish axioms about special elementary systems. These correspond to the *particles* of the non-relativistic theory. In other terms, the group of transformations under which the theory is invariant is the Galilean group, not the Poincaré group. We will return to this point later. For physics this description is adequate until speeds do not reach the order of the speed of light (about 300.000 km/s). However certain mathematical concepts, like the *Weyl *-algebra* introduced in the forthcoming non-relativistic description, have broader validity: they are employed in a relativistic regime as well, in formulations of *quantum field theory* that we will not discuss.

Complex systems are built by composing elementary ones via the Hilbert tensor product, as we will see subsequently when studying compound systems.

The simplest elementary system in non-relativistic QM consists in a quantum particle of mass $m > 0$ and spin 0. The next axiom holds in this system.

A5. *Consider an inertial frame system $\mathscr{I}$ equipped with orthonormal Cartesian coordinates x_1, x_2, x_3 on the rest space of the frame. A non-relativistic particle of mass $m > 0$ and spin 0 is described as follows.*

(a) *The system's Hilbert space is $\mathsf{H} = L^2(\mathbb{R}^3, dx)$, where $\mathbb{R}^3$ is identified with the rest space of $\mathscr{I}$ via the x_1, x_2, x_3, and dx is the ordinary Lebesgue measure on $\mathbb{R}^3$.*
(b) *The three observables associated to x_1, x_2, x_3 are self-adjoint operators, called* **position operators***:*

$$(X_i\psi)(x_1,x_2,x_3) = x_i\psi(x_1,x_2,x_3)\,, \tag{11.12}$$

$i = 1,2,3$, *of domains:*

$$D(X_i) := \left\{\psi \in L^2(\mathbb{R}^3, dx) \;\middle|\; \int_{\mathbb{R}^3} |x_i\psi(x_1,x_2,x_3)|^2\, dx < +\infty \right\}.$$

(c) *The three observables associated to the momentum components in $\mathscr{I}$, p_1, p_2, p_3, are self-adjoint operators, called* **momentum operators***:*

$$P_k = \overline{-i\hbar\frac{\partial}{\partial x_k}}\,, \tag{11.13}$$

$k = 1,2,3$, *where the operator on the right is the closure of the essentially self-adjoint differential operator:*

$$-i\hbar\frac{\partial}{\partial x_k} : \mathscr{S}(\mathbb{R}^3) \to L^2(\mathbb{R}^3, dx)$$

and $\mathscr{S}(\mathbb{R}^3)$ the Schwartz space on $\mathbb{R}^3$ (see Chapter 3.6).

In the physics' literature vectors (normalised to 1) of $L^2(\mathbb{R}^3, dx)$ associated to a particle are called its **wavefunctions**. Wavefunctions determine (not uniquely, owing to arbitrary numerical factors) the particle's pure states.

Remark 11.10. (1) The discussion of Chapter 5.3 explains that the same notions can be given if we define the operators X_i substituting $\psi \in L^2(\mathbb{R}^3, dx)$ with $\psi \in \mathscr{D}(\mathbb{R}^3)$, or taking $\psi \in \mathscr{S}(\mathbb{R}^3)$ in the domains. In either case one has to take the unique self-adjoint extension of the operator on $\mathscr{D}(\mathbb{R}^3)$ or $\mathscr{S}(\mathbb{R}^3)$.
(2) P_i may be defined equivalently by (see Definition 5.27, Proposition 5.29 and the ensuing discussion):

$$(P_i f)(\mathbf{x}) = -i\hbar\text{w-}\frac{\partial}{\partial x_i} f(\mathbf{x})\,,$$

$$D(P_i) := \left\{ f \in L^2(\mathbb{R}^3, dx) \;\middle|\; \text{w-}\tfrac{\partial}{\partial x_i} f \in L^2(\mathbb{R}^3, dx) \text{ exists} \right\}.$$

As usual w-$\frac{\partial}{\partial x_i}$ denoted the weak derivative. The study of Chapter 5.3 also shows that P_i (see Proposition 5.29) can be defined, equivalently, substituting the Schwartz space with $\mathscr{D}(\mathbb{R}^3)$ and taking the unique self-adjoint extension of the operator obtained, which is still essentially self-adjoint.
(3) Let K_i denote the ith position operator on the codomain of the Fourier-Plancherel transform $\widehat{\mathscr{F}} : L^2(\mathbb{R}^3, dx) \to L^2(\mathbb{R}^3, dk)$, see Chapter 3.6. Then Proposition 5.31 gives

$$P_i = \hbar \widehat{\mathscr{F}}^{-1} K_i \widehat{\mathscr{F}}\,,$$

an alternative definition of momentum.
(4) From Chapter 9.1.5 we know

$$\sigma(X_i) = \sigma_c(X_i) = \mathbb{R}\,, \quad \sigma(P_i) = \sigma_c(P_i) = \mathbb{R} \quad i = 1,2,3. \tag{11.14}$$

■

11.2.1 The canonical commutation relations (CCRs)

The definition of position and momentum is such that there exist *invariant* spaces $\mathsf{H}_0 \subset L^2(\mathbb{R}^3, dx)$ for all six observables, despite the latter's domains are different: $X_i(\mathsf{H}_0) \subset \mathsf{H}_0$ and $P_i(\mathsf{H}_0) \subset \mathsf{H}_0$, $i = 1,2,3$. For instance take the Schwartz space $\mathsf{H}_0 = \mathscr{S}(\mathbb{R}^3)$. Checking this is immediate by definition of Schwartz space. On $\mathscr{S}(\mathbb{R}^3)$ a direct computation that uses (11.12) and (11.13) yields Heisenberg's **canonical commutation relations** (CCRs):

$$[X_i, P_j] = i\hbar\delta_{ij} I\,,$$

where $\delta_{ij} = 0$ for $i \neq j$ and $\delta_{ij} = 1$ for $i = j$. More precisely:

Lemma 11.11. *The operators position X_i and momentum P_j, $i, j = 1,2,3$, defined in* **A5***, fulfill* **Heisenberg's commutation relations***:*

$$[X_i, P_j]\psi = i\hbar\delta_{ij}\psi \quad \text{for every } \psi \in D(X_iP_j) \cap D(P_jX_i),\ i, j = 1,2,3\,. \tag{11.15}$$

Equations (11.15) are valid when replacing X_i with $X'_i := X_i + a_i I$ and P_j with $P'_j := P_j + b_j I$, for any constants $a_i, b_j \in \mathbb{R}$.

Proof. A straighforward computation shows $D(X'_iP'_j) \cap D(P'_jX'_i) = D(X_iP_j) \cap D(P_jX_i)$. On $\varphi \in \mathscr{D}(\mathbb{R}^3)$, the operator P'_j acts as $-i\hbar\partial/\partial x_j + b_j I$ by construction; since X'_i multiplies by the shifted coordinate $x_i + a_i$, we obtain $P'_jX'_i\varphi = -i\hbar\delta_{ij}\varphi + X'_iP'_j\varphi$. Therefore

$$\left(P'_jX'_i\varphi - X'_iP'_j\varphi + i\hbar\delta_{ij}\varphi \middle| \psi\right) = 0\,, \quad \varphi \in \mathscr{D}(\mathbb{R}^3),\ \psi \in L^2(\mathbb{R}^3, dx).$$

In turn, if $\psi \in D(X'_iP'_j) \cap D(P'_jX'_i) = D(X_iP_j) \cap D(P_jX_i)$, since P_j and X_i are self-adjoint, the identity reads

$$\left(\varphi \middle| X'_iP'_j\psi - P'_jX'_i\psi - i\hbar\delta_{ij}\psi\right) = 0\,.$$

As $\mathscr{D}(\mathbb{R}^3)$ is dense in $L^2(\mathbb{R}^3, dx)$, (11.15) holds with X'_i, P'_j replacing X_i, P_j. Consequently (11.15) holds by taking $a_j = b_j = 0$, $j = 1,2,3$. □

Pairs of observables solving Heisenberg's relations (11.15), on some invariant domain such as $\mathscr{S}(\mathbb{R}^3)$, are often called **conjugate observables**. The relations are the simplest manifestation of general CCRs for *Bosonic* quantum fields, where position and momentum are defined to befit the theory. From the physical viewpoint it has been noticed over and over during the history of QM, and its evolutions, that the canonical commutation relations are much more important than the operators X_i and P_i themselves.

As the definitions make obvious, position and momentum are unbounded operators, and are not defined on the entire Hilbert space. Technically this is a thorn in the side, for it forces to bring into the picture spectral theory of unbounded operators, which is more involved than the bounded theory. This begs the question whether a substitute definition of X_i, P_i might exists, preserving Heisenberg's relations and making the operators bounded. The answer is no, and the reason is dictated by Heisenberg's CCRs.

Proposition 11.12. *There are no self-adjoint operators X and P such that, on a common invariant subspace, $[X,P] = i\hbar I$ and at the same time X, P are bounded.*

Proof. Suppose $[X,P] = i\hbar I$ on a common invariant space D where they are bounded. Restrict to D (or its closure $\overline{D}$, by extending X,P to self-adjoint operators defined on $\overline{D}$), and consider it as the Hilbert space. The restrictions will now be self-adjoint and bounded. From $[X,P] = i\hbar I$:

$$PX^n - X^nP = -inX^{n-1} .$$

If n is odd, using Proposition 3.38(a) repeatedly (as $X^p = (X^p)^*$ for any natural p) and the norm's properties:

$$n||X||^{n-1} = n||X^{n-1}|| \leq 2||P||||X^n|| \leq 2||P||||X||||X^{n-1}|| = 2||P||||X||||X||^{n-1} .$$

As $||X|| \neq 0$ (because of (11.15)), we obtain the absurd: $n \leq 2||P||||X|| < +\infty$ for any $n = 1,3,5,\ldots$. □

11.2.2 Heisenberg's uncertainty principle as a theorem

A comforting immediate consequence of the CCRs and the formalism is to turn *Heisenberg's Uncertainty Principle* for the variables position and momentum (cf. Chapter 6.4) into a theorem. We shall prove the principle in its classical form on *pure states*, only to reformulate it later under weaker assumptions on vectors and then *extend it to mixed states*.

Theorem 11.13 ("Heisenberg's Uncertainty Principle"). *Let ψ be a unit vector, describing a pure state of a classical particle of spin* 0*, such that:*

$$\psi \in D(X_iP_i) \cap D(P_iX_i) \cap D(X_i^2) \cap D(P_i^2)$$

(in particular $\psi \in \mathscr{S}(\mathbb{R}^3)$). Then

$$(\Delta X_i)_\psi (\Delta P_i)_\psi \geq \frac{\hbar}{2} \quad i=1,2,3 \tag{11.16}$$

(where we wrote ψ instead of ρ_ψ for simplicity.

Proof. The hypotheses imply, in particular, $\psi \in D(X_i^2) \cap D(P_i^2)$, so standard deviations are defined and can be found using formula (11.5). By definition (11.3) we see $(\Delta X_i)_\psi = (\Delta X_i')_\psi$ and $(\Delta P_i)_\psi = (\Delta P_i')_\psi$ for $X_i' := X_i + a_i I$, $P_i' := P_i + b_i I$ with a_i, b_i are real constants. Hence replacing X_i, P_i by X_i', P_i' produces an equivalent formula to (11.16). Let us choose $a_i = -\langle X_i \rangle_\psi$, $b_i = -\langle P_i \rangle_\psi$ and prove (11.16) for the operators X_i'. P_i'. From (11.5) the choices force $(\Delta X_i')_\psi = ||X_i'\psi||$ and $(\Delta P_i')_\psi = ||P_i'\psi||$. So we need to prove

$$||X_i'\psi||\,||P_i'\psi|| \geq \hbar/2\,. \tag{11.17}$$

As X_i', P_i' satisfy (11.15), Schwarz's inequality, the operators' self-adjointness and the properties of the inner product give

$$||X_i'\psi||\,||P_i'\psi|| \geq |(X_i'\psi|P_i'\psi)| \geq |Im(X_i'\psi|P_i'\psi)| = \frac{1}{2}\left|(\psi|X_i'P_i'\psi) - \overline{(\psi|X_i'P_i'\psi)}\right|$$

$$= \frac{1}{2}\left|(\psi|(X_i'P_i' - P_i'X_i')\psi)\right| = \frac{\hbar}{2}(\psi|\psi) = \frac{\hbar}{2}\,.$$

Lemma 11.11 was used in the penultimate equality. So we have found (11.17). □

Remarks 11.14. This proof shows more generally that $\Delta A_\psi \Delta B_\psi \geq \frac{1}{2}|(\psi|[A,B]\psi)|$ for every vector $\psi \in D(AB) \cap D(BA) \cap D(A^2) \cap D(B^2)$ and Hermitian operators A, B on H. ■

11.3 Weyl's relations, the theorems of Stone–von Neumann and Mackey

The CCRs satisfy a remarkable property: in the statement of axiom **A5** it is somehow superfluous to ask the Hilbert space be $L^2(\mathbb{R}^3, dx)$ and that the position and momentum operators have the given form. This information is by some means contained in Heisenberg's relations so long as, loosely put, the representation of position and momentum is *irreducible*. This fact is the heart of the famous theorem of Stone–von Neumann, that we will prove in this section. By dropping irreducibility Mackey proved (as a consequence of more general facts in the theory of *imprimitivity systems*) that the Hilbert space is an orthogonal sum of irreducible representations (countably many if the space is separable). We will prove Mackey's theorem after Stone–von Neumann's.

11.3.1 Families of operators acting irreducibly and Schur's lemma

Before we get going, a few generalities on *families of operators acting irreducibly* are necessary. Tightly linked to this notion is *Schur's lemma*, a very useful result of abstract representation theory of unitary groups we will encounter in the next chapter.

Definition 11.15. *Let* H *be a Hilbert space and* $\mathscr{A} := \{A_i\}_{i\in J}$ *a family of operators* $A_i : \mathsf{H} \to \mathsf{H}$. *The space* H *is* **irreducible** *under* $\mathscr{A}$ *if there is no closed non-trivial subspace in* H *that is simultaneously invariant for all elements in* $\mathscr{A}$.

In other terms there cannot be any closed subspace $\mathsf{H}_0 \subset \mathsf{H}$, different from $\{0\}$ and H, for which $A_i(\mathsf{H}_0) \subset \mathsf{H}_0$ for every $i \in J$.

Here is *Schur's lemma.*

Proposition 11.16 (Schur's lemma). *Let $\mathscr{A} := \{A_i\}_{i\in J} \subset \mathfrak{B}(\mathsf{H})$ be a family of operators on a Hilbert space, closed under Hermitian conjugation ($A_i^* \in \mathscr{A}$ if $A_i \in \mathscr{A}$).*

(a) *$\mathscr{A}$ is irreducible $\Leftrightarrow$ every operator $V \in \mathfrak{B}(\mathsf{H})$ satisfying*

$$VA_i = A_i V \quad \textit{for every } i \in J,$$

has the form $V = \chi I$ for some complex number $\chi \in \mathbb{C}$.

(b) *Let $\mathscr{A}' := \{A'_i\}_{i\in J} \subset \mathfrak{B}(\mathsf{H}')$ be another family on the Hilbert space H', indexed by the same set J and closed under conjugation. Suppose*

$$A_i^* = A_{j_i} \textit{ implies } A_i'^* = A'_{j_i} \textit{ for every } i \in J \textit{ and some } j_i \in J. \tag{11.18}$$

If H and H' are irreducible, then every bounded linear operator $S : \mathsf{H} \to \mathsf{H}'$ such that

$$SA_i = A'_i S \quad \textit{for every } i \in J,$$

has the form $S = rU$ where $U : \mathsf{H} \to \mathsf{H}'$ is unitary and $r \in \mathbb{R}$ (in particular S is null when $r = 0$).

Proof. Let us begin with the more involved (b), which we will employ for (a).

(b) Taking adjoints of $SA_i = A'_i S$ gives $A_i^* S^* = S^* A_i'^*$ for any $i \in J$. That is to say $A_{j_i} S^* = S^* A'_{j_i}$, $i \in J$. Note how j_i covers J as i varies in J, since for every $A_i \in \mathscr{A}$, $(A_i^*)^* = A_i$, so we may rephrase the identity as $A_i S^* = S^* A'_i$ for every $i \in J$. Comparing with $SA_i = A'_i S$ gives $A_i S^* S = S^* S A_i$ and $A'_i SS^* = SS^* A'_i$. From the former the bounded self-adjoint operator $V := S^* S$ commutes with every A_i, so by Theorem 8.54(c) the spectral measure $P^{(V)}$ on $\mathbb{R}$ commutes with each A_i. But then every closed $P^{(V)}_E(\mathsf{H})$ is invariant under each A_i. As the space is irreducible, either $P^{(V)}_E = I$, or $P^{(V)}_E(\mathsf{H}) = \{0\}$ i.e. $P^{(V)}_E = 0$, for any Borel set $E \subset \mathbb{R}$. Suppose the spectrum of V contains at least two $\alpha \neq \alpha'$ and let us use Theorem 9.10(b). Consider two open disjoint real intervals $E \ni \alpha$, $E' \ni \alpha'$. Then $P^{(V)}_E \neq 0$, $P^{(V)}_{E'} \neq 0$ since the intervals intersect the spectrum, and therefore $P^{(V)}_E = P^{(V)}_{E'} = I$. On the other hand $P^{(V)}_E P^{(V)}_{E'} = 0$ since $E \cap E' = \varnothing$. This is absurd, so the spectrum of V (never empty) contains a single isolated point, which is in the point spectrum. Thus $S^* S = V = \lambda I$ for some $\lambda \in [0, +\infty)$, V being clearly positive. In a similar manner we obtain $SS^* = \lambda' I$ for some $\lambda' \in [0, +\infty)$. But then

$$\lambda S^* = S^* SS^* = \lambda' S^* .$$

Consequently either $\lambda = \lambda'$ or $S^* = 0$ and $S = (S^*)^* = 0$. In the second case the proof ends. In the first instance, let $U := \lambda^{-1/2} S$, so that $UU^* = I'$ and $U^* U = I$ where I, I' are the identity operators of H, H'. Therefore (see Definition 3.51) U is unitary. The claim is proved by taking $r = \lambda^{1/2}$.

Let us pass to (a) and assume H is $\mathscr{A}$-irreducible. If $VA_i = A_iV$, then $A_i^*V^* = V^*A_i^*$, meaning $A_iV^* = V^*A_i$ for any $i \in J$, as $\mathscr{A}$ is conjugate-closed. Then the bounded self-adjoint $V_+ := \frac{1}{2}(V+V^*)$ and $V_- := \frac{1}{2i}(V-V^*)$ commute with $\mathscr{A}$, implying that their spectral measures commute with $\mathscr{A}$. Arguing as in part (b) we conclude $V_\pm = \lambda_\pm I$ for some reals $\lambda_\pm$. Then $V = V_+ + iV_- = (\lambda_+ + i\lambda_-)I = \chi I$, $\chi \in \mathbb{C}$. Suppose, conversely, that the only operators commuting with $\mathscr{A}$ are χI. If H_0 is invariant under $\mathscr{A}$ and P is the orthogonal projector on H_0, then $PA_iP = A_iP$ for any $i \in J$. Take adjoints: $PA_i^*P = PA_i^*$. As $\mathscr{A}$ is conjugate-closed and $i \in J$ arbitrary, the identity reads $PA_iP = PA_i$. Comparing with the initial relation gives $PA_i = A_iP$, $i \in J$. Therefore $P = \chi I$ for a $\chi \in \mathbb{C}$. $P^* = P$ implies $\chi \in \mathbb{R}$, and $PP = P$ tells $\chi^2 = \chi$. So there are two possibilities: $P = 0$, and then $\mathsf{H}_0 = \{0\}$, or $P = I$ so $\mathsf{H}_0 = \mathsf{H}$. This means H is $\mathscr{A}$-irreducible. □

Remarks 11.17. Schur's lemma, in cases (a) and (b), is particularly useful in these situations:

(i) $\mathscr{A}$, $\mathscr{A}'$ are images of two representations $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$, $\pi' : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}')$ of the same *-algebra (or C^*-algebra) $\mathfrak{A}$;
(ii) $\mathscr{A}$, $\mathscr{A}'$ are images of two unitary representations $\mathsf{G} \ni g \mapsto U_g$, $\mathsf{G} \ni g \mapsto U'_g$ of one group G.

In either case, closure under conjugation in case (a), and (11.18) in case (b) are automatic if one takes respectively G and $\mathfrak{A}$ as the indexing set I. ■

11.3.2 Weyl's relations from the CCRs

In order to illustrate the Stone–von Neumann theorem we proceed step by step. A relevant technical point is that Heisenberg's commutation relations are too hard to use rigorously, for they involve subtleties about domains. To by-pass these issues we can pass from X_i and P_i to considering the one-parameter unitary groups they generate. Even better, we may take, for $n = 3$, the operators $\sum_{i=k}^n t_kX_k + u_kP_k$, $t_k, u_k \in \mathbb{R}$. These are essentially self-adjoint on $\mathscr{S}(\mathbb{R}^3)$, so we can look at the exponentials of their self-adjoint extensions $\overline{\sum_{i=k}^n t_kX_k + u_kP_k}$. A brute-force, direct, computation based on Heisenberg's (11.15) and the formal Taylor expansion of the exponential (not yet justified) yields the following identity[1]:

$$\exp\left\{ i\overline{\sum_{k=1}^n t_kX_k + u_kP_k} \right\} \exp\left\{ i\overline{\sum_{k=1}^n t'_kX_k + u'_kP_k} \right\}$$
$$= \exp\left\{ -\frac{i\hbar}{2}\left(\sum_{k=1}^n t_ku'_k - t'_ku_k \right) \right\} \exp\left\{ i\overline{\sum_{k=1}^n (t_k+t'_k)X_k + (u_k+u'_k)P_k} \right\} .$$

The above are called *Weyl relations*, and follow formally from Heisenberg's commutation relations.

[1] If the exponentiated operators were $n \times n$ complex matrices the result would follow from the celebrated *Baker–Campbell–Hausdorff formula*: $e^Ae^B = e^{[A,B]/2}e^{A+B}$, valid when the matrix $[A,B]$ commutes with both A and B.

The following proposition proves, completely independently from previous results that involve different techniques, that the operators X_i, P_i are essentially self-adjoint if restricted to $\mathscr{S}(\mathbb{R}^3)$, even in dimension higher than 3 and that the previously mentioned Weyl's relations hold rigorously. For conveniency, we will assume $\hbar = 1$ in the sequel.

Proposition 11.18. *Consider $L^2(\mathbb{R}^n, dx)$, with given $n = 1, 2, \cdots$ and Lebesgue measure dx on $\mathbb{R}^n$. For $k = 1, 2, \cdots, n$ define symmetric operators:*

$$\mathscr{X}_k : \mathscr{S}(\mathbb{R}^n) \to L^2(\mathbb{R}^n, dx) \quad \textit{and} \quad \mathscr{P}_k : \mathscr{S}(\mathbb{R}^n) \to L^2(\mathbb{R}^n, dx)$$

$$(\mathscr{X}_k \psi)(\mathbf{x}) = x_k \psi(\mathbf{x}), \tag{11.19}$$

$$(\mathscr{P}_k \psi)(\mathbf{x}) = -i \frac{\partial \psi}{\partial x_k}(\mathbf{x}). \tag{11.20}$$

Then:

(a) *the symmetric operators, defined on $\mathscr{S}(\mathbb{R}^n)$, $\sum_{k=1}^n t_k \mathscr{X}_k + u_k \mathscr{P}_k$ map $\mathscr{S}(\mathbb{R}^n)$ to itself and are essentially self-adjoint for any $(\mathbf{t}, \mathbf{u}) \in \mathbb{R}^{2n}$.*
(b) *$L^2(\mathbb{R}^n, dx)$ is irreducible under the family of bounded operators:*

$$W((\mathbf{t}, \mathbf{u})) := \exp\left\{ i \overline{\sum_{k=1}^n t_k \mathscr{X}_k + u_k \mathscr{P}_k} \right\}, \quad (\mathbf{t}, \mathbf{u}) \in \mathbb{R}^{2n}. \tag{11.21}$$

(c) *The operators W satisfy* **Weyl's relations***:*

$$W((\mathbf{t}, \mathbf{u}))W((\mathbf{t}', \mathbf{u}')) = e^{-\frac{i}{2}(\mathbf{t} \cdot \mathbf{u}' - \mathbf{t}' \cdot \mathbf{u})} W((\mathbf{t} + \mathbf{t}', \mathbf{u} + \mathbf{u}')), \quad W((\mathbf{t}, \mathbf{u}))^* = W(-(\mathbf{t}, \mathbf{u})). \tag{11.22}$$

(d) *For given $(\mathbf{t}, \mathbf{u}) \in \mathbb{R}^{2n}$, every mapping $\mathbb{R} \ni s \mapsto W(s(\mathbf{t}, \mathbf{u}))$ satisfies:*

$$s\text{-}\lim_{s \to 0} W(s(\mathbf{t}, \mathbf{u})) = W(0). \tag{11.23}$$

Proof. Let us begin with $n = 1$, for the generalisation to finite $n > 1$ is obvious. We will use tools from Chapter 9.1.4. At present we just have the operators $\mathscr{X}$ and $\mathscr{P}$. Both are well defined when restricted to the Schwartz space $\mathscr{S}(\mathbb{R})$, and admit each a self-adjoint extension that coincides with X, P, as previously discussed. We want to construct a dense subspace of analytic vectors for the symmetric operators $a\mathscr{X} + b\mathscr{P} : \mathscr{S}(\mathbb{R}) \to L^2(\mathbb{R}, dx)$, for any $a, b \in \mathbb{R}$. Define, on the dense domain $\mathscr{S}(\mathbb{R})$, the **annihilation operator**, **creation operator** and **number operator**:

$$A := \frac{1}{\sqrt{2}} \left(\mathscr{X} + \frac{d}{dx} \right), \quad A^\star := \frac{1}{\sqrt{2}} \left(\mathscr{X} - \frac{d}{dx} \right), \quad \mathscr{N} := A^\star A. \tag{11.24}$$

By construction $A^* \supset A^\star$, $(A^\star)^* \supset A$ and $\mathscr{N}$ is symmetric. By direct computation the CCRs (or the above definition) imply commutation relations on $\mathscr{S}(\mathbb{R})$, namely:

$$[A, A^\star] = I. \tag{11.25}$$

It is a well-known fact in the theory of orthogonal polynomials that the complete orthonormal system in $L^2(\mathbb{R}^n, dx)$ of Hermite functions $\{\psi_n\}_{n=0,1,...} \subset \mathscr{S}(\mathbb{R})$ (cf. Example 3.32(4)) satisfies $\psi_0 = \pi^{-1/4}e^{-x^2/2}$ and the recursive formula:

$$\psi_{n+1} = (2(n+1))^{-1/2}(x - \frac{d}{dx})\psi_n .$$

By definition of $A^\star$, that is the same as saying Hermite functions arise, once ψ_0 is given, from

$$\psi_n = \sqrt{\frac{1}{n!}}(A^\star)^n \psi_0 . \tag{11.26}$$

At the same time a straighforward computation produces

$$A\psi_0 = 0. \tag{11.27}$$

Equations (11.26), (11.27) and (11.25) justify, by induction, the middle relation in the triple:

$$A^\star \psi_n = \sqrt{n+1}\psi_{n+1} , \quad A\psi_n = \sqrt{n}\psi_{n-1} , \quad \mathscr{N}\psi_n = n\psi_n . \tag{11.28}$$

The right side in the second one is assumed null if $n = 0$, and as one sees easily the first identity is just the recusive relation introduced a few lines above; the third one follows from the other two.

As the ψ_n are normalised to 1, the first two in (11.28) give the inequality:

$$||A_1 A_2 \cdots A_k \psi_n|| \leq \sqrt{n+1}\sqrt{n+2}\cdots\sqrt{n+k} \leq \sqrt{(n+k)!} , \tag{11.29}$$

where every A_i is either A or $A^\star$. Consider a symmetric operator on $\mathscr{S}(\mathbb{R})$ given by an arbitrary real linear combination $T := a\mathscr{X} + b\mathscr{P}$, $a, b \in \mathbb{R}$. By (11.24), if $z := a + ib$ we have

$$T = \frac{\bar{z}A + zA^\star}{\sqrt{2}} . \tag{11.30}$$

This and (11.29) imply, for any Hermite function ψ_n:

$$||T^k\psi_n|| = 2^{-k/2}||(\bar{z}A + zA^\star)^k\psi_n|| \leq 2^{-k/2}2^k|z|^k\sqrt{(n+k)!} = |z|^k\sqrt{2^k(n+k)!} .$$

Hence, for $t \geq 0$:

$$\sum_{k=0}^{+\infty} \frac{t^k}{k!}||T^k\psi_n|| \leq \sum_{k=0}^{+\infty} \frac{(\sqrt{2}|z|t)^k\sqrt{(n+k)!}}{k!} \leq \sum_{k=0}^{+\infty} \frac{(\sqrt{2}|z|t)^k\sqrt{(n+k)^n}}{\sqrt{k!}} < +\infty .$$

The last series has finite sum by computing the convergence radius r via

$$1/r = \lim_{k\to+\infty}\left(\sqrt{\frac{(n+k)^n}{k!}}\right)^{1/k} = \lim_{k\to+\infty} e^{\frac{n\ln(k+n)}{2k} - \frac{\ln k!}{2k}} = \lim_{k\to+\infty} e^{-\frac{\ln k!}{2k}} = 0$$

(in the end we used Stirling's formula). Therefore any finite combination of Hermite functions is analytic for every $T := a\mathscr{X} + b\mathscr{P}$ on $\mathscr{S}(\mathbb{R})$. As the latter are symmetric, they must be essentially self-adjoint on $\mathscr{S}(\mathbb{R})$ by Nelson's theorem (Theorem 5.47). This ends the proof of case (a) for $n = 1$; if $n > 1$ the argument is similar, keeping in mind that Hermite functions in n variables:

$$\psi_{m_1,\ldots,m_n}(x_1,\ldots,x_n) := \psi_{m_1}(x_1)\cdots\psi_{m_n}(x_m)$$

are a complete orthonormal system in $L^2(\mathbb{R}^n, dx)$ (see Example 10.27(1)). What we have seen proves (a), but also (d) and the second identity in (c): in fact

$$W(s(\mathbf{t},\mathbf{u})) = \exp\left\{ is\overline{\sum_{k=1}^{n} t_k\mathscr{X}_k + u_k\mathscr{P}_k} \right\} = \exp\left\{ is\left(\overline{\sum_{k=1}^{n} t_k\mathscr{X}_k + u_k\mathscr{P}_k}\right) \right\}$$

by construction, because by definition of closable operator A we have $\overline{sA} = s\overline{A}$ for every $s \in \mathbb{C}$. Now, as $\overline{\sum_{i=1}^{n} t_k\mathscr{X}_k + u_k\mathscr{P}_k}$ is self-adjoint, Theorem 9.29(a) ensures strong continuity of the one-parameter unitary group $\mathbb{R} \ni s \mapsto W(s(\mathbf{t},\mathbf{u}))$, since $W(0) = \exp\left\{i0\overline{\sum_{i=1}^{n} t_k\mathscr{X}_k + u_k\mathscr{P}_k}\right\} = I$. The second identity in (c) is obvious since $\mathbb{R} \ni s \mapsto W(s(\mathbf{t},\mathbf{u}))$ is a one-parameter unitary group.

To prove (b) we shall invoke Lemma 11.19, which we will prove after the present theorem but relies only on part (a). Suppose there is a non-null closed $\mathsf{H}_0 \subseteq L^2(\mathbb{R}^n)$ invariant under $W((\mathbf{t},\mathbf{u}))$, and let $\psi \neq 0$ be an element. Calling $\phi \in \mathsf{H}_0^{\perp}$, we will prove $\phi = 0$ and so $\mathsf{H}_0 = L^2(\mathbb{R}^n)$. By assumption, H_0 and the orthogonal complement are invariant:

$$(\phi|W((\mathbf{t},0)W((0,\mathbf{u}))\psi) = 0\,, \quad \text{for any } (\mathbf{t},\mathbf{u}) \in \mathbb{R}^{2n}\,.$$

I.e.

$$\left(\phi\,\middle|\, e^{i\overline{\sum_k t_k\mathscr{X}_k}} e^{i\overline{\sum_k u_k\mathscr{P}_k}}\,\psi\right) = 0\,, \quad \text{for any } (\mathbf{t},\mathbf{u}) \in \mathbb{R}^{2n}\,.$$

The left side can be computed with (11.37), (11.38) in Lemma 11.19:

$$\int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}}\overline{\phi(\mathbf{x})}\psi(\mathbf{x}+\mathbf{u})\,dx = 0\,, \quad \text{for any } \mathbf{t},\mathbf{u} \in \mathbb{R}^n.$$

Since the map $\mathbf{x} \mapsto h_{\mathbf{u}}(\mathbf{x}) := \overline{\phi(\mathbf{x})}\psi(\mathbf{x}+\mathbf{u})$ is in $L^1(\mathbb{R}^n, dx)$, as product of $L^2(\mathbb{R}^n, dx)$ maps, and given $\mathbf{t} \in \mathbb{R}^n$ is arbitrary, the identity simply tells that the Fourier transform of $h_{\mathbf{u}} \in L^1(\mathbb{R}^n, dx)$ is zero. By Proposition 3.81(f) $h_{\mathbf{u}}$ is null almost everywhere. In other terms:

$$\overline{\phi(\mathbf{x})}\psi(\mathbf{x}+\mathbf{u}) = 0 \text{ almost everywhere for any given } \mathbf{u} \in \mathbb{R}^n\,. \tag{11.31}$$

Call $E \subset \mathbb{R}^n$ the set on which ψ is not null, and F the set where ϕ never vanishes. (Both are measurable as pre-images of the open set $\mathbb{C} \setminus \{0\}$ under measurable maps.) Denote by m the Lebesgue measure of $\mathbb{R}^n$, so $m(E) > 0$ by assumption. To satisfy (11.31) we must have:

$$m(F \cap (E - \mathbf{u})) = 0 \text{ for any } \mathbf{u} \in \mathbb{R}^n, \text{ i.e. } \int_{\mathbb{R}} \chi_F(\mathbf{x})\chi_E(\mathbf{x}+\mathbf{u})dx = 0 \text{ for any } \mathbf{u} \in \mathbb{R}^n\,.$$

Integrating in $\mathbf{u}$ gives

$$\int_{\mathbb{R}} dy \int_{\mathbb{R}} \chi_F(\mathbf{x})\chi_E(\mathbf{x}+\mathbf{u})dx = 0\,.$$

As integrands are non-negative and the double integral is finite, Fubini–Tonelli allows to swap integrals and we use Lebesgue's invariance under translations to obtain:

$$\begin{aligned}0 &= \int_{\mathbb{R}} dx\chi_F(\mathbf{x}) \int_{\mathbb{R}} \chi_E(\mathbf{x}+\mathbf{u})du = \int_{\mathbb{R}} dx\chi_F(\mathbf{x}) \int_{E-\mathbf{x}} 1du \\ &= \int_{\mathbb{R}} dx\chi_F(\mathbf{x}) \int_{\mathbb{R}} \chi_E(\mathbf{u})dy = m(F)m(E)\,.\end{aligned}$$

As $m(E) > 0$, we have $m(F) = 0$. Thus ϕ is null almost everywhere, hence the null vector of $L^2(\mathbb{R}^n, dx)$. So $\mathsf{H}_0 = L^2(\mathbb{R}^n, dx)$, proving irredubility in (b).

There remains to show

$$W((\mathbf{t},\mathbf{u}))W((\mathbf{t}',\mathbf{u}')) = e^{-\frac{i}{2}(\mathbf{t}\cdot\mathbf{u}'-\mathbf{t}'\cdot\mathbf{u})}W((\mathbf{t}+\mathbf{t}',\mathbf{u}+\mathbf{u}'))\,. \tag{11.32}$$

For this we need two steps. Introduce

$$U((\mathbf{t},\mathbf{u})) := e^{-\frac{i}{2}(\mathbf{t}\cdot\mathbf{u})}W((\mathbf{t},0))W((0,\mathbf{u}))\,.$$

Step one will prove that

$$U((\mathbf{t},\mathbf{u}))U((\mathbf{t}',\mathbf{u}')) = e^{-\frac{i}{2}(\mathbf{t}\cdot\mathbf{u}'-\mathbf{t}'\cdot\mathbf{u})}U((\mathbf{t}+\mathbf{t}',\mathbf{u}+\mathbf{u}'))\,. \tag{11.33}$$

Step two consists in showing

$$U((\mathbf{t},\mathbf{u})) = W((\mathbf{t},\mathbf{u}))\,, \tag{11.34}$$

which will conclude the overall proof.

Exactly as in part (b), Lemma 11.19 implies:

$$(U((\mathbf{t},\mathbf{u}))\psi)(\mathbf{x}) = e^{\frac{i}{2}\mathbf{t}\cdot\mathbf{u}}e^{i\mathbf{t}\cdot\mathbf{x}}\psi(\mathbf{x}+\mathbf{u})\,. \tag{11.35}$$

Hence, for given $\psi \in L^2(\mathbb{R}^n, dx)$,

$$U((\mathbf{t},\mathbf{u}))U((\mathbf{t}',\mathbf{u}'))\psi = e^{-\frac{i}{2}(\mathbf{t}\cdot\mathbf{u}'-\mathbf{t}'\cdot\mathbf{u})}U((\mathbf{t}+\mathbf{t}',\mathbf{u}+\mathbf{u}'))\psi\,.$$

This is the same as (11.33), which is eventually justified. Let us pass to (11.34). Consider, for $\mathbf{t},\mathbf{u}$ fixed, the unitary family $U_s := U(s(\mathbf{t},\mathbf{u}))$, $s \in \mathbb{R}$. Directly from (11.35) we compute $U_{s+s'} = U_sU_{s'}$ and $U_0 = I$. Therefore $\{U_s\}_{s\in\mathbb{R}}$ is a one-parameter unitary group. The strategy is now to prove that the group is strongly continuous, find its generator and show it concides with the generator of $\{W(s(\mathbf{t},\mathbf{u}))\}_{s\in\mathbb{R}}$. By Stone's theorem (Theorem 9.29) the two groups will be the same, hence (11.34). As for strong continuity, note that for any $\psi, \phi \in L^2(\mathbb{R}^2, dx)$:

$$\begin{aligned}(\phi|U_s\psi) &= e^{\frac{is^2}{2}\mathbf{t}\cdot\mathbf{u}}\left(\phi\,\middle|\, e^{is\overline{\sum_k t_k \mathscr{X}_k}}e^{is\overline{\sum_k u_k \mathscr{P}_k}}\psi\right) \\ &= e^{\frac{is^2}{2}\mathbf{t}\cdot\mathbf{u}}\left(e^{is\overline{\sum_k t_k \mathscr{X}_k}}\phi\,\middle|\, e^{is\overline{\sum_k u_k \mathscr{P}_k}}\psi\right) \to (\phi|\psi) \quad \text{as } s \to 0\,,\end{aligned}$$

because the scalar product is continuous, and one-parameter groups generated by the self-adjoint operators $\overline{\sum_k u_k \mathscr{P}_k}$ and $\overline{\sum_k t_k \mathscr{X}_k}$ are strongly continuous. Proposition 9.23 guarantees $\{U_s\}_{s\in\mathbb{R}}$ is strongly continuous. Consider $\psi \in \mathscr{S}(\mathbb{R}^n)$, and let us check

$$\lim_{s\to 0}\left|\left|\frac{U_s\psi-\psi}{s} - i\left(\sum_k t_k \mathscr{X}_k + u_k \mathscr{P}_k\right)\psi\right|\right|^2 = 0\,. \tag{11.36}$$

A few passages give

$$\left|\left|\frac{U_s\psi-\psi}{s} - i\left(\sum_k t_k \mathscr{X}_k + u_k \mathscr{P}_k\right)\psi\right|\right|^2$$

$$= \int_{\mathbb{R}^n}\left|\frac{e^{is^2\mathbf{t}\cdot\mathbf{u}/2}e^{is\mathbf{tx}}\psi(\mathbf{x}+s\mathbf{u})-\psi(\mathbf{x})}{s} - i\mathbf{t}\cdot\mathbf{x}\psi(\mathbf{x}) - \mathbf{u}\cdot\nabla_{\mathbf{x}}\psi\right|^2 dx\,.$$

Hence

$$\left|\left|\frac{U_s\psi-\psi}{s} - i\left(\sum_k t_k \mathscr{X}_k + u_k \mathscr{P}_k\right)\psi\right|\right|^2$$

$$\leq \int_{\mathbb{R}^n}\left|e^{is^2\mathbf{t}\cdot\mathbf{u}/2}e^{is\mathbf{tx}}\frac{\psi(\mathbf{x}+s\mathbf{u})-\psi(\mathbf{x})}{s} - \mathbf{u}\cdot\nabla_{\mathbf{x}}\psi\right|^2 dx$$

$$+2\int_{\mathbb{R}^n}\left|e^{is^2\mathbf{t}\cdot\mathbf{u}/2}e^{is\mathbf{tx}}\frac{\psi(\mathbf{x}+s\mathbf{u})-\psi(\mathbf{x})}{s} - \mathbf{u}\cdot\nabla_{\mathbf{x}}\psi\right|\left|\frac{e^{is^2\mathbf{t}\cdot\mathbf{u}/2}e^{is\mathbf{tx}}-1}{s\mathbf{t}\cdot\mathbf{x}} - i\right||\mathbf{t}\cdot\mathbf{x}\psi(\mathbf{x})|dx$$

$$+\int_{\mathbb{R}^n}\left|\frac{e^{is^2\mathbf{t}\cdot\mathbf{u}/2}e^{is\mathbf{t}\cdot\mathbf{x}}-1}{s\mathbf{t}\cdot\mathbf{x}} - i\right|^2 |\mathbf{t}\cdot\mathbf{x}\psi(\mathbf{x})|^2 dx\,.$$

Consider the integrals on the right. The middle one, by Schwarz's inequality, tends to zero when the other two do, because its square is less than the product of the other two. By dominated convergence the last integral is infinitesimal as $s \to 0$, because the integrand tends to 0 pointwise and is unifomrly bounded by the L^1 map $C|\mathbf{t}\cdot\mathbf{x}\psi(\mathbf{x})|^2$, for some constant $C > 0$. The first integrand also tends to 0 pointwise, as $s \to 0$. We want to use Lebesgue's theorem, so we need an L^1 upper bound, uniform in s around 0 (hence independent of s). Decomposing the integral and recalling $\psi \in \mathscr{S}(\mathbb{R}^n)$, it suffices to find an L^1 uniform bound in $s \in [-\varepsilon, \varepsilon]$ for the expressions

$$\left|\frac{\psi(\mathbf{x}+s\mathbf{u})-\psi(\mathbf{x})}{s}\right| \quad \text{and} \quad \left|\frac{\psi(\mathbf{x}+s\mathbf{u})-\psi(\mathbf{x})}{s}\right|^2$$

in order to obtain a bound of the whole integrand. Assume ψ real (if not, decompose ψ in real and imaginary parts) and invoke the mean value theorem:

$$\left|\frac{\psi(\mathbf{x}+s\mathbf{u})-\psi(\mathbf{x})}{s}\right| = \left|\mathbf{u}\cdot\nabla\psi|_{\mathbf{x}+s_0\mathbf{u}}\right|\,,$$

where $s_0 \in [-\varepsilon, \varepsilon]$. Since $\psi \in \mathscr{S}(\mathbb{R}^n)$, for any $p = 1, 2, \ldots$ there is $K_p \geq 0$ with

$$|\mathbf{u} \cdot \nabla \psi|_{\mathbf{x}}| \leq \frac{K_p}{1 + ||\mathbf{x}||^p} .$$

If we fix $\varepsilon > 0$, $\mathbf{u} \in \mathbb{R}^n$ and $p = 2, 3, \ldots$, there is $C_{p,\varepsilon} > 0$ such that

$$\frac{1}{1 + ||\mathbf{x} + s_0 \mathbf{u}||^p} \leq \frac{C_{p,\varepsilon}}{1 + ||\mathbf{x}||^{p-1}} \quad \text{for any } \mathbf{x} \in \mathbb{R}^n,\ s_0 \in [-\varepsilon, \varepsilon].$$

Therefore, for a certain constant $C \geq 0$:

$$\left| \frac{\psi(\mathbf{x} + s\mathbf{u}) - \psi(\mathbf{x})}{s} \right| \leq \frac{C}{1 + ||\mathbf{x}||^{n+1}} , \quad \mathbf{x} \in \mathbb{R}^n,\ s \in [-\varepsilon, \varepsilon].$$

The map on the right and its square are in $L^1(\mathbb{R}^n, dx)$, and this is what we wanted in order to apply Lebesgue's theorem. Hence (11.36) is proved.

Summing up, the self-adjoint generator of the strongly continuous group $\{U(s(\mathbf{t}, \mathbf{u}))\}_{s \in \mathbb{R}}$ coincides with the generator of $\{W(s(\mathbf{t}, \mathbf{u}))\}_{s \in \mathbb{R}}$ on $\mathscr{S}(\mathbb{R}^n)$. Since the second generator is essentially self-adjoint on that space, and as such it admits a unique self-adjoint extension, the generators coincide everywhere. Consequently the groups concide, for both arise by exponentiating the same self-adjoint generator. □

The proof of parts (b), (c) rely on the following lemma, itself a consequence of (a). We state it aside given its technical usefulness.

Lemma 11.19. *Retaining the assumptions of Proposition 11.18, if* $\psi \in L^2(\mathbb{R}^n, dx)$ *and* $\mathbf{t}, \mathbf{u} \in \mathbb{R}^n$*:*

$$\left(e^{i\overline{\sum_k t_k \mathscr{X}_k}} \psi \right)(\mathbf{x}) = e^{i\mathbf{t} \cdot \mathbf{x}} \psi(\mathbf{x}) , \tag{11.37}$$

and

$$\left(e^{i\overline{\sum_k u_k \mathscr{P}_k}} \psi \right)(\mathbf{x}) = \psi(\mathbf{x} + \mathbf{u}) . \tag{11.38}$$

Proof. By direct calculation the group $\{U_s\}_{s \in \mathbb{R}}$,

$$(U_s \psi)(\mathbf{x}) := e^{is\mathbf{t} \cdot \mathbf{x}} \psi(\mathbf{x}) , \quad \forall \psi \in L^2(\mathbb{R}^n, dx)$$

is strongly continuous and satisfies

$$-i \lim_{s \to 0} \frac{1}{s} (U_s \psi - \psi) = \left(\sum_k t_k \mathscr{X}_k \right) \psi$$

on $\mathscr{S}(\mathbb{R}^n)$. In fact:

$$\left|\left| \frac{1}{s} (U_s \psi - \psi) - i \left(\sum_k t_k \mathscr{X}_k \right) \psi \right|\right|^2 = \int_{\mathbb{R}^3} \left| \frac{e^{is\mathbf{t} \cdot \mathbf{x}} - 1}{s} - i\mathbf{t} \cdot \mathbf{x} \right|^2 |\psi(\mathbf{x})|^2 dx$$

$$= \int_{\mathbb{R}^3} \left| \frac{e^{is\mathbf{t}\cdot\mathbf{x}} - 1}{s\mathbf{t}\cdot\mathbf{x}} - i \right|^2 |\mathbf{t}\cdot\mathbf{x}|^2 |\psi(\mathbf{x})|^2 dx \to 0 \quad \text{as } s \to 0\,,$$

where we used three ingredients: $\mathbf{x} \mapsto |\mathbf{t}\cdot\mathbf{x}|^2|\psi(\mathbf{x})|^2$ is L^1 as $\psi \in \mathscr{S}(\mathbb{R}^n)$; the map

$$\mathbb{R} \times \mathbb{R}^3 \ni (s, \mathbf{x}) \mapsto \left| \frac{e^{is\mathbf{t}\cdot\mathbf{x}} - 1}{s\mathbf{t}\cdot\mathbf{x}} - i \right|^2$$

is bounded and pointwise (in $\mathbf{x}$) tends to 0, $s \to 0$; Lebesgue dominated convergence. By Stone's theorem the generator of U_s is a self-adjoint extension of $\sum_k t_k \mathscr{X}_k$. At the same time, $\sum_k t_k \mathscr{X}_k$ is essentially self-adjoint by (a) in the theorem above, so the unique extension is its closure. Thus $\{U_s\}_{s\in\mathbb{R}}$ is generated by $\overline{\sum_k t_k \mathscr{X}_k}$, proving (11.37).

Now the second identity. By (3.65)–(3.68), because the Fourier-Plancherel transform $\hat{\mathscr{F}}$ is a Fourier transform $\mathscr{F}$ on the $\mathscr{F}$-invariant $\mathscr{S}(\mathbb{R}^n)$:

$$\sum_k u_k \mathscr{P}_k = \hat{\mathscr{F}}^{-1} \sum_k u_k \mathscr{K}_k \, \hat{\mathscr{F}}\,,$$

where $\mathscr{K}_k$ is $\mathscr{X}_k$ (the new name reflects the fact the variable of the transformed map is $\mathbf{k}$ not $\mathbf{x}$). The Fourier transform is an isomorphism, so

$$\overline{\sum_k u_k \mathscr{P}_k} = \hat{\mathscr{F}}^{-1} \overline{\sum_k u_k \mathscr{K}_k} \, \hat{\mathscr{F}}\,.$$

By Corollary 9.33

$$e^{i\overline{\sum_k u_k \mathscr{P}_k}} = \hat{\mathscr{F}}^{-1} e^{i\overline{\sum_k u_k \mathscr{K}_k}} \, \hat{\mathscr{F}}\,. \tag{11.39}$$

Reducing to $\psi \in \mathscr{S}(\mathbb{R}^n)$, where $\hat{\mathscr{F}}$ and its inverse are computed by the Fourier integral and reduce to $\mathscr{F}$ and inverse (cf. Definition 3.79), equation (11.39) implies

$$\left(e^{i\overline{\sum_k u_k \mathscr{P}_k}}\psi\right)(\mathbf{x}) = \left(\mathscr{F}^{-1} e^{i\overline{\sum_k u_k \mathscr{K}_k}} \hat{\psi}\right)(\mathbf{x}) = \psi(\mathbf{x}+\mathbf{u})\,, \quad \text{for any } \psi \in \mathscr{S}(\mathbb{R}^n). \tag{11.40}$$

Recall $\mathscr{S}(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n, dx)$, and $\mathscr{S}(\mathbb{R}^n) \ni \psi_n \to \psi$ in L^2. Then $\psi_n(\,\cdot + \mathbf{u}) \to \psi(\,\cdot + \mathbf{u})$ is in L^2, because Lebesgue's measure is translation-invariant and the continuity of $e^{i\overline{\sum_k u_k \mathscr{P}_k}}$ implies (11.38) by (11.40). □

11.3.3 The theorems of Stone–von Neumann and Mackey

In this part we show how Weyl's relations, valid for bounded operators $W((\mathbf{t},\mathbf{u}))$, with $(\mathbf{t},\mathbf{u}) \in \mathbb{R}^{2n}$, that form an irreducible set on a complex Hilbert space H and such that $s \mapsto W(s(\mathbf{t},\mathbf{u}))$ are strongly continuous at $s = 0$, imply that H is isomorphic to $L^2(\mathbb{R}^n, dx)$ under the identification sending $W((\mathbf{t},\mathbf{u}))$ to $e^{i\overline{\sum_k t_k X_k + u_k P_k}}$. In particular, the Hilbert space H turns out to be separable.

The theorem will be stated in a slightly more general form, for which we need symplectic geometry.

Let us recall some facts about symplectic vector spaces.

Definition 11.20. *A pair* (X, σ) *is called a (real)* **symplectic vector space** *if* X *is a real vector space and the* **symplectic form** $\sigma : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ *is a bilinear, skew-symmetric and* **weakly non-degenerate** *map:* $\sigma(u, v) = 0\ \forall u \in \mathsf{X} \Rightarrow v = 0$.
If (X', σ') *is another symplectic vector space, we call a linear map* $f : \mathsf{X} \to \mathsf{X}'$ *a* **symplectic linear map** *if it preserves the symplectic forms:* $\sigma'(f(x), f(y)) = \sigma(x, y)$, $x, y \in \mathsf{X}$.
A **symplectomorphism** *is an invertible symplectic linear map.*

Note that any symplectic linear map $f : \mathsf{X} \to \mathsf{X}'$ is one-to-one (see Exercise 11.4), so the image $(f(\mathsf{X}), \sigma')$ is a symplectic subspace of (X', σ') isomorphic to (X, σ). If X is a normed space (infinite-dimensional), there exists a stronger concept of *non-degeneracy*: it requires (a) $\sigma(\cdot, v) \in \mathsf{X}'$ for any $v \in \mathsf{X}$, and (b) $\mathsf{X} \ni v \mapsto \sigma(\cdot, v) \in \mathsf{X}'$ is bijective. In finite dimension weak non-degeneracy is the same as this strong non-degeneracy.

The next result is due to Darboux (and is related to a more famous theorem on symplectic manifolds, which we shall not be concerned about [FaMa06]).

Theorem 11.21 (Darboux). *If* (X, σ) *is a (real) symplectic vector space with* $\dim \mathsf{X} = 2n$ *finite, there exists a basis (infinitely many, actually), called* **standard symplectic basis**, $\{\mathbf{e}_1, \cdots, \mathbf{e}_n, \mathbf{f}_1, \cdots, \mathbf{f}_n\} \subset \mathsf{X}$*, in which* σ *assumes the following* **canonical form***:*

$$\sigma(\mathbf{z}, \mathbf{z}') := \left(\sum_{i=1}^{n} t_i u_i' - t_i' u_i \right) \quad \text{for any } \mathbf{z}, \mathbf{z}' \in \mathsf{X}, \tag{11.41}$$

where $\mathbf{z} = \sum_{i=1}^n t_i \mathbf{e}_i + \sum_{i=1}^n u_i \mathbf{f}_i$*,* $\mathbf{z}' = \sum_{i=1}^n t_i' \mathbf{e}_i + \sum_{i=1}^n u_i' \mathbf{f}_i$*.*

It is not hard to prove that an automorphism of a symplectic vector space is a symplectomorphism if and only if it preserves Darboux bases.

Now we can state the Stone–von Neumann theorem, whose proof is postponed to after we have introduced Weyl *-algebras. In a dedicated section ensuing the proof we will make mathematically- and physically-related comments on the theorem.

Theorem 11.22 (Stone–von Neumann). *Let* H *be a Hilbert space and* (X, σ) *a (real)* $2n$*-dimensional symplectic vector space. Suppose* H *admits a family of operators* $\{W(\mathbf{z})\}_{\mathbf{z} \in \mathsf{X}} \subset \mathfrak{B}(\mathsf{H})$ *with the following properties:*

(a) H *is irreducible under* $\{W(\mathbf{z})\}_{\mathbf{z} \in \mathsf{X}}$.
(b) *The Weyl relations*

$$W(\mathbf{z})W(\mathbf{z}') = e^{-\frac{i}{2}\sigma(\mathbf{z},\mathbf{z}')} W((\mathbf{z}+\mathbf{z}')), \quad W(\mathbf{z})^* = W(-\mathbf{z}), \quad \mathbf{z}, \mathbf{z}' \in \mathsf{X} \tag{11.42}$$

hold.
(c) *For given* $\mathbf{z} \in \mathsf{X}$*, every mapping* $\mathbb{R} \ni s \mapsto W(s\mathbf{z})$ *satisfies*

$$s\text{-}\lim_{s \to 0} W(s\mathbf{z}) = W(0). \tag{11.43}$$

Then, in a given standard symplectic basis of X *for which* $\mathbf{z} \in \mathsf{X}$ *is determined by* $(\mathbf{t}^{(\mathbf{z})}, \mathbf{u}^{(\mathbf{z})}) \in \mathbb{R}^n \times \mathbb{R}^n$*, there exists a Hilbert space isomorphism* $S : \mathsf{H} \to L^2(\mathbb{R}^n, dx)$ *such that:*

$$S W(\mathbf{z}) S^{-1} := \exp\left\{ i \overline{\sum_{k=1}^{n} t_k^{(\mathbf{z})} \mathscr{X}_k + u_k^{(\mathbf{z})} \mathscr{P}_k} \right\}, \quad \textit{for any } \mathbf{z} \in \mathsf{X}. \tag{11.44}$$

where the symmetric operators $\mathscr{X}_i$*,* $\mathscr{P}_i$ *are as of Proposition 11.18.*

Consequently H must be necessarily separable, as $L^2(\mathbb{R}^n, dx)$ is.

To complement the Stone–von Neumann theorem we state another result, proved by Mackey, that treats reducible representations of the Weyl *-algebra. The notion of Hilbert sum used below is the one found in Definition 7.34.

Theorem 11.23 (Mackey). *Assume the hypotheses of Theorem 11.22, with (a) replaced by one of the following equivalent facts.*

(a1) *Every generator* $W(z)$*,* $\mathbf{z} \in \mathsf{X}$*, has trivial kernel.*
(a2) *Every generator* $W(z)$ *is unitary.*
(a3) $W(0)$ *is the identity operator on* H.

Then the Hilbert space H *is the Hilbert sum of a family (at most countable if* H *is separable) of closed, irreducible and* $W(z)$*-invariant subspaces. On each such component the Stone–von Neumann theorem holds with respect to the restricted operators* $W(z)$*.*

Important remark. With the Darboux theorem in mind, an alternative way to formulate the Stone–von Neumann theorem, more often encountered in the literature, goes as follows. Mackey's theorem has a similar reformulation as well, which we omit but the reader can easily reconstruct. ■

Theorem 11.24 (Alternative version of the Stone–von Neumann theorem). *Let* H *be a complex Hilbert space and suppose* $\{U(\mathbf{t})\}_{\mathbf{t}\in\mathbb{R}^n}, \{V(\mathbf{u})\}_{\mathbf{u}\in\mathbb{R}^n} \subset \mathfrak{B}(\mathsf{H})$ *satisfy the following properties.*

(a) H *is irreducible under* $\{U(\mathbf{t})\}_{\mathbf{t}\in\mathbb{R}^n} \cup \{V(\mathbf{u})\}_{\mathbf{u}\in\mathbb{R}^n}$.
(b) *The relations (also called Weyl relations):*

$$U(\mathbf{t})V(\mathbf{u}) = V(\mathbf{u})U(\mathbf{t})e^{i\mathbf{t}\cdot\mathbf{u}}, \quad \mathbf{t}, \mathbf{u} \in \mathbb{R}^n,$$
$$U(\mathbf{t})U(\mathbf{t}') = U(\mathbf{t}+\mathbf{t}') \quad V(\mathbf{u})V(\mathbf{u}) = V(\mathbf{u}+\mathbf{u}'), \quad \mathbf{t}, \mathbf{u}, \mathbf{t}'\mathbf{u}' \in \mathbb{R}^n$$

hold.
(c) *For any pair* $\mathbf{t} \in \mathbb{R}^n$*,* $\mathbf{u} \in \mathbb{R}^n$*:*

$$s\text{-}\lim_{s\to 0} U(s\mathbf{t}) = U(0) \quad \textit{and} \quad s\text{-}\lim_{s\to 0} V(s\mathbf{u}) = V(0).$$

Then there there exists an isomorphism $S_1 : \mathsf{H} \to L^2(\mathbb{R}^n, dx)$ *such that:*

$$S_1\, U(\mathbf{t})\, S_1^{-1} := \exp\left\{ i \overline{\sum_{k=1}^{n} t_k \mathscr{X}_k} \right\} \quad \textit{and} \quad S_1\, V(\mathbf{u})\, S_1^{-1} := \exp\left\{ i \overline{\sum_{k=1}^{n} u_k \mathscr{P}_k} \right\}$$

where the symmetric operators $\mathscr{X}_i$*,* $\mathscr{P}_i$ *are defined as in Proposition 11.18.*

Let us explain how the two versions are equivalent. Assume the Hilbert spaces H of the statements are the same. We begin by proving that Theorem 11.22 implies 11.24. From the hypotheses of 11.24 and its Weyl relations it is immediate to see the $W((\mathbf{t},\mathbf{u})) := e^{i\mathbf{t}\cdot\mathbf{u}/2}U(\mathbf{t})V(\mathbf{u})$ fulfill Theorem 11.22 over the symplectic vector space $(\mathbb{R}^n \times \mathbb{R}^n, \sigma_c)$, where σ_c is the symplectic form already in canonical form:

$$\sigma_c((\mathbf{t},\mathbf{u}),(\mathbf{t}',\mathbf{u}')) = \left(\sum_{i=1}^{n} t_i u_i' - t_i' u_i\right)$$

in the standard basis of $\mathbb{R}^n \times \mathbb{R}^n$. If we choose the symplectic basis to be the standard one on $\mathbb{R}^n \times \mathbb{R}^n$, then Theorem 11.22 implies Theorem 11.24 by taking $S_1 = S$.

So let us prove 11.24 implies 11.22. Choose a standard symplectic basis on X and identify elements in X with pairs $(\mathbf{t},\mathbf{u})$ in $\mathbb{R}^n \times \mathbb{R}^n$. If the $W((\mathbf{t},\mathbf{u}))$ satisfy Theorem 11.22, then the new operators $V(\mathbf{t}) := W((\mathbf{t},0))$ and $U(\mathbf{u}) := W((0,\mathbf{u}))$ fulfill Theorem 11.24. A direct computation shows 11.24 implies Theorem 11.22 for $S = S_1$.

11.3.4 The Weyl *-algebra

The statement of the Stone–von Neumann theorem contains an extremely important notion, both for the proof but also in view of further developments of QM towards quantum field theory. We are talking about *Weyl *-algebras*. Let us spend some time on this.

Definition 11.25. *Let X be a (non-trivial) real vector space of arbitrary dimension (possibily infinite) and $\sigma : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ a symplectic form on it. A *-algebra (Definition 3.40) $\mathscr{W}(\mathsf{X},\sigma)$ is called* **Weyl *-algebra of** (X,σ) *if there exists a family $\{W(u)\}_{u\in\mathsf{X}}$ of non-zero elements, called the* **generators** *, such that:*

(i) **Weyl's (commutation) relations***:*

$$W(u)W(v) = e^{-\frac{i}{2}\sigma(u,v)}W(u+v)\,, \quad W(u)^* = W(-u)\,, \quad u,v \in \mathsf{X} \qquad (11.45)$$

hold;

(ii) $\mathscr{W}(\mathsf{X},\sigma)$ *is* **generated** *by $\{W(u)\}_{u\in\mathsf{X}}$, i.e. $\mathscr{W}(\mathsf{X},\sigma)$ coincides with the linear span of finite combinations of finite products of $\{W(u)\}_{u\in\mathsf{X}}$.*

What we show now, amongst other things, is that a symplectic vector space (X,σ) determines a unique Weyl *-algebra up to *-isomorphisms (Definition 3.40).

Theorem 11.26. *Let X be a (non-trivial) real vector space of arbitrary dimension (possibily infinite) and $\sigma : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ a symplectic form.*

(a) *There exists, always, a Weyl *-algebra $\mathscr{W}(\mathsf{X},\sigma)$ associated to (X,σ).*
(b) *Any Weyl *-algebra $\mathscr{W}(\mathsf{X},\sigma)$ has a unit $\mathbb{I}$, and:*

$$W(0) = \mathbb{I}\,, \quad W(u)^* = W(-u) = W(u)^{-1}\,, \, u \in \mathsf{X}\,. \qquad (11.46)$$

The generators $\{W(u)\}_{u\in\mathsf{X}}$ are linearly independent, so in particular $W(u) \neq W(v)$ if $u \neq v$.

(c) *If $\mathscr{W}(\mathsf{X},\sigma)$, generated by $\{W(u)\}_{u\in\mathsf{X}}$, and $\mathscr{W}'(\mathsf{X},\sigma)$, generated by $\{W'(u)\}_{u\in\mathsf{X}}$, are Weyl *-algebras of (X,σ), there is a unique *-isomorphism $\alpha : \mathscr{W}(\mathsf{X},\sigma) \to \mathscr{W}'(\mathsf{X},\sigma)$, which is determined by imposing:*

$$\alpha(W(u)) = W'(u)\,, \quad \text{for any } u \in \mathsf{X}.$$

(d) *Every representation (Definition 3.48) of $\mathscr{W}(\mathsf{X},\sigma)$ on a Hilbert space H*

$$\pi : \mathscr{W}(\mathsf{X},\sigma) \to \mathfrak{B}(\mathsf{H})$$

is either faithful of null.
(e) *Let $\mathscr{W}(\mathsf{X}',\sigma')$ be a Weyl *-algebra of the symplectic vector space (X',σ'). If $f : \mathsf{X} \to \mathsf{X}'$ is a symplectic linear map, there exists a *-homomorphism (*-isomorphism if f is a symplectomorphism) $\alpha_f : \mathscr{W}(\mathsf{X},\sigma) \to \mathscr{W}(\mathsf{X}',\sigma')$ that is completely determined by:*

$$\alpha_f(W(u)) = W'(f(u))\,, \quad u \in \mathsf{X} \tag{11.47}$$

(with obvious notation). Furthermore, α_f is injective.

Proof. (a) Consider the $\mathbb{C}$-Hilbert space $\mathsf{H} := L^2(\mathsf{X},\mu)$ where μ is the counting measure of the set X. With $u \in \mathsf{X}$ consider $W(u) \in \mathfrak{B}(L^2(\mathsf{X},\mu))$ defined by $(W(u)\psi)(v) := e^{i\sigma(u,v)}\psi(u+v)$ for any $\psi \in L^2(\mathsf{X},\mu)$, $v \in \mathsf{X}$. It is immediate that such operators are non-null and satisfy Weyl's commutation relations (11.46), by using Hermitian conjugation as involution. Finite combinations of finite products form a Weyl *-algebra of (X,σ).
(b) From the first equation in (11.45) we have $W(u)W(0) = W(0) = W(0)W(u)$ and $W(u)W(-u) = W(0) = W(-u)W(u)$, because the $W(u)$ do not vanish and generate the whole *-algebra. Hence $W(0) = \mathbb{I}$ and $W(-u) = W(u)^{-1}$. The latter, bearing in mind the second equation in (11.45), implies $W(u)^* = W(u)^{-1}$. Now let us prove the generators' linear independence. Consider a subset of n generators $\{W(u_j)\}_{j=1,\ldots,n}$, with $u_1,\ldots,u_n$ all distinct, and let us show the $W(u_j)$ are independent. Over arbitrary subsets (and finite combinations) the claim is proved. Consider the null combination $\sum_{j=1}^n a_jW(u_j) = 0$ and let us prove, by induction, $a_j = 0$ for $j = 1,\ldots,n$. If $n = 1$ this is true as every $W(u)$ is non-null by definition. Suppose the claim holds for $n-1$ generators, however chosen, and let us prove the assertion for n. Without loss of generality (relabelling if necessary) we may assume, by contradiction, $a_n \neq 0$. Then $\sum_{j=1}^n a_jW(u_j) = 0$ implies

$$W(u_n) = \sum_{j=1}^{n-1} \frac{-a_j}{a_n} W(u_j)\,.$$

Consequently

$$\mathbb{I} = W(u_n)^*W(u_n) = \sum_{j=1}^{n-1} \frac{-a_j}{a_n} W(u_n)^*W(u_j) = \sum_{j=1}^{n-1} \frac{-a_j}{a_n} e^{-i\sigma(-u_n,u_j)/2} W(u_j - u_n)$$
$$= \sum_{j=1}^{n-1} b_j W(u_j - u_n)\,,$$

where $b_j := \frac{-a_j}{a_n} e^{-i\sigma(-u_n,u_j)/2}$. To prove the claim it suffices to show $b_j = 0$ for every $j = 1,2,\ldots,n-1$. To do so, let us fix a $u \in \mathsf{X}$, so by the above identity

$$\mathbb{I} = W(u)\mathbb{I}W(-u) = \sum_{j=1}^{n-1} b_j W(u)W(u_j - u_n)W(-u)$$
$$= \sum_{j=1}^{n-1} b_j e^{-i\sigma(u,u_j-u_n)/2} W(u_j - u_n)\,.$$

Comparing the expressions obtained for $\mathbb{I}$ we have

$$\sum_{j=1}^{n-1} b_j W(u_j - u_n) = \sum_{j=1}^{n-1} b_j e^{-i\sigma(u,u_j-u_n)/2} W(u_j - u_n)\,.$$

Multiply by $W(u_n)$ and simplify:

$$\sum_{j=1}^{n-1} b_j W(u_j) = \sum_{j=1}^{n-1} b_j e^{-i\sigma(u,u_j-u_n)/2} W(u_j)\,.$$

As the generators $W(u_j)$, $j = 1,2,\ldots,n-1$, are linearly independent, we have $b_j(1 - e^{-i\sigma(u,u_j-u_n)/2}) = 0$. If $b_j \neq 0$ for some j then we would have $1 = e^{-i\sigma(u,u_j-u_n)/2}$, and so $\frac{\sigma(u,u_j-u_n)}{2\pi} = k(u) \in \mathbb{Z}$. But the left-hand side is linear in $u \in \mathsf{X}$, so the mapping $\mathsf{X} \ni u \mapsto k(u)$ must be linear. Being $\mathbb{Z}$-valued it is the zero map. Therefore $\sigma(u, u_j - u_n) = 0$ for any $u \in \mathsf{X}$. Non-degeneracy of σ implies $u_j - u_n = 0$, an absurd.
(c) The Weyl generators are linearly independent, and the product of two is a complex multiple of a generator (by the first Weyl identity), whence generators form a basis for the Weyl *-algebra. Consider the unique linear map $\alpha : \mathscr{W}(\mathsf{X},\sigma) \to \mathscr{W}'(\mathsf{X},\sigma)$ defined by $\alpha(W(u)) = W'(u)$ for any $u \in \mathsf{X}$. As $\{W(u)\}_{u\in\mathsf{X}}$ and $\{W'(u)\}_{u\in\mathsf{X}}$ are bases of the corresponding *-algebras, α is a vector-space isomorphism. But products of elements of the two *-algebras are combinations of the generators, by the first set of Weyl relations (the same for both *-algebras), so α must preserve products. $\alpha(W(0)) = W'(0)$ implies α preserves multiplicative neutral elements. Eventually $\alpha(W(-u)) = W'(-u)$ and the second Weyl set imply α commutes with involutions as well. The procedure also shows that α is uniquely determined by fixing $\alpha(W(u)) = W'(u)$ for every $u \in \mathsf{X}$.
(d) Consider a representation $\pi : \mathscr{W}(\mathsf{X},\sigma) \to \mathfrak{B}(\mathsf{H})$. By construction the operators $\{\pi(W(u))\}_{u\in\mathsf{X}}$ satisfy Weyl's relations. If every $\pi(W(u))$ is non-null, they define a Weyl *-algebra of (X,σ). By part (c) the representation π, when the codomain restricts to $\pi(\mathscr{W}(\mathsf{X},\sigma))$, is a *-isomorphism, making π injective. If, on the contrary, $\pi(W(u)) = 0$ for some $u \in \mathsf{X}$, then π is the zero representation. That is because if $z \in \mathsf{X}$, setting $z - u =: v$ implies $\pi(W(z)) = e^{\frac{i}{2}\sigma(u,v)}\pi(W(u))\pi(W(v)) = e^{\frac{i}{2}\sigma(u,v)} 0 \pi(W(v)) = 0$ by Weyl's relations. Hence π is null as the $W(v)$ form a basis for $\mathscr{W}(\mathsf{X},\sigma)$.
(e) As the generators of the Weyl *-algebra form a basis, as we said in (c), there is one and only one linear map $\alpha_f : \mathscr{W}(\mathsf{X},\sigma) \to \mathscr{W}(\mathsf{X}',\sigma')$, completely determined by (11.47). Using the Weyl relations, recalling f preserves symplectic forms, we obtain α_f is a *-homomorphism. Its uniqueness is clear, since any *-homomorphism

is linear, and (11.47) determine α_f for they fix its values on given bases. Injectivity goes like this: if $\alpha(\sum_i a_i W(u_i)) = 0$ (summing over an arbitrary, finite, set) then $\sum_i a_i \alpha(W(u_i)) = 0$, i.e. $\sum_i a_i W'(f(u_i)) = 0$, where $f(u_i) \neq f(u_j)$ for $i \neq j$ as f is one-to-one (σ' is weakly non-degenerate). Since the $W'(u')$ are linearly independent, $a_i = 0$ for every i and $\alpha(\sum_i a_i W(u_i)) = 0$ implies $\sum_i a_i W(u_i)$, as we wanted. □

Remark 11.27. (1) In the sense of (a), (c) above, the pair (X, σ) and equations (11.45) determine the Weyl *-algebra of (X, σ) universally (up to isomorphisms). Any concrete Weyl *-algebra of (X, σ) is sometimes called a **realisation** of the Weyl *-algebra of (X, σ). A **realisation on a Hilbert space** H is a *faithful representation* of the abstract Weyl *-algebra on H.
(2) If $\mathscr{W}(\mathsf{X}, \sigma)$ is a realisation of the Weyl *-algebra of (X, σ) by means of operators on the Hilbert space H, *it is in general false that* the identity operator $I : \mathsf{H} \to \mathsf{H}$ coincides with the neutral element $\mathbb{I} = W(0)$ of the algebra. The matter is relevant also because Weyl's relations imply $\mathbb{I} = I$ *if and only if every generator* $W(u)$ *is a unitary operator*.

Let us show a simple counterexample. Suppose $\mathscr{W}(\mathsf{X}, \sigma)$ is a representation on the Hilbert space $(\mathsf{H}, (\cdot|\cdot))$ of the Weyl *-algebra of (X, σ), where $W(0) = \mathbb{I} = I$. Consider the Hilbert space $\mathsf{H}' := \mathsf{H} \oplus \mathbb{C}$ with product $\langle(\psi, z)|(\psi', z')\rangle = (\psi|\psi') + \overline{z}z$. A representation on H' of the Weyl *-algebra of (X, σ) is generated by the operators $W(u)' : (\psi, z) \mapsto (W(u)\psi, 0)$, where the $W(u)$ generate the Weyl *-algebra on H. In this case $\mathbb{I} = W(0)' : (\psi, z) \mapsto (\psi, 0)$, so $W(0)'$ is not the identity on H', but just the orthogonal projector on H. Nevertheless, there exist also representations for which generators are unitary; for instance, the so-called *GNS representations*, that we will encounter later,[2] are fundamental in formulations of quantum field theories.
(3) *If* $\mathscr{W}(\mathsf{V}, \sigma)$ *is a realisation on the Hilbert space* H *of the Weyl* **-algebra of* (V, σ) *and* H *is irreducible for the generating set* $\{W(u)\}_{u \in \mathsf{V}}$*, then* $\mathbb{I} = I$ *and the* $W(u)$ *are unitary.*

Note $W(u) \neq 0$, for any $u \in \mathsf{V}$, for otherwise the representation would be null, hence reducible. If $\mathbb{I} \neq I$, the operator $W(0) = \mathbb{I}$ would be a projector, different from the zero and the identity of H, that commutes with every $W(u)$ because $W(u)W(0) = W(0) = W(0)W(u)$. Then the closed projection space of $W(0)$ (by assumption other than $\{0\}$ and H) would be invariant under every $W(u)$, hence $\mathscr{W}(\mathsf{V}, \sigma)$-invariant, contradicting the assumption.
(4) *If* $\mathscr{W}(\mathsf{V}, \sigma)$ *is a realisation on the Hilbert space* H *of the Weyl* **-algebra of* (V, σ)*, then* $\mathbb{I} = I$ *(the* $W(u)$ *are unitary) precisely when each generator* $\{W(u)\}_{u \in \mathsf{V}}$ *has trivial kernel* $\{0\}$.

The proof is straightforward. If every $W(u)$ has trivial null space, the orthogonal projector $W(0) = \mathbb{I}$ has trivial kernel and must coincide with the projector I. Conversely if $W(0) = I$ then the $W(u)$ are unitary, hence their null spaces are trivial.
(5) The Weyl *-algebra $\mathscr{W}(\mathsf{V}, \sigma)$ of a symplectic vector space (V, σ) admits a norm rendering the algebra's Banach completion a C^*-algebra: the **Weyl C^*-algebra** of (V, σ). Take, for example, the closure of the realisation of (V, σ) in $\mathfrak{B}(L^2(\mathsf{X}, \mu))$ de-

[2] In such a case the concrete construction of the representation and the existence of a cyclic vector force $W(u)$ to be unitary.

scribed in the proof of Theorem 11.26(a). The important fact, proved in Chapter 14, is that this C^*-algebra is determined by (V,σ), for one can prove there is a unique norm on a Weyl *-algebra satisfying the C^* identity $||a^*a|| = ||a||^2$. Moreover, the *-isomorphism of Theorem 11.26(c) extends to an (isometric) *-isomorphism of the C^*-algebras. Weyl C^*-algebras are but one starting point to build the quantum theory of Bosonic fields [BrRo02]. See [Str05a] for an example of C^*-algebras used in QM. ■

11.3.5 Proof of the theorems of Stone–von Neumann and Mackey

In this section we prove the Stone–von Neumann theorem as given by 11.22, and then Mackey's Theorem 11.23. Part of the arguments are mere reworkings of the analogous in [Str05a]. We will make a few remarks at the end, of mathematical nature and physical alike, on the relevance of these results and their broader reach.

Proof of Theorem 11.22 (Stone–von Neumann). Begin by observing every operator $W(\mathbf{z}) \in \mathfrak{B}(\mathsf{H})$ is non-zero: for if $W(\mathbf{z}_0) = 0$, for every $\mathbf{z} \in \mathsf{X}$ with $\mathbf{z} - \mathbf{z}_0 =: \mathbf{v}$, we would have $W(\mathbf{z}) = e^{\frac{i}{2}\sigma(\mathbf{z}_0,\mathbf{v})} W(\mathbf{z}_0)W(\mathbf{v}) = e^{\frac{i}{2}\sigma(\mathbf{z}_0,\mathbf{v})} 0 W(\mathbf{v}) = 0$. Then H would not be irreducible for the entire family $W(\mathbf{z}) \in \mathfrak{B}(\mathsf{H})$. By Definition 11.25, the set of $W(\mathbf{z}) \in \mathfrak{B}(\mathsf{H})$ is a generating system for a realisation of a Weyl *-algebra $\mathfrak{A}$ of the symplectic vector space (X,σ). This is given by finite combinations of finite products of the $W(\mathbf{z})$ and realised as the image of a faithful representation $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ of $\mathfrak{A}$. Fix a basis in X, so to associate bijectively every $\mathbf{z} \in \mathsf{X}$ to its components $(\mathbf{t}^{(\mathbf{z})},\mathbf{u}^{(\mathbf{z})}) \in \mathbb{R}^n \times \mathbb{R}^n$. Consider the Hilbert space $L^2(\mathbb{R}^n, dx)$. The family of non-null (unitary) operators $\left\{\exp\left\{i\overline{\sum_{k=1}^n t_k^{(\mathbf{z})}\mathscr{X}_k + u_k^{(\mathbf{z})}\mathscr{P}_k}\right\}\right\}_{\mathbf{z}\in\mathsf{X}}$ defines, by Proposition 11.18, another realisation of the same *-algebra $\mathfrak{A}$ and a corresponding faithful representation $\pi_1 : \mathfrak{A} \to \mathfrak{B}(L^2(\mathbb{R}^n, dx))$. We denote by $a_\mathbf{z} \in \mathfrak{A}$ the generators of $\mathfrak{A}$, so that $\pi(a_\mathbf{z}) = W(\mathbf{z})$ and also $\pi_1(a_\mathbf{z}) = \exp\left\{i\overline{\sum_{k=1}^n t_k^{(\mathbf{z})}\mathscr{X}_k + u_k^{(\mathbf{z})}\mathscr{P}_k}\right\}$ for any $\mathbf{z} \in \mathsf{X}$. Suppose now there are two non-zero vectors $\Phi_0 \in \mathsf{H}$, $\Psi_0 \in L^2(\mathbb{R}^n, dx)$ such that: (i) $\mathscr{D} := \pi(\mathfrak{A})\Phi_0$ is dense in H, (ii) $\mathscr{D}_1 := \pi_1(\mathfrak{A})\Psi_0$ is dense in $L^2(\mathbb{R}^n, dx)$, and (iii):

$$(\Phi_0|\pi(a)\Phi_0) = (\Psi_0|\pi_1(a)\Psi_0)\,, \quad a \in \mathfrak{A}. \tag{11.48}$$

Let us show that, consequently, there is a linear map $\widetilde{S} : \mathscr{D} \to \mathscr{D}_1$

$$\widetilde{S}\pi(a)\Phi_0 = \pi_1(a)\Psi_0\,, \quad a \in \mathfrak{A}, \tag{11.49}$$

extending by continuity to a Hilbert isomorphism from H to $L^2(\mathbb{R}^n, dx)$ satisfying (11.44), and thus proving the theorem.

The mapping is well defined: suppose $\pi(a)\Phi_0 = \pi(b)\Phi_0$. For (11.49) to be well defined we must have $\pi_1(a)\Psi_0 = \pi_1(b)\Psi_0$. From $\pi(a)\Phi_0 = \pi(b)\Phi_0$ follows, for any $c \in \mathfrak{A}$:

$$(\pi(c)\Phi_0|\pi(a)\Phi_0) = (\pi(c)\Phi_0|\pi(b)\Phi_0)\,.$$

Since π is a representation of *-algebras, so $\pi(c^*) = \pi(c)^*$ and $\pi(f)\pi(d) = \pi(fd)$, the displayed equation is equivalent to

$$(\Phi_0|\pi(c^*a)\Phi_0) = (\Phi_0|\pi(c^*b)\Phi_0)$$

so by (11.48) we have $(\Psi_0|\pi_1(c^*a)\Psi_0) = (\Psi_0|\pi_1(c^*b)\Psi_0)$. Proceeding backwards, for any $c \in \mathfrak{A}$:

$$(\pi_1(c)\Psi_0|\pi_1(a)\Psi_0) = (\pi_1(c)\Psi_0|\pi_1(b)\Psi_0)\ .$$

As $\pi_1(c)\Psi_0$ roams the dense space $\mathscr{D}_1$, necessarily $\pi_1(a)\Psi_0 = \pi_1(b)\Psi_0$, as required. Therefore $\widetilde{S}$ in (11.49) is well defined. It is immediate to see that $\widetilde{S}$ is linear, for π, π_1 are representations. By construction $\widetilde{S}$ preserves the inner product, and so is isometric:

$$\left(\widetilde{S}\pi(a)\Phi_0|\widetilde{S}\pi(b)\Phi_0\right) = (\pi_1(a)\Psi_0|\pi_1(b)\Psi_0) = (\Psi_0|\pi_1(a)^*\pi_1(b)\Psi_0)$$

$$= (\Psi_0|\pi_1(a^*)\pi_1(b)\Psi_0) = (\Psi_0|\pi_1(a^*b)\Psi_0) = (\Phi_0|\pi(a^*b)\Phi_0)$$

$$= (\Phi_0|\pi(a^*)\pi(b)\Phi_0) = (\Phi_0|\pi(a)^*\pi(b)\Phi_0) = (\pi(a)\Phi_0|\pi(b)\Phi_0)\ .$$

By Proposition 2.44 we can extend, by linearity and continuity, the transformation $\widetilde{S}$ from the dense domain $\mathscr{D}$ to the Hilbert space, obtaining a linear map $S : \mathsf{H} \to L^2(\mathbb{R}^n, dx)$. The extension S stays isometric by inner product's continuity. Similarly, we can construct on the dense $\mathscr{D}_1$ first, then on $L^2(\mathbb{R}^n, dx)$, a linear isometry $S' : L^2(\mathbb{R}^n, dx) \to \mathsf{H}$ by extending

$$\widetilde{S}'\pi_1(a)\Psi_0 = \pi(a)\Phi_0 \quad \text{for any } a \in \mathfrak{A}. \tag{11.50}$$

Since $\widetilde{S}\widetilde{S}' = I_{\mathscr{D}_1}$, $\widetilde{S}'\widetilde{S} = I_{\mathscr{D}}$ on the dense spaces $\mathscr{D}_1$, $\mathscr{D}$, these are valid by continuity on the extended domains: $SS' = I_{L^2(\mathbb{R}^n,dx)}$, $S'S = I_{\mathsf{H}}$. Overall, $S : \mathsf{H} \to L^2(\mathbb{R}^n, dx)$ is a Hilbert isomorphism satisfying

$$S\pi(a)\Phi_0 = \pi_1(a)\Psi_0 \quad \text{for any } a \in \mathfrak{A}. \tag{11.51}$$

Invert the identity for $b \in \mathfrak{A}$ to obtain $\pi(b)\Phi_0 = S^{-1}\pi_1(b)\Psi_0$. Substituting in (11.51), and replacing $\pi(a)$ by $\pi(ab) = \pi(a)\pi(b)$ on the left and $\pi_1(a)$ by $\pi_1(ab) = \pi_1(a)\pi_1(b)$ on the right, finally produces:

$$S\pi(a)S^{-1}\,\pi_1(b)\Psi_0 = \pi_1(a)\,\pi_1(b)\Psi_0\,.$$

The vectors $\pi_1(b)\Psi_0$ define a dense space in $L^2(\mathbb{R}^n, dx)$, so

$$S\pi(a)S^{-1} = \pi_1(a) \quad \text{for any } a \in \mathfrak{A}\,.$$

Picking as $a \in \mathfrak{A}$ a generic Weyl generator transforms the identity into (11.44).

To end the proof we have to exhibit vectors Φ_0, Ψ_0 satisfying (11.48) and generating, under the respective representations, dense subspaces. Let Φ_0 be any non-zero vector. The closed $\overline{\pi(\mathfrak{A})\Phi_0}$ is invariant by any $\pi(a)$, and in particular by any

$\pi(W(\mathbf{z}))$, by construction. Since H is irreducible for these vectors, then $\overline{\pi(\mathfrak{A})\Phi_0} = \mathsf{H}$, i.e. $\mathscr{D} := \pi(\mathfrak{A})\Phi_0$ is dense in H. A similar argument says $\mathscr{D}_1 := \pi_1(\mathfrak{A})\Psi_0$ is dense in $L^2(\mathbb{R}^n, dx)$ for every non-zero $\Psi_0 \in L^2(\mathbb{R}^n, dx)$. There remains to determine Φ_0, Ψ_0 fulfilling (11.48). Consider in $L^2(\mathbb{R}^n, dx)$ the vector

$$\Psi_0(\mathbf{x}) = \psi_0(x_1)\cdots\psi_0(x_n) = \pi^{-n/4}e^{-|\mathbf{x}|^2/2}$$

where ψ_0 is the first Hermite function. A straightforward calculation based on Lemma 11.11 gives

$$\left(\Psi_0 \middle| \exp\left\{i\overline{\sum_{k=1}^n t_k\mathscr{X}_k + u_k\mathscr{P}_k}\right\}\Psi_0\right) = \pi^{-n/2}\int_{\mathbb{R}^n} e^{i\mathbf{t}\cdot\mathbf{x}}e^{-|\mathbf{x}+\mathbf{u}|^2/2}dx = e^{-|\mathbf{t}|^2/4-|\mathbf{u}|^2/4}$$

and so

$$\left(\Psi_0 \middle| \exp\left\{i\overline{\sum_{k=1}^n t_k\mathscr{X}_k + u_k\mathscr{P}_k}\right\}\Psi_0\right) = e^{-(|\mathbf{t}|^2+|\mathbf{u}|^2)/4}, \quad \text{for any } (\mathbf{t},\mathbf{u}) \in \mathbb{R}^n\times\mathbb{R}^n. \tag{11.52}$$

If we manage to find a $\Phi_0 \in \mathsf{H}$ such that

$$(\Phi_0|W(\mathbf{z})\,\Phi_0) = e^{-(|\mathbf{t}^{(\mathbf{z})}|^2+|\mathbf{u}^{(\mathbf{z})}|^2)/4}, \quad \text{for any } \mathbf{z}\in\mathsf{X}, \tag{11.53}$$

then (11.48) holds by linearity, as any $\pi_1(a)$ is a combination of elements $\pi_1(a_\mathbf{z})$ and the corresponding $\pi(a)$ is a combination (same coefficients) of elements $\pi(a_\mathbf{z})$. At this point the existence of such a Φ_0 is warranted by the next proposition.

Proposition 11.28. *Under the assumptions of Theorem 11.22, if a basis on* X *has been fixed so to map every* $\mathbf{z}\in\mathsf{X}$ *to its components* $(\mathbf{t}^{(\mathbf{z})},\mathbf{u}^{(\mathbf{z})})\in\mathbb{R}^n\times\mathbb{R}^n$*, there exists* $\Phi_0\in\mathsf{H}$ *satisfying (11.53).*

Proof. First, the operators $W(\mathbf{z})$ are unitary with $W(0) = I$, by Remark 11.27(3) and because H is $W(\mathbf{z})$-irreducible. We claim $\mathsf{X}\ni\mathbf{z}\mapsto W(\mathbf{z})$ is continuous in the strong topology (the regularity assumption s-$\lim_{s\to0}W(s\mathbf{z}) = W(0) = I$ is only apparently weaker than strong continuity at $\mathbf{z}=0$, since the limit might not be uniform along directions tending to the origin). Let us set $W((\mathbf{t}^{(\mathbf{z})},\mathbf{u}^{(\mathbf{z})})) := W(\mathbf{z})$ in the sequel. Let us begin by proving $\mathbb{R}^n\ni\mathbf{t}\mapsto W((\mathbf{t},0))$ and $\mathbb{R}^n\ni\mathbf{u}\mapsto W((0,\mathbf{t}))$ are strongly continuous. We will prove it for $\mathbb{R}^n\ni\mathbf{t}\mapsto W((\mathbf{t},0))$ only, as the other case is identical. Weyl's relations imply additivity: $W((\mathbf{t},0))W((\mathbf{t}',0)) = W((\mathbf{t}+\mathbf{t}',0))$. If $\mathbf{e}_1,\dots,\mathbf{e}_n$ are the basis vectors expressing $\mathbf{t}=\sum_{k=1}^n t_k\mathbf{e}_k$ we can write $W((\mathbf{t},0)) = W((t_1\mathbf{e}_1,0))\cdots W((t_n\mathbf{e}_n,0))$. Each map $\mathbb{R}\ni t_k\mapsto W((t_k\mathbf{e}_k,0))$ is strongly continuous by regularity, i.e. s-$\lim_{s\to0}W(s\mathbf{z}) = W(0) = I$ in Theorem 11.22. Take $\psi\in\mathsf{H}$ and let us show $||W((\mathbf{t},0))\psi-\psi|| \to 0$ as $\mathbf{t}\to 0$. We have

$$||W((\mathbf{t},0))\psi-\psi|| = \left|\left|\prod_{k=1}^{n} W((t_k\mathbf{e}_k,0))\psi-\psi\right|\right|$$
$$\leq \left|\left|\prod_{k=1}^{n} W((t_k\mathbf{e}_k,0))\psi-\prod_{k=1}^{n-1} W((t_k\mathbf{e}_k,0))\psi\right|\right|$$
$$+\left|\left|\prod_{k=1}^{n-1} W((t_k\mathbf{e}_k,0))\psi-\prod_{k=1}^{n-2} W((t_k\mathbf{e}_k,0))\psi\right|\right| + \cdots + ||W((t_1\mathbf{e}_1,0))\psi-\psi||$$
$$= ||W((t_n\mathbf{e}_n,0))\psi-\psi|| + ||W((t_{n-1}\mathbf{e}_{n-1},0))\psi-\psi|| + \cdots + ||W((t_1\mathbf{e}_1,0))\psi-\psi||\,.$$

In the last passage we used that $W((t_k\mathbf{e}_k,0))$ is unitary, so it preserves the norm; in particular

$$\left|\left|\prod_{k=1}^{n} W((t_k\mathbf{e}_k,0))\psi-\prod_{k=1}^{n-1} W((t_k\mathbf{e}_k,0))\psi\right|\right| = \left|\left|\prod_{k=1}^{n-1} W((t_k\mathbf{e}_k,0))\,(W((t_n\mathbf{e}_n,0))\psi-\psi)\right|\right|$$
$$= ||W((t_n\mathbf{e}_n,0))\psi-\psi||\,.$$

The inequality

$$||W((\mathbf{t},0))\psi-\psi|| \leq \sum_{k=1}^{n} ||W((t_k\mathbf{e}_k,0))\psi-\psi||$$

and the continuity of $W((t_k\mathbf{e}_k,0))\psi$ for $t_k\to 0$ imply

$$W((\mathbf{t},0))\psi \to \psi \quad \text{as } \mathbf{t}\to 0,$$

working on products of intervals along the Cartesian axes as neighbourhoods of $\mathbf{z}=0$. Therefore the function $\mathsf{X}\ni\mathbf{z}\mapsto(\phi_1\,|W(\mathbf{z})\,\phi_2) = \left(W((\mathbf{t}^{(\mathbf{z})},0))^*\phi_1\,\middle|\,W((0,\mathbf{u}^{(\mathbf{z})}))\,\phi_2\right)$ is continuous at $\mathbf{z}=0$ for any ϕ_1,ϕ_2. Hence $\mathsf{X}\ni\mathbf{z}\mapsto W(\mathbf{z})$ is strongly continuous everywhere, in fact

$$||W(\mathbf{z})\phi - W(\mathbf{z}_0)\phi||^2 = ||e^{i\sigma(\mathbf{z}_0,\mathbf{z})/2}W(\mathbf{z}-\mathbf{z}_0)\phi-\phi||^2$$
$$= 2||\phi||^2 - e^{-i\sigma(\mathbf{z}_0,\mathbf{z})/2}\overline{(\phi|W(\mathbf{z}-\mathbf{z}_0)\phi)} - e^{i\sigma(\mathbf{z}_0,\mathbf{z})/2}(\phi|W(\mathbf{z}-\mathbf{z}_0)\phi) \to 0 \text{ as } \mathbf{z}\to\mathbf{z}_0,$$

for any $\phi\in\mathsf{H}$, by the Weyl relations and the unitarity of $W(\mathbf{z})$. We can then apply Proposition 9.27 and define

$$P := (2\pi)^{-n}\int_{\mathbb{R}^{2n}} dz e^{-|\mathbf{z}|^2/4}W(\mathbf{z})\,. \tag{11.54}$$

By construction $P\in\mathfrak{B}(\mathsf{H})$, and Proposition 9.27 implies $P^*=P$:

$$(\phi_1\,|P^*\phi_2\,) = \overline{(\phi_2|P\phi_1)} = (2\pi)^{-n}\int_{\mathbb{R}^{2n}} e^{-|\mathbf{z}|^2/4}\overline{(\phi_2\,|W(\mathbf{z})\,\phi_1)}\,dz$$
$$= (2\pi)^{-n}\int_{\mathbb{R}^{2n}} e^{-|\mathbf{z}|^2/4}\,(\phi_1\,|W(\mathbf{z})\,\phi_2)\,dz = (\phi_1|P\phi_2)\,,$$

where we used

$$\overline{(\phi_2\,|W(\mathbf{z})\ \phi_1)} = (W(\mathbf{z})\phi_1 \mid \phi_2) = (\phi_1 \mid W(\mathbf{z})^*\phi_2) = (\phi_1 \mid W(-\mathbf{z})\phi_2)\ ,$$

and that the measure dz and $\exp -|\mathbf{z}^2/4|$ are unchanged by the reflection $\mathbf{z} \to -\mathbf{z}$. Notice $P \neq 0$, for otherwise

$$0 = \left(\phi_1 | W(\mathbf{z}')PW(\mathbf{z}')\phi_2\right) = (2\pi)^{-n}\int_{\mathbb{R}^{2n}} e^{-|\mathbf{z}|^2/4}\left(\phi_1\left|W(\mathbf{z}')W(\mathbf{z})W(\mathbf{z}')\ \phi_2\right.\right) dz$$

i.e., by Weyl's relations:

$$0 = (2\pi)^{-n}\int_{\mathbb{R}^{2n}} e^{-|\mathbf{z}|^2/4} e^{i\mathbf{t}^{(\mathbf{z}')}\cdot\mathbf{t}^{(\mathbf{z})} - i\mathbf{u}^{(\mathbf{z}')}\cdot\mathbf{u}^{(\mathbf{z})}}\left(\phi_1\,|W(\mathbf{z})\ \phi_2\right) dz\,.$$

In other terms the Fourier transform of the L^1 function

$$\mathbf{z} \mapsto e^{-|\mathbf{z}|^2/4}\left(\phi_1\,|W(\mathbf{z})\ \phi_2\right)$$

is null. Then by Proposition 3.81(f) $\mathbf{z} \mapsto (\phi_1\,|W(\mathbf{z})\ \phi_2) = 0$ almost everywhere. Since the map is continuous it must vanish everywhere, so $W(\mathbf{z}) = 0$. As earlier mentioned this cannot be. To finish the proof we need to justify:

$$PW(\mathbf{z})P = e^{-|\mathbf{z}|^2/4}P\,. \tag{11.55}$$

Choosing $\mathbf{z} = 0$ in (11.55) gives $PP = P$, making P a non-null orthogonal projector. If $\Phi_0 \in P(\mathsf{H}) \setminus \{0\}$ with $||\Phi_0|| = 1$, as $P\Phi_0 = \Phi_0$, equation (11.55) implies, for any $\mathbf{z} \in \mathsf{X}$

$$(\Phi_0|W(\mathbf{z})\Phi_0) = e^{-|\mathbf{z}|^2/4} = e^{-(|\mathbf{t}^{(\mathbf{z})}|^2+|\mathbf{u}^{(\mathbf{z})}|^2)/4}\,.$$

Hence our Φ_0 satisfies (11.53). Now let us prove (11.55). By definition of P Proposition 9.27(b) and Weyl's relations give

$$(2\pi)^n PW(\mathbf{z})P = \int_{\mathbb{R}^n} dz' e^{-\mathbf{z}'^2/4} PW(\mathbf{z})W(\mathbf{z}') = \int_{\mathbb{R}^{2n}} dz' e^{-\mathbf{z}'^2/4} e^{-i\sigma(\mathbf{z},\mathbf{z}')/2} PW(\mathbf{z}+\mathbf{z}')\,.$$

Recalling (11.54) we can solve for P the integrand. By Proposition 9.27(b):

$$(\phi_1|PW(\mathbf{z})P\phi_2)$$

$$= \frac{1}{(2\pi)^{2n}}\int dz' dz'' e^{-(\mathbf{z}'^2+\mathbf{z}''^2)/4} e^{-i\sigma(\mathbf{z},\mathbf{z}')/2} e^{-i\sigma(\mathbf{z}'',\mathbf{z}+\mathbf{z}')/2}\left(\phi_1|W(\mathbf{z}+\mathbf{z}'+\mathbf{z}'')\phi_2\right) \tag{11.56}$$

for any $\phi_1, \phi_2 \in \mathsf{H}$. We have passed from an iterated integral to an integral in the product measure using Fubini–Tonelli: this is because the integrand vanishes absolutely and exponentially as the product measure's variables go to infinity, due to the exponentials and the estimate $|(\phi_1|W(\mathbf{z}+\mathbf{z}'+\mathbf{z}'')\phi_2)| \leq ||\phi_1||\,||\phi_2||$. Set $\mathbf{z} = (\alpha, \beta)$, $\mathbf{z}' = (\gamma', \delta')$ and $\mathbf{z}'' = (\gamma, \delta)$. The right side of (11.56) reads:

$$\int_{\mathbb{R}^{4n}} \frac{d\gamma d\delta d\gamma' d\delta'}{(2\pi)^{2n}} e^{-(|\gamma|^2-|\delta|^2-|\gamma'|^2-|\delta'|^2)/4} e^{-\frac{i}{2}(\alpha\cdot\delta'-\beta\cdot\gamma'+\gamma\cdot\beta+\gamma\cdot\delta'-\delta\cdot\alpha-\delta\cdot\gamma')}$$

$$\times\left(\phi_1\left|W((\alpha+\gamma+\gamma', \beta+\delta+\delta'))\ \phi_2\right.\right)\,.$$

Changing variables to $\kappa, \nu, \mu, \lambda \in \mathbb{R}^n$, where $\gamma = (\kappa+\mu-\alpha)/2$, $\gamma' = (\kappa-\mu-\alpha)/2$, $\delta = (\nu+\lambda-\beta)/2$, $\delta' = (\nu-\lambda-\beta)/2$, the integral can be computed explicitly, because the integrals in μ, λ decouple to produce Gaussian integrals. The right-hand side of (11.56) equals, eventually:

$$\frac{e^{-(|\alpha|^2+|\beta|^2)/4}}{(2\pi)^n}\int_{\mathbb{R}^{2n}} d\kappa d\nu e^{-(|\kappa|^2+|\nu|^2)/4}\left(\phi_1 \left| W((\kappa,\nu))\, \phi_2\right.\right) = e^{-|\mathbf{z}|^2/4}(\phi_1|P\phi_2)$$

which produces (11.55) since $\phi_1, \phi_2 \in \mathsf{H}$ are free. □

This concludes the proof of Theorem 11.22 (Stone–von Neumann). □

Proof of Theorem 11.23 (Mackey). The hypotheses (a1), (a2), (a3) are equivalent because of Remark 11.27(4). With those assumptions the $W(\mathbf{z})$ are unitary, with $W(0) = I$. By this we can go through the proof of 11.26, which only used that the $W(\mathbf{z})$ were unitary with $W(0) = I$, and did not rely on the representation's irreducibility, and build the orthogonal projector $P \neq 0$:

$$P = (2\pi)^{-n}\int_{\mathbb{R}^{2n}} dz e^{-|\mathbf{z}|^2/4} W(\mathbf{z})\,, \quad \text{for any } \mathbf{z} \in \mathbb{R}^{2n}\,,$$

so that every $\Phi_0 \in P(\mathsf{H})$ satisfies

$$(\Phi_0|W(\mathbf{z})\Phi_0) = e^{-|\mathbf{z}|^2/4} = e^{-(|\mathbf{t}^{(\mathbf{z})}|^2+|\mathbf{u}^{(\mathbf{z})}|^2)/4}\,,$$

as we have seen. First we show the closed $\mathsf{H}_0 := \overline{< \{W(\mathbf{z})P(\mathsf{H})\}_{z\in\mathsf{X}} >}$ coincides with H. H_0 is, by construction, invariant under $W(\mathbf{z})$. Then $\mathsf{H}_0^\perp$ is also invariant. If $\mathsf{H}_0^\perp \neq \{0\}$, working in $\mathsf{H}_0^\perp$ as ambient Hilbert space, using the restrictions $W(\mathbf{z})\restriction_{\mathsf{H}_0^\perp}$ (note $W(0)\restriction_{\mathsf{H}_0^\perp} = I\restriction_{\mathsf{H}_0^\perp} \neq 0$ if $\mathsf{H}_0^\perp \neq \{0\}$), we construct the unique orthogonal projector $P' \neq 0$ such that

$$(\phi_1'|P'\phi_2') = (2\pi)^{-n}\int_{\mathbb{R}^{2n}} e^{-|\mathbf{z}|^2/4}\left(\phi_1' \left| W(\mathbf{z})\, \phi_2'\right.\right) dz\,, \quad \text{for any } \mathbf{z} \in \mathbb{R}^{2n},\ \phi_1, \phi_2 \in \mathsf{H}_0^\perp\,.$$

We know the integral on the right equals $(\phi_1'|P\phi_2')$, i.e. zero, because $\phi_2' \in \mathsf{H}_0^\perp = (P(\mathsf{H}))^\perp$. Hence $P' = 0$, but this contradicts $P' \neq 0$. We conclude $\mathsf{H}_0^\perp = \{0\}$, and so $\mathsf{H}_0 = \mathsf{H}$. Take a basis $\{\Phi_k\}_{k\in I}$ of $P(\mathsf{H})$ and consider the closed spaces $\mathsf{H}_k := \overline{< \{W(\mathbf{z})\Phi_k\}_{\mathbf{z}\in\mathsf{X}} >}$ invariant under $W(\mathbf{z})$. Notice $\Phi_k \in \mathsf{H}_k$, since $W(0) = I$, so $\mathsf{H}_k \neq \{0\}$ for any $k \in I$. By (11.55):

$$(\Phi_j|W(\mathbf{z})\Phi_k) = (\Phi_j|PW(\mathbf{z})P\Phi_k) = e^{-|\mathbf{z}|^2/4}(\Phi_j|P\Phi_k) = 0 \quad \text{if } j \neq k\,.$$

We have found closed subspaces $\mathsf{H}_j \neq \{0\}$ that are mutually orthogonal (in particular j varies in a countable set if H is separable). By construction, as $\overline{< \{W(\mathbf{z})P(\mathsf{H})\}_{z\in\mathsf{X}} >} = \mathsf{H}$ and $\{\Phi_k\}_{k\in I}$ is a basis in $P(\mathsf{H})$, the space of finite combinations of vectors in the mutually orthogonal H_k is dense in H: H is thus a Hilbert sum $\oplus_{k\in I}\mathsf{H}_k$ of the closed H_k, $k \in I$ (Definition 7.34). To finish, on every H_k we can replicate the proof of Stone–von Neumann with H replaced by H_k and $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ replaced by $\pi_k : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_k)$, restriction of the image of each operator in $\pi(\mathfrak{A})$ to H_k. The only difference is that now $\pi_k(\mathfrak{A})\Phi_k$ is dense in H_k by assumption, whereas in the theorem it descends from the irreducibility of $\pi_k(\mathfrak{A})$. Thus the restriction $\pi_k(\mathfrak{A})$ of $\pi(\mathfrak{A})$ to H_k is isomorphic to the standard representation of the Weyl algebra on $L^2(\mathbb{R}^n, dx)$. As the latter is irreducible, so must be every π_k. This ends the proof. □

11.3.6 More on "Heisenberg's principle": weakening the assumptions and extension to mixed states

The formalism developed to prove the Stone–von Neumann theorem allows to generalise Theorem 11.13, i.e. Heisenberg's principle, by taking weaker assumptions on the set to which ψ belongs (the existence of $(\Delta X_i)_\psi$, $(\Delta P_j)_\psi$ suffices). It also enables to extend it to include mixed states. Let us begin with a technical lemma.

Lemma 11.29. *Let X_i, P_j be the position and momentum operators of axiom* **A5**, *and define $X'_i := X_i + a_i I$, $P'_j := P_j + b_j I$, with $a_i, b_j \in \mathbb{R}$. If $\psi, \phi \in D(X_i) \cap D(P_j)$ the canonical commutation relations hold, here written using quadratic forms:*

$$(X'_i\psi|P'_j\varphi) - (P'_j\psi|X'_i\varphi) = i\hbar\delta_{ij}(\psi|\varphi)\,. \tag{11.57}$$

Proof. Notice $D(X_i) = D(X'_i)$, $D(P_j) = D(P'_j)$. In case $a_i, b_j = 0$, consider (11.22), so

$$\begin{aligned}&(W((-\mathbf{t},\mathbf{0}))\psi\,|W((\mathbf{0},\mathbf{u}))\varphi\,) - (W((\mathbf{0},-\mathbf{u}))\psi\,|W((\mathbf{t},\mathbf{0}))\varphi\,)\\ &\quad = (1 - e^{-i(\mathbf{t}\cdot\mathbf{u})/2})\,(W((-\mathbf{t},\mathbf{0}))\psi\,|W((\mathbf{0},\mathbf{u}))\varphi\,)\,.\end{aligned}$$

Using Stone's theorem $(X'_i\psi|P'_j\varphi) - (P'_j\psi|X'_i\varphi) = i\hbar\delta_{ij}(\psi|\varphi)$. Add a_iI and b_jI to the operators inside the scalar products on the left. Since X_i, P_j are Hermitian, the terms on the right cancel out, yielding (11.57) in the general case. □

Theorem 11.30. *Let X_i and P_j be the position and momentum operators of axiom* **A5**. *If the unit vector $\psi \in \mathsf{H}_S$ is such that $(\Delta X_i)_\psi$ and $(\Delta P_i)_\psi$ exist, then Heisenberg's principle holds:*

$$(\Delta X_i)_\psi(\Delta P_i)_\psi \geq \hbar/2\,.$$

Proof. By part (i) in Proposition 11.8(a) if $(\Delta X_i)_\psi$ and $(\Delta P_i)_\psi$ are defined then $\psi \in D(X_i) \cap D(P_i)$. Referring to Lemma 11.29 we choose $a_i = -(\psi|X_i\psi)$, $b_j = -(\psi|P_i\psi)$. By definition of standard deviation (11.3) and Theorem 9.4(f) we have $(\Delta X_i)^2_\psi = \int(\lambda - a_i)^2 d\mu^{(A)}_\psi(\lambda) = ||X'_i\psi||^2$. Similarly, $||P'_i\psi||^2 = (\Delta P_i)^2_\psi$. On the other hand (for any a_i, b_i) from (11.57) we infer:

$$||X'_i\psi||\,||P'_i\psi|| \geq |(X'_i\psi|P'_i\psi)| \geq |Im(X'_i\psi|P'_i\psi)| = \frac{\hbar}{2}\,. \tag{11.58}$$

Since $(\Delta X_i)_\psi(\Delta P_i)_\psi = ||X'_i\psi||\,||P'_i\psi||$, the claim is proved. □

So now we can extend "Heisenberg's principle" to mixed states as well.

Theorem 11.31. *Let X_i and P_j be the position and momentum operators of axiom* **A5**. *If ρ is a mixed state for the spin-zero particle such that $(\Delta X_i)_\rho$ and $(\Delta P_i)_\rho$ exist, then:*

$$(\Delta X_i)_\rho (\Delta P_i)_\rho \geq \frac{\hbar}{2}\,.$$

Proof. Let us notice, preliminarily, that if $(\Delta X_i)_\rho$ and $(\Delta P_i)_\rho$ can be defined, then also $\langle (X_i)^k \rangle_\rho$ and $\langle (P_i)^k \rangle_\rho$, $k = 0,1,2$, are defined, as is easy to see using Definition 11.6, because measures are finite. Furthermore, $Ran(\rho) \subset D(X_i) \cap D(P_i)$ by (ii) in Proposition 11.8(b), as $Ran(\rho^{1/2}) \supset Ran(\rho)$. Set $X_i' := X_i + a_i I$, $P_i' := P_i + b_i I$, and choose $a_i := -\langle X_i \rangle_\rho$, $b_i := -\langle P_i \rangle_\rho$. A direct computation relying on Definition 11.6 tells that $(\Delta X_i)^2_\rho = \langle (X_i')^2 \rangle_\rho$ and $(\Delta P_i)^2_\rho = \langle (P_i')^2 \rangle_\rho$. Write $\rho = \sum_n p_n \psi_n (\psi_n|\;)$ in a basis of unit eigenvectors. We argue as in Proposition 11.8 when we proved (11.11). As $A = X_i', P_i'$ and $f(\lambda) = \lambda$, and since $\mu^{(A)}_\rho(E) = tr(P^{(A)}(E)\rho) = \sum_n p_n \mu_{\psi_n}(E)$, using that $p_n \geq 0$ we can prove:

$$\int |f(\lambda)|^2 d\mu^{(A)}_\rho(\lambda) = \sum_{n=0}^{+\infty} p_n \int |f(\lambda)|^2 d\mu^{(A)}_{\psi_n}(\lambda) = \sum_{n=0}^{+\infty} p_n (f(A)\psi_n | f(A)\psi_n) \leq +\infty\,,$$

where $\psi_n \in D(X_i') \cap D(P_i') = D(X_i) \cap D(P_i)$, because $\psi_n \in Ran(\rho) \subset D(X_i) \cap D(P_i)$. Therefore:

$$(\Delta X_i)^2_\rho = \langle (X_i')^2 \rangle_\rho = \sum_n p_n (X_i' \psi_n | X_i' \psi_n)$$

and

$$(\Delta P_i)^2_\rho = \langle (P_i')^2 \rangle_\rho = \sum_m p_m (P_i' \psi_m | P_i' \psi_m)\,.$$

Schwarz's inequality plus (11.58) imply the claim, because

$$\langle (X_i')^2 \rangle^{1/2}_\rho \langle (P_i')^2 \rangle^{1/2}_\rho \geq \sum_n p_n^{1/2} p_n^{1/2} (X_i' \psi_n | X_i' \psi_n)^{1/2} (P_i' \psi_n | P_i' \psi_n)^{1/2} \geq \sum_n p_n \frac{\hbar}{2} = \frac{\hbar}{2}$$

(note $p_n \geq 0$ and $\sum_n p_n = 1$). □

11.3.7 The Stone–von Neumann theorem revisited, via the Heisenberg group

Our approach to the proof of Stone–von Neumann relies on the structure of (Weyl) *-algebra. There is, however, another point of view, due to Weyl, in which the **(Weyl-)Heisenberg group** plays the algebra's role. The Heisenberg group in $\mathbb{R}^{2n+1}$, which we indicate by $\mathscr{H}(n)$, is the simply connected Lie group diffeomorphic to $\mathbb{R}^{2n+1}$ with product law

$$(\eta, \mathbf{t}, \mathbf{u}) \circ (\eta', \mathbf{t}', \mathbf{u}') = \left(\eta + \eta' + \frac{1}{2} \sum_{i=1}^{n} u_i t_i' - u_i' t_i\,,\, \mathbf{t} + \mathbf{t}'\,,\, \mathbf{u} + \mathbf{u}' \right)$$

(as usual $\mathbf{t} \in \mathbb{R}^n$, $\mathbf{u} \in \mathbb{R}^n$ whilst $\eta \in \mathbb{R}$). A direct computation of its Lie algebra shows there is a basis of $2n+1$ generators $\mathbf{x}_i$, $\mathbf{p}_i$, $\mathbf{e}$, $i = 1, 2, \ldots, n$ that satisfy:

$$[\mathbf{x}_i, \mathbf{p}_j] = \delta_{ij}\mathbf{e}\,, \quad [\mathbf{x}_i, \mathbf{e}] = [\mathbf{p}_i, \mathbf{e}] = 0\,, \quad i, j = 1, 2, \ldots, n\,.$$

The linear mapping determined by $\mathbf{e} \mapsto -iI$, $\mathbf{x}_k \mapsto -iX_k$, $\mathbf{p}_k \mapsto -iP_k$ is an isomorphism from the Heisenberg Lie algebra to the Lie algebra of finite real combinations of the conjugate self-adjoint operators $-iI, -iX_k, -iP_k$, restricted to the common, dense and invariant domain $\mathscr{S}(\mathbb{R}^n)$, with commutator $[\cdot,\cdot]$ as Lie bracket. This map induces a Lie group isomorphism. By direct inspection, in fact, if the operators $W((\mathbf{t},\mathbf{u}))$ are defined by Proposition 11.18, the map

$$\mathbb{R}^{2n+1} \ni (\eta, \mathbf{t}, \mathbf{u}) \mapsto e^{i\eta} W((\mathbf{t},\mathbf{u})) =: H((\eta, \mathbf{t}, \mathbf{u})) \tag{11.59}$$

is a faithful (one-to-one) and irreducible unitary representation of the $2n+1$-dimensional Heisenberg group on $L^2(\mathbb{R}^n, dx)$. Moreover,

$$\text{s-}\lim_{s\to 0} H(s(\eta, \mathbf{t}, \mathbf{u})) = I \quad \text{for any given } (\eta, \mathbf{t}, \mathbf{u}) \in \mathbb{R}^{2n+1}\,. \tag{11.60}$$

Conversely,

Proposition 11.32. *An irreducible unitary representation of the Heisenberg group $\mathscr{H}(n) \ni (\eta, \mathbf{t}, \mathbf{u}) \mapsto H((\eta, \mathbf{t}, \mathbf{u}))$ on the Hilbert space H, such that (11.60) holds, has the form (11.59):*

$$H((\eta, \mathbf{t}, \mathbf{u})) = e^{i\eta} W((\mathbf{t},\mathbf{u}))\,,$$

where the $W((\mathbf{t},\mathbf{u}))$ satisfy the Stone–von Neumann Theorem 11.22.

Proof. Equation (11.59) holds because the centre $\mathbb{R}$ of the Heisenberg group is represented by a unitary Abelian subgroup. As the elements of $\mathbb{R}$ commute with the Weyl group every element $H((\eta, \mathbf{0}, \mathbf{0}))$ commutes with the whole representation. But the latter is irreducible, so Schur's lemma forces $H((\eta, \mathbf{0}, \mathbf{0})) = \chi(\eta)I$, with $\chi(\eta) \in \mathbb{C}$, and $|\chi(\eta)| = 1$ as $H((\eta, \mathbf{0}, \mathbf{0}))$ is unitary. Eventually, since $\eta \to H((\eta, \mathbf{0}, \mathbf{0}))$ is strongly continuous, Stone's theorem implies $\chi(\eta) = e^{ic\eta}$ for every $\eta \in \mathbb{R}$ and some constant c. The group's commutation rules require $c = 1$, but also make the $W((\mathbf{t},\mathbf{u}))$ obey Weyl's relations. □

In this framework we have an alternative statement of Stone–von Neumann, first proved by Weyl.

Theorem 11.33. *Every irreducible unitary representation of the Heisenberg group $\mathscr{H}(n)$ satisfying (11.60) is unitarily equivalent to the representation on $L^2(\mathbb{R}^n, dx)$:*

$$\mathbb{R}^{2n+1} \ni (\eta, \mathbf{t}, \mathbf{u}) \mapsto e^{i\eta} W((\mathbf{t},\mathbf{u}))\,,$$

where the $W((\mathbf{t},\mathbf{u}))$ are operators of Proposition 11.18.

Proof. Immediate consequence of Proposition 11.32, in this setup. □

Remark 11.34. (1) The Stone–von Neumann theorem proves that the non-relativistic elementary particle with spin 0 is described by an *irreducible* representation of a certain Lie group. The same happens for particles with spin, charge, ..., provided one picks the right group. *Elementary systems* are thus described by *irreducible representations* of a group, typically related to the symmetries of the physical system. This point of view has proved – thanks to Wigner in particular – incredibly rewarding for the development of relativistic quantum theories, where irreducible representations of the Poincaré group are employed to define elementary particles, and irreducibility is a characteristic feature of elementary systems.
(2) There exist more or less rigorous formulations of the Stone–von Neumann theorem that rely only on Heisenberg's relations (11.15) and do not need exponentials. To set up these formulations, though, the technical assumptions on domains (spaces of analytic vectors) and on the existence of self-adjoint extensions are not al all obvious, nor they have a straightforward physical meaning. Beside the foundational work of E. Nelson [Nel59], an important and thorough result is that of J. Dixmier [Dix56], which we shall return to in the next chapter. In a nutshell the theorem, generalised to arbitrary finite dimension, states that if P, Q are symmetric on a dense invariant space on which Heisenberg's relations hold, and there $P^2 + Q^2$ is essentially self-adjoint, then P, Q give a strongly continuous representation of the Weyl algebra on the Hilbert space; hence, up to isomorphisms, P, Q have the usual form on $L^2(\mathbb{R}, dx)$. ■

11.3.8 Dirac's correspondence principle and Weyl's calculus

The formulation of QM we have presented leaves open the question of how to pick out operators on H that correspond to observables of physical interest, other than position and momentum. Several important authors have written much about procedures allowing to pass from relevant classical observables to major quantum observables. But that is somewhat like fighting a losing battle: from a physical perspective Quantum Mechanics is 'more central' than Classical Mechanics, whence the latter should be seen as a limiting case of the former. Even this fact is by no means easy to prove, apart in a few general cases: one such is *Eherenfest's theorem*, whose precise formulation was found only recently [FK09]. Thus one expects there should be quantum entities, observables in particular, without classical counterparts (for instance the "parity" of elementary particles, and in many respects also spin).

That said, certain quantum observables for the spin-zero particle will, in principle, be "functions" of the observables X_i, P_i. The common belief is that the quantum quantity corresponding to the classical $F(x, p)$ should look something like $F(X, P)$. But going down this road is a real challenge, more than what mathematics prospects. In fact: (1) is it not at all obvious what meaning one should assign to a function of X and P when these operators have non-commuting spectral measures (in the commuting case there are ways out that use *joint measures*, like (11.1); (2) naïve recipes in this direction do not produce self-adjoint, not even symmetric, operators when the operators do not commute.

For the sake of clarity consider the classical quantity $x \cdot p$. Which observable – i.e. *self-adjoint* operator – should it correspond to? Passing to the spectral measures is ill-advised, because they do not commute. So let us try to use the operators themselves, restricted to an invariant and dense subspace where they are both defined. The hope is to produce an essentially self-adjoint operator, or at least symmetric, and then in some way or another choose among its self-adjoint extensions (if any at all, in case the operator is symmetric). The tentative answer:

$$x \cdot p \text{ corresponds to } X \cdot P (= \sum_{i=1}^{n} X_i P_i)$$

is totally inadequate, even if we view the operators on the invariant dense space $\mathscr{S}(\mathbb{R}^3)$. That is because $X \cdot P$ is not symmetric on $\mathscr{S}(\mathbb{R}^3)$, for X_i and P_i do not commute (exercise). Neither would it make sense to seek self-adjoint extensions of $X \cdot P$. Another possibility is to associate to $x \cdot p$ the symmetric operator $(X \cdot P + P \cdot X)/2$ defined on $\mathscr{S}(\mathbb{R}^3)$, and study its self-adjoint extensions. When examining more complicated situations, like $x_k^2 p_k$, this recipe reveals itself very ambiguous, because *a priori* there are several possibilities: $(X_k^2 P_k + P_k X_k^2)/2$ is symmetric on the domain $\mathscr{S}(\mathbb{R}^3)$, but also $X_k(X_k P_k + P_k X_k)/4 + (X_k P_k + P_k X_k)X_k/4$ is, and there are others. These choices correspond to “symmetrised” products, of sorts, of (non-commuting) operators, that should produce an operator that is at least symmetric.

A criterion, helpful but not decisive to solve the issues raised, was found by Dirac, and goes under the accepted name of “Dirac’s correspondence principle”. To present it, let us recall that a **Lie algebra** $(\mathsf{V}, [\cdot,\cdot])$ is a vector space (here, over $\mathbb{R}$) equipped with a skew-symmetric bilinear map $[\cdot,\cdot] : \mathsf{V} \times \mathsf{V} \to \mathsf{V}$, called **Lie bracket**, that satisfies the Jacobi identity:

$$[u,[v,w]] + [v,[w,u]] + [w,[u,v]] = 0\,, \quad \text{for any } u,v,w \in \mathsf{V}.$$

In studying the phase space $\mathscr{F}$ of the classical particle (although the setup is fully general), Dirac considered the real vector space $\mathscr{G}(\mathscr{F})$ of sufficiently regular maps from $\mathscr{F}$ to $\mathbb{R}$ with *Poisson bracket*:

$$\{f,g\} := \sum_i \frac{\partial f}{\partial x_i}\frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial x_i}\frac{\partial f}{\partial p_i}\,, \quad f,g \in \mathscr{G}(\mathscr{F})\,.$$

He noticed $(\mathscr{G}(\mathscr{F}), \{\cdot,\cdot\})$ is a *Lie algebra*. In particular the CCRs hold:

$$\{x_i, p_j\} = \delta_{ij}\,.$$

These equations are Heisenberg’s relations when we substitute $x_i \to X_i$, $p_i \to P_i$ and $\{\cdot,\cdot\} \to -i\hbar^{-1}[\cdot,\cdot]$. The idea behind “Dirac’s correspondence principle” is the following.

Let $\hat{f}$ denote the quantum analogue (an operator at least symmetric, and defined on a dense invariant domain, independently of the specific quantity) of the generic classical quantity $f \in \mathscr{G}(\mathscr{F})$. Under Dirac’s correspondence if

$$h = \{f,g\}$$

for classical $f,g,h \in \mathscr{G}(\mathscr{F})$, in the quantum realm the corresponding $\hat{f},\hat{g},\hat{h}$ satisfy

$$\hat{h} = -i\hbar^{-1}[\hat{f},\hat{g}]\,.$$

Just as example, consider the usual classical particle. The components of the classical angular momentum

$$l_i = \sum_{j,k=1}^{3} \varepsilon_{ijk} x_j p_k$$

correspond to

$$L_i = \sum_{j,k=1}^{3} \varepsilon_{ijk} X_j P_k \,,$$

which are essentially self-adjoint operators on $\mathscr{S}(\mathbb{R}^3)$. The classical commutation relations

$$\{l_i, l_j\} = \sum_{k=1}^{3} \varepsilon_{ijk} l_k$$

have analogue quantum counterparts on $\mathscr{S}(\mathbb{R}^3)$:

$$[L_i, L_j] = i\hbar \sum_{k=1}^{3} \varepsilon_{ijk} L_k \,.$$

Dirac's principle could be explained for observables corresponding to the generators of unitary transformations in a symmetry group of the system; these, though, do not exhaust all possible observables. (All this might be more enlightening after reading the book's final three chapters.) In that case it is only natural to request that (a) the Lie algebra of the symmetry group, (b) the Lie algebra of the unitary representation of transformations on the quantum system, and (c) the Lie algebra of generators of the group of classical canonical transfomations that correspond to symmetries of the classical system, are all isomorphic.

Although we will not push the study any further, we cannot not mention that serious technical hurdles crop up in trying to interpret Dirac's idea literally. Suppose, in particular, of working with polynomial functions of arbitrarily large degree in the canonical variables x_i, p_j. Then [Stre07] it is not possible to define a "symmetrised product" of self-adjoint operators corresponding to canonical variables (so to produce operators that are at least symmetric) that does not depend on the degree and that yields the isomorphism $f \mapsto \hat{f}$.

At the same time we have to emphasise that some ideas underlying Dirac's correspondence principle have found a rigorous treatment within certain quantisation procedures called *Weyl Quantisation* or *Weyl calculus* (in particular see [Jef04], [ZFC05], [Gra04] and [DA10]). The following formula, proved by Weyl and based on the Fourier transform, tells how to associate to a function $f = f(x_1, \dots, x_n, p_1, \dots, p_n)$ the operator

$$f(X_1, \dots, X_n, P_1, \dots, P_n)$$
$$:= \frac{1}{(2\pi)^{2n}} \int_{\mathbb{R}^{2n}} \exp\left\{ -i \overline{\sum_{k=1}^{n} t_k \mathscr{X}_k + u_k \mathscr{P}_k} \right\} \widetilde{f}(t_1, \dots, t_n, u_1, \dots, u_n) dt du$$

that is function of the operators $X_1, \ldots, X_n, P_1, \ldots, P_n$. Above, $\widetilde{f}$ is the Fourier transform of f, and the integral is meant in the sense of Proposition 9.27, assuming f is suitable, e.g. a Schwartz function on $\mathbb{R}^{2n}$. Using duality theorems the definition extends to Schwartz distributions f, for which also polynomial functions can be considered (see Chapter 2 of [Jef04] for a brief and precise technical account of Weyl calculus, and Chapter 1 for other, related procedures). Weyl's procedure does provide operators that can be viewed as functions of the non-commuting X_k, P_k, but has problems. First of all, it maps real functions to self-adjoint operators, but it does not preserve positivity (positive functions are not sent, in general, to positive operators). Weyl's procedure, furthermore, maps the Poisson bracket of two polynomials to the commutator of the corresponding functions of operators only if the polynomials are at most quadratic.

In conclusion, despite interesting and remarkable technical attempts the broad validity of Dirac's correspondence principle remains dubious; its many snags are critical and apparently inescapable, and are borne by the endeavour to provide a serious framework to Dirac's original idea, even in its most rigorous versions such as Weyl's calculus.

Exercises

11.1. Let $\mathfrak{R}$ be the von Neumann algebra (of bounded observables on a physical system) on the Hilbert space H and $\mathfrak{P}_{\mathfrak{R}} := \mathfrak{R} \cap \mathfrak{P}(\mathsf{H})$. Prove $\mathfrak{R} = \mathfrak{P}''_{\mathfrak{R}}$.

Hint. If $T \in \mathfrak{R}$, $T = A + iB$ with A, B self-adjoint and bounded. Show the spectral measures of A and B belong to $\mathfrak{R}$, and consequently $T \in \mathfrak{P}''_{\mathfrak{R}}$.

11.2. Let $\mathfrak{R}$ be the von Neumann algebra (of bounded observables on a physical system) with non-trivial centre $\mathfrak{Z} := \mathfrak{R} \cap \mathfrak{R}'$. Prove $\mathfrak{Z}$ contains non-trivial orthogonal projectors (that can describe superselection rules).

Hint. If $A \in \mathfrak{Z}$, $\frac{i}{2}(A + A^*)$ and $\frac{i}{2}(A - A^*)$ belong to $\mathfrak{Z}$. Consider the spectral measures of these self-adjoint operators.

11.3. Consider a particle moving on the real line, and suppose the pure state represented by the differentiable function $\psi \in D(X^2) \cap D(P^2) \cap D(XP) \cap D(PX)$, with $||\psi|| = 1$, satisfies $(\Delta X)_\psi (\Delta P)_\psi = \hbar/2$. Prove

$$\psi(x) = (\pi\hbar\gamma)^{-1/4} e^{i\frac{\langle P\rangle_\psi x}{\hbar}} e^{-\frac{(x-\langle X\rangle_\psi)^2}{2\hbar\gamma}}$$

for some $\gamma > 0$.

Hint. Referring to the proof of Theorem 11.13, note that we can have $(\Delta X)_\psi (\Delta P)_\psi = \hbar/2$ only if $||X'\psi||\,||P'\psi|| = |(X'\psi|P'\psi)|$, plus $Re(X'\psi|P'\psi) = 0$. The first condition implies, by Proposition 3.3(i), that $X'\psi = cP'\psi$ for some $c \in \mathbb{C}$. Since $\sigma_p(X) = \varnothing$ and $\psi \neq 0$, the second condition implies $Re(c) = 0$. Solving the differential equation $X'\psi = iIm(c)P'\psi$, and using $||\psi|| = 1$, leads to the required expression for ψ.

11.4. Prove that a symplectic linear map $f : (\mathsf{X}, \sigma) \to (\mathsf{X}', \sigma')$ is one-to-one.

Hint. Remember that symplectic forms are weakly non-degenerate, and if $f(x) = 0$ then $\sigma(y,x) = \sigma'(f(y), 0) = 0$ for any $y \in \mathsf{X}$.

11.5. Consider the Hilbert space $\mathsf{H} := L^2([a,b], dx)$ and the self-adjoint operator X on H defined by $(X\psi)(x) := x\psi(x)$, for any $\psi \in \mathsf{H}$ such that $X\psi \in \mathsf{H}$. Prove there is no self-adjoint extension P of the symmetric operator $-i\frac{d}{dx}$, defined on the subspace of C^1 maps either vanishing, or periodic, at the boundary of $[a,b]$, so that the one-parameter unitary groups $U(u) := e^{iuX}$, $V(v) := e^{ivP}$ satisfy Weyl's relations: $U(u)V(v) = V(v)U(u)e^{iuv}$ for any $u, v \in \mathbb{R}$.

Hint. First note that, trivially, $V(sv), U(su) \to I$ in strong sense, as $s \to 0$, because one-parameter unitary groups generated by self-adjoint operators are strongly continuous. There are various ways to solve the exercise. For example we can prove $\sigma(X) = [a,b]$. This is impossible if P as above exists, because by Theorem 11.24 there should be a unitary operator S mapping X and P into the operators on $L^2(\mathbb{R}, dx)$ of axiom **A5** (passing from $\mathbb{R}^3$ to $\mathbb{R}^1$ in the obvious way); another possibility is to split $L^2([a,b], dx)$ in a Hilbert sum of closed X- and P-invariant spaces, on each of which there is the aforementiond unitary S. In either case we can prove $\sigma(X) = \sigma(SXS^{-1}) = \mathbb{R} \neq [a,b]$.

11.6. Refer to the proof of Proposition 11.18 and adapt the definitions of A, $A^\star$ by considering $L^2(\mathbb{R}, dx)$ with Hermite functions $\{\psi_n\}_{n\in\mathbb{N}}$ as basis, and the Bargmann-Hilbert space B_1 (see Example 3.32(6)) with entire functions $\{u_n\}_{n\in\mathbb{N}}$ as basis:

$$u_n(z) := \frac{z^n}{\sqrt{n!}} \quad \text{for any } z \in \mathbb{C}.$$

Call **Segal–Bargmann transformation** the unitary operator

$$U : L^2(\mathbb{R}, dx) \to \mathsf{B}_1$$

determined by $U\psi_n := u_n$, $n = 0, 1, 2, \ldots$. Prove

$$UA^\star U^* = z \quad \text{and} \quad UAU^* = \frac{d}{dz} \tag{11.61}$$

over the dense spans of finite combinations of elements of the two bases.

11.7. On the Bargmann-Hilbert space B_1 (see Example 3.32(6)), consider

$$K_0 := z\frac{d}{dz},$$

defined on $D(H_0) = \{f \in \mathsf{B}_1 \,|\, zdf/dz \in \mathsf{B}_1\}$. Prove that it is essentially self-adjoint and find its spectrum. Does $2\overline{K_0} + I$ have any physical meaning?

Hint. Prove it is symmetric, and show $\{u_n\}_{n\in\mathbb{N}}$ is an eigenvector basis of H_0 (hence of analytic vectors). Up to a factor, $2\overline{K_0} + I$ is the Hamiltonian of the harmonic oscillator.

This page intentionally left blank

12

Introduction to Quantum Symmetries

Mathematical sciences, in particular, display order, symmetry and clear limits: and these are the uppermost instances of beauty.

Aristotle

This chapter continues in the description of the mathematical structure of Quantum Mechanics, by introducing fundamental notions and tools of great relevance.

Section one is devoted to defining and characterising *quantum symmetries*. We will present examples, and discuss what happens in presence of superselection rules, define *Kadison symmetries* and *Wigner symmetries*. We shall then prove the *theorems of Wigner and Kadison*, which show that the previous notions actually coincide, and manifest themselves via unitary or antiunitary operators.

In section two we pass to the problem of representing symmetry groups, by introducing *projective representations*, *projective unitary representations* and $U(1)$ *central extensions of a (symmetry) group*.

A part will be dedicated to *topological groups* and the study of strongly continuous projective unitary representations. We will examine the special case of the Abelian group $\mathbb{R}$, that has important applications in QM. Next, after recalling the basics on *Lie* groups and algebras, we will discuss key results due to Bargmann, Gårding and Nelson (and a few generalisations thereof) about projective unitary and unitary representations of Lie groups. We will consider the Peter–Weyl theorem on strongly continuous unitary representations of compact Lie groups (or better, compact Hausdorff topological groups).

As an example of primary importance in physics, we will see in the third section unitary representations of the symmery group $SO(3)$ in relationship to the *spin*.

Eventually we will apply the machinery to the *Galilean group*, and prove *Bargmann's superselection rule of the mass*.

12.1 Definition and characterisation of quantum symmetries

A truly crucial notion on QM, also in view of the subsequent development in quantum field theories, is that of *symmetry* of a quantum system. There are two notions of symmetry: one is dynamic, while the other, more elementary one does not involve temporal evolution, i.e. the dynamics of the physical system. In this first section we will deal with the first kind, and tackle the static type in the following chapter.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_12, © Springer-Verlag Italia 2013

Consider a physical system S described on the Hilbert space H_S, with $\mathfrak{S}(\mathsf{H}_S)$ denoting the space of states and $\mathfrak{S}_p(\mathsf{H}_S)$ that of pure states. When we act by a transformation g on S we alter its quantum state. To the physical transformation g corresponds a map $\gamma_g : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$ of the space of states, or $\gamma_g : \mathfrak{S}_p(\mathsf{H}_S) \to \mathfrak{S}_p(\mathsf{H}_S)$ if we restrict to pure states. The relationship between g and γ_g is not relevant at present, and we will take it for granted; at any rate, it will depend upon the description of S. If γ_g obeys certain conditions, γ_g is called a *symmetry* of the system. Abusing the terminology we will often say g is a symmetry of S. Two are the requisites for γ_g to be a symmetry:

(a) γ_g must be bijective.
(b) γ_g should preserve some mathematical structure of the space of states $\mathfrak{S}(\mathsf{H}_S)$ or of pure states $\mathfrak{S}_p(\mathsf{H}_S)$; for the moment we will not specify which structure, although this will have a precise physical interpretation.

In physics, requisite (a) can actually be forced upon the transformation g acting on the system, and corresponds to asking g be *reversible*, in other words (i) there must exist an inverse transformation g^{-1}, associated to $\gamma_g^{-1} : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$, taking back to the original state, and (ii) any quantum state should be reachable via γ_g, by choosing the initial state suitably.

The differences between the several symmetry notions known depend on the interpretation of condition (b), i.e. on the γ_g-invariant structure. There are at least three possibile choices (see [Sim76]). The simplest structure the map can preserve is the convexity of the space of states, physically corresponding to the fact that a state arises from mixing states with certain statistical weights. Symmetry operations modify the constituent states, but do not change the weights. This sort of quantum symmetries were studied by Kadison [Kad51], and are nowadays called "Kadison symmetries". Another type, due to Wigner [Wig59], refers to functions on $\mathfrak{S}_p(\mathsf{H}_S)$. For these one requires that the metric structure of the projective space of rays be preserved. We will call such "Wigner symmetries". In the language of physics Wigner symmetries modify pure states but do not change the transition probabilities of pure pairs. A third type, which we shall not discuss, was discovered by Segal, and concerns the *Jordan algebra* structure of observables (see [Sim76]). In the sequel we will study the first two classes; we will prove that mathematically they reduce to the same, and that they are described by unitary or antiunitary operators (hence Wigner symmetries may be extended to Kadison symmetries defined on the entire space of states). This characterisation in terms of (anti)unitary operators is hugely important in physics, and is formulated in two results known as *Kadison's theorem* and *Wigner's theorem*. The latter is much more renowned in the physical community, despite the former is equally important.

Remarks 12.1. The notion of quantum symmetry could also be defined referring to the *observables* of a quantum system S rather than to the *states*. Although we will not follow this path, we should mention that a quantum symmetry (associated to the transformation g) may also be defined by an isomorphism of bounded, orthocomplemented σ-complete lattices $\alpha_g : \mathfrak{P}(\mathsf{H}_S) \to \mathfrak{P}(\mathsf{H}_S)$, see Definition 7.11. It is easy to prove that this implies:

$$\alpha_g\left(s\text{-}\sum_{i=1}^{+\infty}P_i\right)=s\text{-}\sum_{i=1}^{+\infty}\alpha_g(P_i)\quad\text{for } \{P_i\}_{i\in\mathbb{N}}\subset\mathfrak{P}(\mathsf{H}_S)\text{ with } P_iP_j=0 \text{ if } i\neq j\,,$$

where $\alpha_g(P_i)\alpha_g(P_j)=0$ if $i\neq j$. By the above identities, on quantum systems obeying Gleason's Theorem 7.24 (systems associated to separable complex Hilbert spaces of dimension ≥ 3) each symmetry, viewed as acting on observables, induces a corresponding symmetry acting on states by a duality process. In fact, if $\mu:\mathfrak{P}(\mathsf{H})\to[0,1]$ is a quantum state in the sense of axiom **A2** (preliminary form) and $\alpha_g:\mathfrak{P}(\mathsf{H}_S)\to\mathfrak{P}(\mathsf{H}_S)$ satisfies the previous identity, $\mu\circ\alpha:\mathfrak{P}(\mathsf{H})\to[0,1]$ is still a state for axiom **A2** (preliminary). If the Hilbert space has dimension ≥ 3, for every mixed state $\rho\in\mathfrak{S}(\mathsf{H}_S)$ there exists a unique mixed state $\alpha_g^*(\rho)$ determined by $tr\left(\alpha_g^*(\rho)P\right)=tr(\rho\,\alpha_g(P))$ for any $P\in\mathfrak{P}(\mathsf{H}_S)$. Then $\alpha_g^*:\mathfrak{S}(\mathsf{H}_S)\to\mathfrak{S}(\mathsf{H}_S)$ is immediately bijective (α_g is) and maps convex combinations of states to convex combinations, preserving statistical weights. And, the mapping $\alpha_g\mapsto\alpha_g^*$ is injective. Put differently, every symmetry α_g of observables corresponds one-to-one to a symmetry α_g^* of states, typically mixed. A consequence of Kadison's theorem, proved in the sequel, is that $\alpha_g\mapsto\alpha_g^*$, viewed from the set of symmetries of observables to the set of "Kadison symmetries", is surjective as well.

From now on we will think of symmetries as transformations of states, and we will discuss the action on observables only after the theorems of Wigner and Kadison.

■

12.1.1 Examples

Before going into mathematical subtleties, let us describe a few examples of physical operations that are (both Wigner and Kadison) symmetries for quantum systems.

Suppose we take an isolated physical system S in a certain inertial frame system $\mathscr{I}$. A transformation that is known to generate a symmetry of S is any rigid translation of S along a given vector, or the rotation about a fixed real axis. In other terms any continuous isometry of the rest space of inertial frames produces a quantum symmetry. Another instance is the change of inertial reference system (in relativistic theories as well), in the following sense: the isolated system S in the inertial frame $\mathscr{I}$ is transformed so that the final system appears, in a different inertial system $\mathscr{I}'\neq\mathscr{I}$, as it appeared at the beginning in $\mathscr{I}$. A third transformation giving symmetries, for isolated systems in inertial frames, is time translation (not to be confused with time evolution) which we will see later.

All these transformations are *active*, meaning that they *change* the system S, or better, its quantum state.

It must be absolutely clear that the transformations we are talking about do not occur because the system's state evolves in time: they are idealised transformations, pure mathematical notions. By the way, some of them could never occur in reality in a system that evolves under its own dynamics, and some could hardly exist. A classical example is the *inversion of parity*. This physical transformation, loosely speaking, substitutes a system S with its mirror image. Sometimes the only way to realise parity inversion, ideally, is to destroy the system and rebuild its symmetrical image from scratch. And sometimes this abstract operation, too, is physically hollow owing to the nature itself of physical laws. Particles that interact under the weak force, surprisingly, constitute systems whose states do not admit parity transformation as a symmetry, in a rather radical sense: the space of states has no transformation γ representing the ideal physical transformation of parity inversion. This simply means that the alleged symmetry is *not* a true symmetry of the system.

Another type of transformation that shares some features with parity inversion, and that is at times associated to symmetries, is *time reversal*. The examples seen so far have to do with spacetime isometries. Albeit active on states, they are related to *passive* transformations of frame systems (or just of coordinates) by means of passive isometries of spacetime. In this case one expects (not always true, as we saw) active transformations on states to be symmetries, precisely because the various frame or coordinate systems – relative to passive (Galilean or Poincaré) transformations used to describe reality (at least macroscopically) – are equivalent. In other words: if we act on the physical system S by an active transformation, we can always revoke the outcome by changing reference system (or just coordinates), knowing the new framing is physically equivalent to the original one.

In contrast to all this, there exist transformations related to symmetries which are neither associated to spacetime isometries, nor reversed by changing frames. A standard example is charge conjugation, which flips the sign of all charges (of the type considered) present in S, and thus changes the superselection sector of the charge. There are, eventually, even more abstract transformations relative to internal symmetries and gauge symmetries, on which we will not spend any time.

In conclusion we wish to underline an important physical fact. The lesson weak interactions teach us is this: deciding whether a transformation acting ideally on a system is indeed a quantum symmetry, is ultimately to be decided – after (b) has been specified – experimentally.

After proving the theorems of Kadison and Wigner we will describe symmetries in terms of (anti)unitary operators, in the case physical transformations form an abstract, topological or Lie *group* [War75, NaSt82].

In the next chapter we shall treat dynamical symmetries, which emerge when one defines the *time evolution* of the quantum state of a system S. In that context we will recover the tight link between dynamical symmetries and associated conservation laws. It is well known, in the classical setup, that this relationship is encoded into the various formulations of the celebrated *Nöther's theorem*.

12.1.2 Symmetries in presence of superselection rules

As was observed already in Chapter 7, if M is a closed subspace in the Hilbert space H we can identify $\mathfrak{S}(\mathsf{M})$ (or $\mathfrak{S}_p(\mathsf{M})$) with a subset of $\mathfrak{S}(\mathsf{H})$ (resp. $\mathfrak{S}_p(\mathsf{H})$) in a natural manner, i.e. by viewing $\mathfrak{S}(\mathsf{M})$ ($\mathfrak{S}_p(\mathsf{M})$) as the collection of states $\rho \in \mathfrak{S}(\mathsf{H})$ ($\mathfrak{S}_p(\mathsf{H})$) such that $Ran(\rho) \subset \mathsf{M}$. This is the same as extending each $\rho \in \mathfrak{S}(\mathsf{M})$ to an operator on

H by declaring it zero on $\mathsf{M}^\perp$. In the remaining part of the chapter we will implicitly make this identification, which is useful in the next situation.

In certain circumstances the possible state of a physical system is not any element in $\mathfrak{S}(\mathsf{H}_S)$ ($\mathfrak{S}_p(\mathsf{H}_S)$ if pure), because some convex combinations are forbidden. This is the case when we have *superselection rules* (see Chapter 7.4.5 and 11.1). Without repeating what we explained earlier, let us only say that in presence of superselection rules H_S splits into a direct sum of closed orthogonal subspaces, at most countably many, called *coherent sectors*:

$$\mathsf{H}_S = \bigoplus_{k\in K} \mathsf{H}_{Sk} \,.$$

Then we can define the spaces of states and pure states of each sector, $\mathfrak{S}(\mathsf{H}_{Sk})$ and $\mathfrak{S}_p(\mathsf{H}_{Sk})$. Note $\mathfrak{S}(\mathsf{H}_{Sk}) \cap \mathfrak{S}(H_{Sj}) = \varnothing$ and $\mathfrak{S}_p(\mathsf{H}_{Sk}) \cap \mathfrak{S}_p(\mathsf{H}_{Sj}) = \varnothing$ if $k \neq j$. Concerning physically-admissible pure states for the superselection rule, these will be precisely the constituents of the disjoint union

$$\bigsqcup_{k\in K} \mathfrak{S}_p(\mathsf{H}_{Sk}) \,.$$

Admissible mixtures by the superselection rule for the system S on H, instead, will be all possible convex combinations (also infinite, in some operator topology) of the set

$$\bigsqcup_{k\in K} \mathfrak{S}(\mathsf{H}_{Sk}) \,.$$

The previous is equivalent to imposing that admissible states are the ρ in $\mathfrak{S}(\mathsf{H}_S)$ (or $\mathfrak{S}_p(\mathsf{H}_S)$) under which every H_{Sk} is invariant.

In this case the symmetries must respect the coherent decomposition of H, and one allows symmetries that swap sectors, i.e. functions $\gamma_{kk'} : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{S}(\mathsf{H}_{Sk'}), k, k' \in K$, possibly with $k' \neq k$. Every mapping $\gamma_{kk'} : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{S}(\mathsf{H}_{Sk'})$ must be invertible and satisfy Wigner's or Kadison's invariance.

12.1.3 Kadison symmetries

Consider a quantum system S described on the Hilbert space H_S, with space of states $\gamma(\mathsf{H}_S)$. A (weak, physically) demand to have a symmetry refers to the mixing procedure of quantum states. An operation on S defines a symmetry if the procedure is invariant under it. More precisely:

If a state is obtainable as mixture of certain states with given statistical weights, then by transforming the system under an operation that generates a symmetry, the same state must be obtainable as mixture of the transformed constituent states with the same statistical weights.

Put equivalently a bijection $\gamma : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$ is a symmetry when it preserves the *convex structure* of $\mathfrak{S}(\mathsf{H}_S)$: if $\rho_i \in \mathfrak{S}(\mathsf{H}_S)$, $0 \le p_i \le 1$ and $\sum_{i\in J} p_i = 1$, then

$$\gamma\left(\sum_{i\in J} p_i\rho_i\right) = \sum_{i\in J} p_i\gamma(\rho_i) \,.$$

Henceforth J will be finite. Then it is obvious, without loss of generality, that we may take J made of two points. Now we can present the formal definition in the general case, when coherent superselection sectors are present.

Definition 12.2 (Kadison symmetry). *Consider a quantum physical system S described on the Hilbert space H_S. Suppose H_S splits in coherent sectors $\mathsf{H}_S = \oplus_{k\in K}\mathsf{H}_{Sk}$. A* **symmetry** *of S (according to Kadison) from the sector H_{Sk} to the sector $\mathsf{H}_{Sk'}$, $k,k' \in K$, is a map $\gamma : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{S}(\mathsf{H}_{Sk'})$ such that:*

(a) *γ is bijective.*
(b) *γ preserves the convex structure of $\mathfrak{S}(\mathsf{H}_{Sk})$ and $\mathfrak{S}(\mathsf{H}_{Sk'})$. Equivalently:*

$$\gamma(p_1\rho_1 + p_2\rho_2) = p_1\gamma(\rho_1) + p_2\gamma(\rho_2)$$
$$\text{for } \rho_1,\rho_2 \in \mathfrak{S}(\mathsf{H}_S)\,, p_1 + p_2 = 1, p_1,p_2 \in [0,1]\,. \quad (12.1)$$

If the Hilbert space H does not have coherent sectors, every bijection $\gamma : \mathfrak{S}(\mathsf{H}) \to \mathfrak{S}(\mathsf{H})$ preserving convexity is called a **Kadison automorphism** *on H.*

A symmetry according to Definition 12.2 is induced by an operator $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$ that is either unitary or antiunitary (Definition 5.39):

$$\gamma^{(U)}(\rho) := U\rho U^{-1}\,, \quad \rho \in \mathfrak{S}(\mathsf{H}_{Sk}). \quad (12.2)$$

To prove this we need an elementary lemma.

Lemma 12.3. *Let $U : \mathsf{H} \to \mathsf{H}'$ be an antiunitary operator from H to H', and $N \subset \mathsf{H}$ a basis. Then $U = VC$, where $V : \mathsf{H} \to \mathsf{H}'$ is unitary and $C : \mathsf{H} \to \mathsf{H}$ is the natural conjugation (Definition 5.41) associated to N:*

$$C\psi := \sum_{z\in N} \overline{(z|\psi)}z\,.$$

Proof. Define $V\psi := \sum_{z\in N}(z|\psi)Uz$; the proof is immediate, because an antiunitary operator is anti-isometric and continuous and by elementary properties of bases. Note that $\{Uz\}_{z\in N}$ is basis of H'. □

Proposition 12.4. *Let $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$ be unitary (isometric and onto), or antiunitary, where H_{Sk} and $\mathsf{H}_{Sk'}$ are coherent sectors of the Hilbert space H_S associated to the quantum system S with space of states $\mathfrak{S}(\mathsf{H})$. Then $\gamma^{(U)} : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{B}(\mathsf{H})$ as defined in (12.2) is a symmetry according to Kadison for S, between H_{Sk} and $\mathsf{H}_{Sk'}$.*

Proof. Property (12.1) is trivial under either assumption on U (not so, though, if we allowed complex coefficients p_i). Let us prove $\gamma^{(U)}(\rho) \in \mathfrak{S}(\mathsf{H}_{Sk'})$ for $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$. Begin by assuming U unitary. If ρ is of trace class on H_S so must be $U\rho U^{-1}$ as well, since trace-class operators form an ideal in $\mathfrak{B}(\mathsf{H}_S)$ by Theorem 4.32(b) if we view $U\rho U^{-1}$ as composite in $\mathfrak{B}(\mathsf{H}_S)$. For this it suffices to think ρ in $\mathfrak{S}(\mathsf{H}_S)$ vanishing on the complement to H_{Sk}, and $\rho(\mathsf{H}_{Sk}) \subset \mathsf{H}_{Sk}$, then extend U and U^{-1} trivially on the orthogonal to H_{Sk} and $\mathsf{H}_{Sk'}$ respectively, hence viewing them as in $\mathfrak{B}(\mathsf{H}_S)$.

If $\rho \geq 0$ then $(\psi|U\rho U^{-1}\psi) = (U^*\psi|\rho U^*\psi) \geq 0$, so $\gamma^{(U)}(\rho) \geq 0$. Using the basis formed by a basis on H_{Sk} and one on $(\mathsf{H}_{Sk})^{\perp}$ we obtain $tr\left(\gamma^{(U)}(\rho)\right) = tr\left(U\rho U^{-1}\right) = tr(U^{-1}U\rho) = tr(\rho) = 1$. In the last passage we used $U^{-1}U\restriction_{\mathsf{H}_{Sk}} = I\restriction_{\mathsf{H}_{Sk}}$ in computing the trace, and that $\rho = 0$ on $(\mathsf{H}_{Sk})^{\perp}$. Therefore $\gamma^{(U)}(\rho) \in \mathfrak{S}(\mathsf{H}_{Sk'})$ for $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$. Now assume U antiunitary. Decompose U as in Lemma 12.3: $U = VC$ with respect to a basis $N \subset \mathsf{H}_S$ specified later. We claim $U\rho U^{-1}$ is positive, trace class with trace one. As V is unitary (in which case the claim holds by what we have just seen) and $U\rho U^{-1} = V(C\rho C^{-1})V^{-1}$, it is enough to prove the claim for $U = C$. Choose N to be made of eigenvectors ψ for ρ (Theorem 4.18 ensures its existence). Hence

$$\rho\phi = \sum_{\psi \in N} p_\psi (\psi|\phi)\psi\,,$$

$\phi \in \mathsf{H}$. Now recall C is continuous and antilinear, $CC = I$, $(f|g) = \overline{(Cf|Cg)}$ by definition of conjugation, every eigenvector p_ψ of ρ is real (positive), and $C\psi = \psi$. Consequently

$$C\rho C^{-1}\phi = \sum_{\psi \in N} p_\psi \overline{(\psi|C\phi)}C\psi = \sum_{\psi \in N} p_\psi \overline{(CC\psi|C\phi)}C\psi =$$

$$= \sum_{\psi \in N} p_\psi (C\psi|\phi)C\psi = \sum_{\psi \in N} p_\psi (\psi|C\phi)\psi = \rho\phi\,.$$

We proved $C\rho C^{-1} = \rho$, so $C\rho C^{-1}$ is of trace class, positive and has trace 1 for $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$. □

Example 12.5. If the superselection rule regards the electrical charge of a system there will be (infinitely many, in general) sectors H_q, one for each value q of the charge. *Charge conjugation* can be constructed as a collection of symmetries of type $\gamma^{(U_q)}$, and $U_q : H_q \to H_{-q}$ for any q. ■

12.1.4 Wigner symmetries

Now let us pass to quantum symmetries according to Wigner. Consider the usual quantum system S, described on the Hilbert space H_S and with space of states $\mathfrak{S}(\mathsf{H}_S)$. We focus on pure states $\mathfrak{S}_p(\mathsf{H}_S)$ (the rays of H_S). Let us restrict to transformations

$$\delta : \mathfrak{S}_p(\mathsf{H}_S) \to \mathfrak{S}_p(\mathsf{H}_S)\,.$$

From the experimental viewpoint we can control the transition probability $|(\psi|\psi')|^2 = tr(\rho\rho')$ of two pure states $\rho = \psi(\psi|\ )$, $\rho' = \psi'(\psi'|\ )$. Wigner's condition for a bijection $\delta : \mathfrak{S}_p(\mathsf{H}_S) \to \mathfrak{S}_p(\mathsf{H}_S)$ to be a symmetry is that it preserves transition probabilities.

If two pure states have a certain transition probability, when transforming the system by a physical operation that determines a symmetry the transformed states must maintain the same transition probability.

The next definition takes into account coherent sectors.

Definition 12.6 (Wigner symmetry). *Consider a quantum system S described on the Hilbert space H_S with space of states $\mathfrak{S}(\mathsf{H}_S)$. Assume H_S splits coherently as $\mathsf{H}_S = \oplus_{k\in K}\mathsf{H}_{Sk}$.*
A **symmetry** *of S (according to Wigner) from H_{Sk} to $\mathsf{H}_{Sk'}$, $k,k' \in K$, is a mapping $\delta : \mathfrak{S}_p(\mathsf{H}_{Sk}) \to \mathfrak{S}_p(\mathsf{H}_{Sk'})$ with the following properties:*

(a) *δ is bijective.*
(b) *δ preserves transition probabilities. That is to say:*

$$tr(\rho_1\rho_2) = tr(\delta(\rho_1)\delta(\rho_2)) , \quad \rho_1, \rho_2 \in \mathfrak{S}_p(\mathsf{H}_{Sk}) . \tag{12.3}$$

If H has no coherent sectors every bijection $\delta : \mathfrak{S}_p(\mathsf{H}) \to \mathfrak{S}_p(\mathsf{H})$ that preserves transition probabilities is a **Wigner automorphism** *on H.*

An example according to Definition 11.6, as with Kadison symmetries, is the symmetry induced by the (anti)unitary $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$ (Definition 5.32), where:

$$\delta^{(U)}(\rho) := U\rho U^{-1} , \quad \rho \in \mathfrak{S}_p(\mathsf{H}_{Sk}). \tag{12.4}$$

In contrast to Kadison's symmetries, here the proof is really straighforward.

Remark 12.7. (1) Since pure states have the form $\psi(\psi|\ )$, $||\psi|| = 1$, the action of $\delta^{(U)}$ on pure states can be described, equivalently though sloppily, by saying $\delta^{(U)}$ sends the pure state ψ to the pure state $U\psi$. This is the way in which QM books often describe symmetries induced by (anti)unitary operators.
(2) Every Kadison symmetry transforms pure states to pure states, so it defines a bijective map on the space of pure states. However, we do not know *a priori* that this will define a Wigner symmetry, because it is far from evident that it wil preserve transition probabilities. On the other hand, a Wigner symmetry does not extend naturally from pure to mixed states. Therefore it is not obvious that the two notions coincide. Yet every unitary or antiunitary operator determines at the same time a Wigner symmetry and a Kadison symmetry by means of $\rho \mapsto U\rho U^{-1}$. ■

To finish, here is a general notion of Wigner symmetry.

Definition 12.8 (General Wigner symmetry). *Suppose the Hilbert space H_S of system S decomposes in coherent sectors, so that admissible pure states are the elements of*

$$\mathfrak{S}_p(\mathsf{H}_S)_{adm} := \bigsqcup_{k\in K} \mathfrak{S}_p(\mathsf{H}_{Sk}) .$$

A **symmetry** *according to Wigner (no mention of sectors) is a bijective map δ from $\mathfrak{S}_p(\mathsf{H}_S)_{adm}$ to itself that preserves transition probabilities.*

We can recover the above definition using Wigner symmetries between pairs of sectors, as follows.

Proposition 12.9. *Let δ be a Wigner symmetry of S, and suppose the Hilbert space H_S of S splits coherently in such a way that admissible pure states are only those in:*

$$\mathfrak{S}_p(\mathsf{H}_S)_{adm} = \bigsqcup_{k\in K} \mathfrak{S}_p(\mathsf{H}_{Sk}) .$$

There exists a bijection $f : K \to K$ and a family of Wigner symmetries

$$\delta_{f,f(k)} : \mathfrak{S}_p(\mathsf{H}_{Sk}) \to \mathfrak{S}_p(\mathsf{H}_{Sf(k)}) , \quad k \in K,$$

with fixed sectors, such that $\delta\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})} = \delta_{f,f(k)}$ for every k. In this sense δ is just a collection of Wigner symmetries exchanging sectors and not overlapping.

Proof. Define on $\mathfrak{S}_p(\mathsf{H}_S)$ the distance

$$d(\rho,\rho') := ||\rho - \rho'||_1 := tr(|\rho - \rho'|) ,$$

where $|| \ ||_1$ is the canonical norm of trace-class operators. Then the sets $\mathfrak{S}_p(\mathsf{H}_{Sk})$ are the connected components of $\mathfrak{S}_p(\mathsf{H}_S)_{adm}$ (Exercise 12.4). The map $\delta : \mathfrak{S}_p(\mathsf{H}_S)_{adm} \to \mathfrak{S}_p(\mathsf{H}_S)_{adm}$ is a surjective isometry for d as follows from Proposition 12.32 (the latter is independent of the present result). In particular δ is a homeomorphism. As such, it preserves maximal connected subsets, and so it must split in isometric bijections between distinct sectors, i.e. Wigner symmetries between distinct sectors. □

12.1.5 The theorems of Wigner and Kadison

We begin by Wigner's theorem. Using that, we will prove Kadison's result along the lines of [Sim76]. The proof of Wigner's theorem is quite direct. Although there are more elegant, but indirect arguments, our approach has the advantage of showing explicitly how to manifacture U with a basis. Let us remark, in passing, that several authors (inclusing Emch, Piron and Bargmann) proved that a slightly modified version of Wigner's theorem holds within QM formulations based on real and quaternionic Hilbert spaces.

Theorem 12.10 (Wigner). *Consider a quantum system S described on the (separable, complex) Hilbert space H_S. Suppose H_S coherently splits as $\mathsf{H}_S = \oplus_{k\in K}\mathsf{H}_{Sk}$ (if $K = \{1\}$ one should replace H_{Sk}, $\mathsf{H}_{Sk'}$ by H in the statement). Assume*

$$\delta : \mathfrak{S}_p(\mathsf{H}_{Sk}) \to \mathfrak{S}_p(\mathsf{H}_{Sk'})$$

is a (Wigner) **symmetry** *of S from H_{Sk} to $\mathsf{H}_{Sk'}$, $k,k' \in K$. Then:*

(a) *there exists an operator $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$, unitary or antiunitary (depending on δ), such that:*

$$\delta(\rho) = U\rho U^{-1} \quad \text{for any pure state } \rho \in \mathfrak{S}_p(\mathsf{H}_{Sk}). \tag{12.5}$$

(b) *U is determined up to a phase factor, i.e. U_1, U_2 (both unitary or antiunitary) satisfy (12.5) (replacing each with U) if and only if $U_2 = \chi U_1$ for some $\chi \in \mathbb{C}$, $|\chi| = 1$.*
(c) *If $\{\psi_n\}_{n\in\mathbb{N}}$ is a basis on H_{Sk} and $\psi'_n \in \mathsf{H}_{Sk'}$ are chosen so that $\psi'_n(\psi'_n| \ ) = \delta(\psi_n(\psi_n| \ ))$, then $\{\psi'_n\}_{n\in\mathbb{N}}$ is a basis on $\mathsf{H}_{Sk'}$. Furthermore, an operator U satisfying (12.5) is*

$$U : \psi = \sum_{n\in\mathbb{N}} a_n\psi_n \mapsto \sum_{n\in\mathbb{N}} a_n\psi'_n \quad \text{in the unitary case,}$$

or

$$U : \psi = \sum_{n\in\mathbb{N}} a_n\psi_n \mapsto \sum_{n\in\mathbb{N}} \overline{a_n}\psi'_n \quad \text{in the antiunitary case.}$$

Proof. (b) Begin by showing that U, if it exists, is unique up to a choice of a phase factor. Clearly if U_1 satisfies the thesis for δ, then $U_2 := \chi U_1$ will do the same for any $\chi \in \mathbb{C}$, $|\chi| = 1$. We claim this is the only possibility. Suppose there are U_1, U_2 (both unitary or antiunitary) for δ. If $\rho = \psi(\psi|\ )$ then, setting $L := U_1^{-1}U_2$ we have $L\psi(\psi|L^{-1}\phi) = \psi(\psi|\phi)$ for any unit ψ, ϕ. Hence $L\psi(L\psi|\phi) = \psi(\psi|\phi)$, as L is unitary. Since $L\psi(L\psi|\ ) = \psi(\psi|\ )$, $L\psi$ and ψ determine the same pure state, so $L\psi = \chi_\psi \psi$ i.e. $U_1\psi = \chi_\psi U_2\psi$, or $U_1\psi = \overline{\chi_\psi} U_2\psi$ (antiunitary case), for every $\psi \in \mathsf{H}_{Sk}$ and for some unit $\chi_\psi \in \mathbb{C}$. The result remains valid even if $||\psi|| \neq 1$, by linearity. Let us prove χ_ψ does not depend on ψ. Notice that $\chi_\psi = \chi_{c\psi}$ if $c \in \mathbb{C}$. Choose ψ, ψ' linearly independent and $a, b \in \mathbb{C} \setminus \{0\}$. The linearity of L implies

$$\chi_{a\psi+b\psi'}(a\psi + b\psi') = L(a\psi + b\psi') = aL\psi + bL\psi' = a\chi_\psi\psi + b\chi_{\psi'}\psi' .$$

Therefore

$$a(\chi_{a\psi+b\psi'} - \chi_\psi)\psi = b(\chi_{\psi'} - \chi_{a\psi+b\psi'})\psi' .$$

As ψ, ψ' are linearly independent and $a, b \neq 0$, we have $(\chi_{a\psi+b\psi'} - \chi_\psi) = 0$ and $(\chi_{\psi'} - \chi_{a\psi+b\psi'}) = 0$, so $\chi_\psi = \chi_{\psi'}$. Hence we have, for some unit $\chi \in \mathbb{C}$,

$$U_2\psi = \chi U_1\psi \quad \text{for any } \psi \in \mathsf{H}_{Sk}.$$

(a)–(c) Let us build an operator U representing δ. Take a basis $\{\psi_n\}_{n\in\mathbb{N}}$ in H_{Sk}. To each ψ_n associate the pure state $\rho_n := \psi_n(\psi_n|\ )$. Let δ act on these states, obtaining pure states $\delta(\rho_n) = \psi'_n(\psi'_n|\ ) \in \mathfrak{S}_p(\mathsf{H}_{Sk'})$, where the unit $\psi'_n \in \mathsf{H}_{Sk'}$ are determined up to a phase factor. Fix once for all this phase, arbitrarily. Note $\{\psi'_n\}_{n\in\mathbb{N}}$ is a basis of $\mathsf{H}_{Sk'}$: the vectors are in fact orthonormal, because $|(\psi'_n|\psi'_m)|^2 = tr(\delta(\rho_n)\delta(\rho_m)) = tr(\rho_n\rho_m) = |(\psi_n|\psi_m)|^2 = \delta_{nm}$. We show that $\psi' \perp \psi'_n$ implies $\psi' = 0$. Let $\psi' \perp \psi'_n$ for every $n \in \mathbb{N}$. If $\psi' \neq 0$, without loss of generality we assume $||\psi'|| = 1$ and define $\rho' := \psi'(\psi'|\ ) \in \mathfrak{S}_p(\mathsf{H}_{Sk'})$. Since δ is onto, then $\rho' = \delta(\rho)$ with $\rho = \psi(\psi|\ )$ for some $\psi \in \mathsf{H}_{Sk}$, $||\psi|| = 1$. Therefore:

$$|(\psi'|\psi'_n)|^2 = tr(\delta(\rho)\delta(\rho_n)) = tr(\rho\rho_n) = |(\psi|\psi_n)|^2 = 0$$

and then $\psi = 0$, for $\{\psi_n\}_{n\in\mathbb{N}}$ is a basis. But this is impossible as $||\psi|| = 1$. Consequently $\psi' = 0$, and $\{\psi'_n\}_{n\in\mathbb{N}}$ is a basis.

Using the bases $\{\psi_n\}_{n\in\mathbb{N}}$ and $\{\psi'_n\}_{n\in\mathbb{N}}$ we will define the operator U in stages. First define unit vectors

$$\Psi_k := 2^{-1/2}(\psi_0 + \psi_k) , \quad k \in \mathbb{N} \setminus \{0\}$$

and corresponding pure states: $(\Psi_k|\ )\Psi_k$, $k \in \mathbb{N} \setminus \{0\}$. The transformed $\delta(\Psi_k(\Psi_k|\ )) = \Psi'_k(\Psi'_k|\ )$ satisfies, in particular:

$$|(\Psi'_k|\psi'_n)|^2 = tr\left(\Psi'_k(\Psi'_k|\ )\delta(\rho_n)\right) = tr(\delta(\Psi_k(\Psi_k|\ ))\delta(\rho_n)) = |(\Psi_k|\psi_n)|^2 = \frac{\delta_{0n} + \delta_{kn}}{2} ,$$

plus $||\Psi'_k|| = 1$. Decomposing $\Psi'_k = \sum_n a_n \psi'_n$, the only possibility is

$$\Psi'_k = \chi'_k 2^{-1/2}(\psi'_0 + \chi_k \psi'_k)$$

with $|\chi'_k| = |\chi_k| = 1$. The χ_k are given by δ, while the χ'_k can be chosen as we want. The χ_k carry the information of δ, and we shall employ them soon.

Let us define U on ψ_n and $(\psi_0 + \psi_k)/\sqrt{2}$ by declaring:

$$U\psi_0 := \psi'_0 \,, \quad U\psi_k := \chi_k \psi'_k \,, \quad U(2^{-1/2}(\psi_0 + \psi_k)) := 2^{-1/2}(\psi'_0 + \chi_k \psi'_k) \,, \quad (12.6)$$

$k \in \mathbb{N} \setminus \{0\}$. With this we are sure that if ϕ is one of the above arguments of U and ρ_ϕ its pure state, then $\delta(\rho_\phi)$ is associated to $U\phi$.

Now we extend U to any $\psi = \sum_{n\in\mathbb{N}} a_n \psi_n \in \mathsf{H}_{Sk}$, so that U continues to represent δ. Assume as before $||\psi|| = 1$ and $a_0 \in \mathbb{R} \setminus \{0\}$. Let $\psi' \in \mathsf{H}_{Sk'}$ with $||\psi'|| = 1$ be such that $\psi'(\psi'|\;) = \delta(\rho_\psi)$. Then

$$\psi' = \sum_{n\in\mathbb{N}} a'_n \psi'_n \,. \quad (12.7)$$

The coefficients a'_k are given, up to a global phase factor, by the coefficients a_n and by δ. With our assmptions on δ we have

$$|(\psi'|\psi'_n)|^2 = tr(\delta(\rho_\psi)\delta(\rho_n)) = tr(\rho_\psi \rho_n) = |(\psi|\psi_n)|^2 \,.$$

In other words, $|a'_n| = |a_n|$. Using this, together with the first two of (12.6) in the right-hand side of (12.7), leads to

$$\psi' = \chi \left(a_0 U\psi_0 + \sum_{n\in\mathbb{N}\setminus\{0\}} \chi_n^{-1} a'_n U\psi_n \right) \,,$$

where χ, $|\chi| = 1$, is arbitrary. Now define

$$U\psi := a_0 U\psi_0 + \sum_{n\in\mathbb{N}\setminus\{0\}} \chi_n^{-1} a'_n U\psi_n \,. \quad (12.8)$$

This ensures, by construction, $U\psi(U\psi|\quad) = \delta(\rho_\psi)$, and one verifies the definition extends (12.6). However, $U\psi$ is not completely fixed, because we still do not know the coefficients a'_n in terms of the components a_n of ψ. At this point we can define it in full, though. By construction of U and in the hypotheses on δ, $|(\Psi_k|\psi)| = |(U\Psi_k|U\psi)|$. By (12.8) this means

$$|a_0 + a_k|^2 = |a_0 + \chi_k^{-1} a'_k|^2 \,.$$

Since $|a_k| = |a'_k|$, the latter implies

$$Re(a_0 a_k) = Re(a_0 \chi_k^{-1} a'_k) \,.$$

Now recalling $a_0 \in \mathbb{R} \setminus \{0\}$, the previous equations occur only in one of these cases:

$$a'_k = \chi_k a_k \quad \text{or} \quad a'_k = \overline{\chi_k a_k} \,.$$

Hence, for any $\psi = \sum_n a_n \psi_n$ with $a_0 \in \mathbb{R} \setminus \{0\}$ we have

$$\psi' = U\psi = \sum_{n\in A_\psi} a_n \psi'_n + \sum_{n\in B_\psi} \overline{a_n} \psi'_n \,.$$

For the given ψ we can always choose one of A_ψ, B_ψ empty[1]. Suppose the contrary. The components of ψ and ψ' satisfy $a'_p = \chi_p a_p$ and $a'_q = \overline{\chi_q a_q}$, for some pair $p \neq q$, where $Im a_p, Im a_q \neq 0$. If $\phi = 3^{-1/2}(\psi_0 + \psi_p + \psi_q)$ then by construction:

$$|(\phi'|\psi')|^2 = |(\phi|\psi)|^2 ,$$

where $\phi' := U\phi = 3^{-1/2}(\psi'_0 + \psi'_p + \psi'_q)$. The displayed equation reads

$$\left|a_0 + a_p + a_q\right|^2 = \left|a_0 + a_p + \overline{a_q}\right|^2 ,$$

i.e.

$$Re(a_p a_q) = Re(a_p \overline{a_q}).$$

This would imply $Im a_q = -Im a_q$, an absurd.

If $\psi = \sum_n a_n \psi_n \in \mathsf{H}_{Sk}$, $||\psi|| = 1$ and $a_0 \in \mathbb{R} \setminus \{0\}$, there are two possible definitions for $U\psi$:

$$U\psi = \sum_{n\in\mathbb{N}} a_n \psi'_n \quad \text{or} \quad U\psi = \sum_{n\in\mathbb{N}} \overline{a_n} \psi'_n . \tag{12.9}$$

We claim the choice does *not* depend on ψ, and so it must depend on the nature of δ. Consider a generic unit $\psi = \sum_n a_n \psi_n \in \mathsf{H}_{Sk}$ and $a_0 \in \mathbb{R} \setminus \{0\}$. Define $\psi^{(nc)}$ associated to $c \in \mathbb{C}$, $Imc \neq 0$, for every $n = 1, 2, \ldots$:

$$\psi^{(nc)} := \frac{1}{\sqrt{1+|c|^2}}(\psi_0 + c\psi_n) .$$

Since $|(\psi|\psi^{(nc)})| = |(U\psi|U\psi^{(nc)})|$ has to hold, necessarily $\psi^{(nc)}$ and ψ are of the same type with respect to the option of (12.9). Therefore all $\psi = \sum_n a_n \psi_n \in \mathsf{H}_{Sk}$, $||\psi|| = 1$, and $a_0 \in \mathbb{R} \setminus \{0\}$ are of the same type.

Define the operator $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$ by:

$$U : \psi = \sum_{n\in\mathbb{N}} a_n \psi_n \mapsto \sum_{n\in\mathbb{N}} a_n \psi'_n \quad \text{in the linear case,}$$

$$U : \psi = \sum_{n\in\mathbb{N}} a_n \psi_n \mapsto \sum_{n\in\mathbb{N}} \overline{a_n} \psi'_n \quad \text{in the antilinear case.}$$

By construction, the former is isometric and onto (unitary), the latter anti-isometric and onto (antiunitary). At this juncture it should be clear that the unitary and antiunitary cases are distinguished by δ, and we cannot represent the same δ by a unitary operator or by an antiunitary one alike. That is because it is impossible that $\psi' := \sum_{n\in\mathbb{N}} a_n \psi_n \mapsto \sum_{n\in\mathbb{N}} \overline{a_n} \psi'_n$ and $\tilde{\psi}' = \sum_{n\in\mathbb{N}} a_n \psi_n \mapsto \sum_{n\in\mathbb{N}} a_n \psi'_n$ differ by a phase factor, for any choice of the a_n, i.e. of ψ, as should happen, on the contrary, if ψ', $\tilde{\psi}'$ determined the same pure state $\delta(\psi(\psi|\ ))$.

By construction, the operator satisfies $U\rho U^{-1} = \delta(\rho)$ as long as $\rho \in \mathfrak{S}_p(\mathsf{H}_{Sk})$ can be written as $\psi(\psi|\ )$, where $a_0 \neq 0$ in $\psi = \sum_{n\in\mathbb{N}} a_n \psi_n$. If so, in fact, we can redefine ψ

[1] There is a certain ambiguity in defining A_ψ and B_ψ, because the subscripts n of the possible real coefficients a_n can be chosen either in A_n or in B_n indifferently.

by changing one global phase, $\tilde{\psi} = \chi\psi$, and without altering $\rho = \psi(\psi|\ ) = \tilde{\psi}(\tilde{\psi}|\ )$, so that $\tilde{a}_0 \in \mathbb{R} \setminus \{0\}$ in the expansion of $\tilde{\psi}$. Now the construction for U implies

$$U\rho U^{-1} = U\psi(\psi|\ )U^{-1} = U\tilde{\psi}(\tilde{\psi}|\ )U^{-1} = U\tilde{\psi}(U\tilde{\psi}|\ ) = \delta(\rho)\,.$$

There remains to prove this fact for pure states associated to $\psi = \sum_{n\in\mathbb{N}} a_n\psi_n$ with $a_0 = 0$. To this end notice that the entire procedure works if we replace ψ_0 with any other basis vector ψ_k. In that case $U\rho U^{-1} = \delta(\rho)$ for pure states associated to $\psi = \sum_{n\in\mathbb{N}} a_n\psi_n$ with $a_k \neq 0$. (Using ψ_k as reference vector instead of ψ_0, the new U cannot be of different character (linear vs. antilinear) from the one of ψ_0. That is because the two operators must behave in the same way on $\psi = \sum_{n\in\mathbb{N}} a_n\psi_n$, if $a_k \neq 0$ and $a_0 \neq 0$, and this fact fixes their character, as we saw above.) This observation essentially ends the proof: take $\rho \in \mathfrak{S}_p(\mathsf{H}_{Sk})$ and $\rho = \psi(\psi|\ )$, with $\psi = \sum_{n\in\mathbb{N}} a_n\psi_n$ and $a_0 = 0$. There must be one $a_k \neq 0$ at least, for $||\psi|| = 1$. Therefore we may go through the previous argument with this ψ_k in place of ψ_0. □

Let us move on to Kadison's theorem, which we will reduce to Wigner's theorem following an idea of Roberts and Roepstorff [RR69], see [Sim76]. We start by proving the theorem in dimension two.

Proposition 12.11. *Let* H *be a two-dimensional Hilbert space. If* $\gamma : \mathfrak{S}(\mathsf{H}) \to \mathfrak{S}(\mathsf{H})$ *is a Kadison automorphism, there exists* $U : \mathsf{H} \to \mathsf{H}$ *unitary, or antiunitary, such that:*

$$\gamma(\rho) = U\rho U^{-1}\,, \quad \rho \in \mathfrak{S}(\mathsf{H}).$$

Proof. Let us characterise states and pure states on H by means of the *Poincaré sphere*. A state $\rho \in \mathfrak{S}(\mathsf{H})$ is, in the present situation, a positive-definite Hermitian matrix with unit trace. The real space of Hermitian matrices has a basis made by I and the Pauli matrices:

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\,,\ \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\,,\ \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\,. \tag{12.10}$$

So for some $a, b_n \in \mathbb{R}$

$$\rho = aI + \sum_{n=1}^{3} b_n\sigma_n\,.$$

The condition $tr(\rho) = 1$ fixes $a = 1/2$, since the σ_n are traceless. Positive definiteness, i.e. the demand the two eigenvalues of ρ are positive, is equivalent to $\sqrt{b_1^1 + b_2^2 + b_3^2} \leq 1/2$, by direct computation. Overall, the elements ρ of $\mathfrak{S}(\mathsf{H})$ are in one-to-one correspondence to vectors $\mathbf{n} \in \mathbb{R}^3$, $|\mathbf{n}| \leq 1$:

$$\rho = \frac{1}{2}\left(I + \mathbf{n}\cdot\sigma\right)\,. \tag{12.11}$$

Having ρ pure, i.e. having a unique eigenvalue 1, is equivalent to $|\mathbf{n}| = 1$, as a direct check shows. Altogether $\mathfrak{S}(\mathsf{H})$ is in one-to-one correspondence with the closed unit ball B in $\mathbb{R}^3$ centred at the origin; and the subset of pure states $\mathfrak{S}_p(\mathsf{H})$ corresponds one-to-one with the surface ∂B. We call B, seen in this way, the **Poincaré sphere**. The correspondence just defined:

$$B \ni \mathbf{n} \mapsto \rho_{\mathbf{n}} \in \mathfrak{S}(\mathsf{H})$$

is a true isomorphism: by (12.11), in fact,

$$\rho_{p\mathbf{n}+q\mathbf{m}} = p\rho_{\mathbf{n}} + q\rho_{\mathbf{m}} \quad \text{for any } \mathbf{n},\mathbf{m} \in B \text{ if } p,q \geq 0,\ p+q=1$$

so the convex geometry of the spaces is preserved. An important property, for later, is the formula

$$tr(\rho_{\mathbf{m}}\rho_{\mathbf{n}}) = \frac{1}{2}(1+\mathbf{m}\cdot\mathbf{n}) \tag{12.12}$$

that descends directly from $tr(\sigma_j) = 0$, $tr(\sigma_i\sigma_j) = 2\delta_{ij}$ (easy to prove). We are ready to characterise Kadison automorphisms. Assigning a Kadison automorphism $\gamma : \mathfrak{S}(\mathsf{H}) \to \mathfrak{S}(\mathsf{H})$ is patently the same as defining a bijection $\gamma' : B \to B$ such that

$$\gamma'(p\mathbf{n}+q\mathbf{m}) = p\gamma'(\mathbf{n}) + q\gamma'(\mathbf{m}) \quad \text{for any } \mathbf{n},\mathbf{m} \in B \text{ if } p,q \geq 0,\ p+q=1\,.$$

If the Kadison automorphism $\gamma : \mathfrak{S}(\mathsf{H}) \to \mathfrak{S}(\mathsf{H})$ fixes $\gamma' : B \to B$, the map $\Gamma : \mathbb{R}^3 \to \mathbb{R}^3$:

$$\Gamma(\mathbf{0}) := \mathbf{0}\,, \quad \Gamma(\mathbf{v}) := |\mathbf{v}|\gamma'\left(\frac{\mathbf{v}}{|\mathbf{v}|}\right), \quad \mathbf{v} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$$

extends γ', is linear and invertible. The proof is straighforward. Kadison automorphisms, being isomorphisms, map extreme elements to extreme elements, so $\Gamma(\mathbf{n}) = \gamma'(\mathbf{n}) = 1$ if $|\mathbf{n}| = 1$, and by linearity:

$$|\Gamma(\mathbf{v})| = |\mathbf{v}|\,, \quad \mathbf{v} \in \mathbb{R}^3\,.$$

In conclusion the linear map $\Gamma : \mathbb{R}^3 \to \mathbb{R}^3$ associated to the Kadison automorphism γ is an isometry of $\mathbb{R}^3$ with the origin as fixed point. This can be if and only if $\Gamma \in O(3)$, the *three-dimensional orthogonal group*. $\Gamma \in O(3)$ implies $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H})}$ is an automorphism according to Wigner. In fact if $\rho_{\mathbf{n}}$ and $\rho_{\mathbf{m}}$ are pure, their transition probability equals $tr(\rho_{\mathbf{n}}\rho_{\mathbf{m}})$, which we can express via (12.12). Since Γ is orthogonal:

$$tr(\gamma(\rho_{\mathbf{n}})\gamma(\rho_{\mathbf{m}})) = \frac{1}{2}(1+\Gamma(\mathbf{n})\cdot\Gamma(\mathbf{m})) = \frac{1}{2}(1+\mathbf{n}\cdot\mathbf{m}) = tr(\rho_{\mathbf{n}}\rho_{\mathbf{m}})\,.$$

Recalling $\gamma'\restriction_{\partial B} = \Gamma\restriction_{\partial B} : \partial B \to \partial B$ is trivially a bijection (this is true for every orthogonal matrix), then $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H})} : \mathfrak{S}_p(\mathsf{H}) \to \mathfrak{S}_p(\mathsf{H})$ is bijective. We conclude $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H})} : \mathfrak{S}_p(\mathsf{H}) \to \mathfrak{S}_p(\mathsf{H})$ is a Wigner automorphism. Wigner's theorem implies the existence of a unitary or antiunitary operator $U : \mathsf{H} \to \mathsf{H}$ such that $\gamma(\rho) = U\rho U^{-1}$ for any $\rho \in \mathfrak{S}_p(\mathsf{H})$. If $\rho \in \mathfrak{S}(\mathsf{H})$ we can decompose it as convex combinations of two pure states associated to the eigenvectors of ρ. If $\rho_1, \rho_2 \in \mathfrak{S}_p(\mathsf{H})$ are the states in question for some $p \in [0,1]$, then $\rho = p\rho_1 + (1-p)\rho_2$, and so

$$\begin{aligned}\gamma(\rho) &= p\gamma(\rho_1) + (1-p)\gamma(\rho_2) = pU\rho_1U^{-1} + (1-p)U\rho_2U^{-1} \\ &= U\left(p\rho_1 + (1-p)\rho_2\right)U^{-1} = U\rho U^{-1}\,.\end{aligned}$$

Therefore the unitary (or antiunitary) operator U satisfies the theorem's claim. □

Let us state and prove Kadison's theorem in general. (Kadison originally proved the non-trivial statements (a) and (b)).

Theorem 12.12 (Kadison). *Consider a quantum system S described on the (separable, complex) Hilbert space H_S. Let H_S decompose in coherent sectors $\mathsf{H}_S = \oplus_{k\in K}\mathsf{H}_{Sk}$ (possibly $K = \{1\}$, in which case H_S should replace every H_{Sk}). Suppose the map*

$$\gamma : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{S}(\mathsf{H}_{Sk'})$$

is a **symmetry** *of S (according to Kadison) from H_{Sk} to $\mathsf{H}_{Sk'}$, $k,k' \in K$. Then*

(a) *there exists an operator $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$, unitary or antiunitary, such that:*

$$\gamma(\rho) = U\rho U^{-1} \quad \text{for every pure state } \rho \in \mathfrak{S}(\mathsf{H}_{Sk}). \tag{12.13}$$

(b) *U is determined up to phase, i.e. U_1 and U_2 (both unitary or antiunitary) satisfy (12.13) (replacing U) if and only if $U_2 = \chi U_1$ with $\chi \in \mathbb{C}$, $|\chi| = 1$.*
(c) *The restriction of γ to pure states is a Wigner symmetry (and the unitary vs. antiunitary character of U in (a) is determined by $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}$).*
(d) *Every Wigner symmetry $\delta : \mathfrak{S}_p(\mathsf{H}_{Sk}) \to \mathfrak{S}_p(\mathsf{H}_{Sk'})$ extends, uniquely, to a Kadison symmetry $\gamma^{(\delta)} : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{S}(\mathsf{H}_{Sk'})$.*

Proof. (b) Supposing U exists and is either unitary or antiunitary, let us prove uniqueness up to phase in the class of operators. If U_1 saisfies the theorem for γ, then $U_2 := \chi U_1$ will satisfy it if $\chi \in \mathbb{C}$, $|\chi| = 1$. We claim there are no other possibilities. Suppose U_1, U_2 (of the same type) satisfy the theorem for γ. In particular, if $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$ then $U_1\rho U_1^{-1} = U_2\rho U_2^{-1}$, so $L\rho L^{-1} = \rho$ where $L := U_1^{-1}U_2$ is linear and unitary. Choosing a pure state $\rho = \psi(\psi|\ )$, the identity reads

$$L\psi(L\psi|\ ) = \psi(\psi|\ )$$

so $L\psi$ belongs to the ray of ψ, and $L\psi = \chi_\psi \psi$ for some unit $\chi_\psi \in \mathbb{C}$. Exactly as in part (b) of Wigner's theorem, χ_ψ does not depend on ψ, ending the proof of (b).

Let us pass to (a) and divide the proof in steps. First, we notice γ is bijective and preserves convexity. Therefore it maps extreme elements to extreme elements and the same for non-extreme ones, i.e. pure states to pure states and mixed ones to mixed ones. We claim that if $\mathsf{M} \subset \mathsf{H}_{Sk}$ is a two-dimensional subspace there is an analogous two-dimensional $\mathsf{M}' \subset \mathsf{H}_{Sk'}$ such that $\gamma(\mathfrak{S}(\mathsf{M})) = \mathfrak{S}(\mathsf{M}')$. If ψ_1, ψ_2 is a basis of M, the generic element in $\mathfrak{S}(\mathsf{M})$ is $\rho = p\psi_1(\psi_1|\ ) + q\psi_2(\psi_2|\ )$, $p+q = 1$ and $p,q \geq 0$. Hence:

$$\gamma(\rho) = p\gamma(\psi_1(\psi_1|\ )) + q\gamma(\psi_2(\psi_2|\ )) = p\psi_1'(\psi_1'|\ ) + q\psi_2'(\psi_2'|\ ),$$

where the unit ψ_1', ψ_2' arise (up to phase) from the corresponding pure states $\gamma(\psi_1(\psi_1|\ ))$, $\gamma(\psi_1(\psi_1|\ ))$. The latter must be distinct, otherwise the bijection $\gamma^{-1} : \mathfrak{S}(\mathsf{H}_{Sk'}) \to \mathfrak{S}(\mathsf{H}_{Sk})$, that preserves convexity, would map pure to mixed. So ψ_1' and ψ_2', both unit, satisfy $\psi_1' \neq a\psi_2'$ for any $a \in \mathbb{C}$, and thus are linearly independent. The space M' is then generated by ψ_1', ψ_2'.

Now we need two lemmas.

Lemma 12.13. *Under our assumptions on γ, there is a Wigner symmetry δ : $\mathfrak{S}_p(\mathsf{H}_{Sk}) \to \mathfrak{S}_p(\mathsf{H}_{Sk'})$ such that $\gamma(\rho) = \delta(\rho)$ for every $\rho \in \mathfrak{S}_p(\mathsf{H}_{Sk})$.*

Proof of Lemma 12.13. Since γ and γ^{-1} preserve extreme and non-extreme sets, $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}: \mathfrak{S}_p(\mathsf{H}_{Sk}) \to \mathfrak{S}_p(\mathsf{H}_{Sk'})$ is invertible, because the left and right inverse is just $\gamma^{-1}\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk'})}: \mathfrak{S}_p(\mathsf{H}_{Sk'}) \to \mathfrak{S}_p(\mathsf{H}_{Sk})$. The proof ends once we show $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}$ preserves transition probabilities. Given $\phi, \psi \in \mathsf{H}_{Sk}$ unit and distinct, let M be their span and $\mathsf{M}' \subset \mathsf{H}_{Sk'}$ the two-dimensional space such that $\gamma(\mathfrak{S}(\mathsf{M})) \subset \mathfrak{S}(\mathsf{M}')$, mentioned above. Call $U : \mathsf{M}' \to \mathsf{M}$ an arbitrary unitary operator. Define

$$\gamma'(\rho) := U\gamma(\rho)U^{-1}, \quad \rho \in \mathfrak{S}(\mathsf{M}).$$

Immediately, γ' is a Kadison symmetry if we restrict to the 2-dimensional Hilbert space $\mathsf{H} = \mathsf{M}$. As shown in Proposition 12.11, Kadison's theorem holds and there is a unitary, or antiunitary, $V : \mathsf{M} \to \mathsf{M}$ such that $\gamma'(\rho) = U\gamma(\rho)U^{-1} = V\rho V^{-1}$. Otherwise said:

$$\gamma(\rho) = UV\rho(UV)^{-1}, \quad \rho \in \mathfrak{S}(\mathsf{M}).$$

In particular, choosing $\rho = \psi(\psi|\ )$ and then $\rho = \phi(\phi|\ )$ gives

$$tr\left(\gamma(\psi(\psi|\ ))\gamma(\phi(\phi|\ ))\right) = tr\left(UV\psi(\psi|\ )(UV)^{-1}UV\phi(\phi|\ )(UV)^{-1}\right) =$$

$$= tr\left(UV\psi(\psi|\ )\phi(\phi|\ )(UV)^{-1}\right) = tr\left(\psi(\psi|\ )\phi(\phi|\ )\right).$$

If $\psi(\psi|\ ) = \phi(\phi|\ )$ the result is the same, as one sees easily. So we proved that $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}$ preserves transition probabilities, and so is a Wigner symmetry. □

By the previous lemma, and invoking Wigner's Theorem 12.10, there exists a unitary, or antiunitary, operator $U : \mathsf{H}_{Sk} \to \mathsf{H}_{Sk'}$ such that

$$\gamma(\rho) = U\rho U^{-1}, \quad \rho \in \mathfrak{S}_p(\mathsf{H}_{Sk}). \tag{12.14}$$

The proof now ends if we prove that the above identity holds also for $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$, and not only for $\mathfrak{S}_p(\mathsf{H}_{Sk})$. For that, note (12.14) is equivalent to:

$$U^{-1}\gamma(\rho)U = \rho, \quad \rho \in \mathfrak{S}_p(\mathsf{H}_{Sk}).$$

Therefore $\Gamma := U^{-1}\gamma(\,\cdot\,)U : \mathfrak{S}(\mathsf{H}_{Sk}) \to \mathfrak{S}(\mathsf{H}_{Sk})$ is a Kadison symmetry (a Kadison automorphism, actually) that reduces to the identity on pure states. Kadison's theorem is eventually proved after we establish the following lemma.

Lemma 12.14. *Let H be a Hilbert space. If a Kadison automorphism $\Gamma : \mathfrak{S}(\mathsf{H}) \to \mathfrak{S}(\mathsf{H})$ restricts to the identity on pure states, it is the identity.*

Proof of Lemma 12.14. Let $\rho = \sum_{k=0}^{N} p_k \psi_k(\psi_k|\ )$ be a finite (convex) combination of pure states. Then

$$\Gamma(\rho) = \Gamma\left(\sum_{k=0}^{N} p_k\psi_k(\psi_k|\ )\right) = \sum_{k=0}^{N} p_k\Gamma\left(\psi_k(\psi_k|\ )\right) = \left(\sum_{k=0}^{N} p_k\right) I = I.$$

Therefore the claim holds for every $\rho \in \mathfrak{S}(\mathsf{H})$ provided finite (convex) combination of pure states are dense in $\mathfrak{S}(\mathsf{H})$ in some topology for which Γ is continuous. Let us show this works if we take the topology of trace-class operators induced by the norm $||T||_1 := tr(|T|)$ (see Chapter 4).

If $\rho \in \mathfrak{S}(\mathsf{H})$ we can decompose the operator spectrally:

$$\rho = \sum_{k\in\mathbb{N}} p_k \psi_k(\psi_k|\ ),$$

where $p_k > 0$, $\sum_{k\in\mathbb{N}} p_k = 1$. Convergence is understood in the strong topology, and also in uniform topology, as we know from Chapter 4. Let us prove, further, we may approximate ρ by finite (convex) combinations $\rho_N \in \mathfrak{S}(\mathsf{H})$ of pure states, so that $||\rho_N - \rho||_1 \to 0$ per $N \to +\infty$. To this end set:

$$\rho_N := \sum_{k=0}^{N} q_k^{(N)} \psi_k(\psi_k|\ ), \quad q_k^{(N)} := \frac{p_k}{\sum_{j=0}^{N} p_j}, \quad \text{N=0,1,2,}\ldots .$$

Evidently $\rho_N \in \mathfrak{S}(\mathsf{H})$ for any $N \in \mathbb{N}$. Since $q_k^{(N)} > p_k$ and the unit vectors ψ_k (adding a basis of $ker(\rho) \supset ker(\rho_N)$) give a basis of H of eigenvectors of ρ, ρ_N and hence of $\rho - \rho_N$. The trace of $|\rho - \rho_N|$ in that basis

$$||\rho - \rho_N||_1 = tr(|\rho - \rho_N|) = \sum_{k=0}^{N} |p_k - q_k^{(N)}| + \sum_{k=N+1}^{+\infty} |p_k|$$

$$= \frac{1 - \sum_{j=0}^{N} p_j}{\sum_{j=0}^{N} p_j} \sum_{k=0}^{N} p_k + \sum_{k=N+1}^{+\infty} p_k \to 0 \quad \text{as } N \to +\infty .$$

The limit exists and vanishes because $p_n > 0$ and $\sum_{n=1}^{+\infty} p_n = 1$.

We will show Γ is continuous for $||\ ||_1$, and conclude. First of all extend Γ from $\mathfrak{S}(\mathsf{H})$ to positive trace-class operators on H, by defining:

$$\Gamma_1(A) := tr(A)\Gamma\left(\frac{1}{trA}A\right), \quad \Gamma_1(0) := 0$$

where $A \in \mathfrak{B}_1(\mathsf{H})$, $A \geq 0$ (so $tr(A) > 0$ if $A \neq 0$). It follows that $\Gamma_1(A) \in \mathfrak{B}_1(\mathsf{H})$, $\Gamma_1(A) \geq 0$, and:

$$\Gamma_1(\alpha A) = \alpha\Gamma_1(A) \quad \text{if } \alpha \geq 0, \text{ and} \quad tr(\Gamma_1(A)) = tr(A) .$$

Since Γ preserves convexity, it is not hard to see

$$\Gamma_1(A+B) = \Gamma_1(A) + \Gamma_1(B) .$$

To conclude extend Γ_1 to self-adjoint trace-class operators:

$$\Gamma_2(A) := \Gamma_1(A_+) - \Gamma_1(A_-) ,$$

where $A_- := -\int_{(-\infty,0)} x dP^{(A)}(x)$ and $A_+ := \int_{[0,+\infty)} x dP^{(A)}(x)$. Observe $A_+ - A_- = A$ and $|A| = A_+ + A_-$ by definition, since $P^{(A)}$ is the PVM of A.

If $A \in \mathfrak{B}_1(\mathsf{H})$ is self-adjoint, then $\Gamma_2(A)$ belongs to $\mathfrak{B}_1(\mathsf{H})$ and is self-adjoint. Moreover:

$$||\Gamma_2(A)||_1 \leq ||\Gamma_1(A_+)||_1 + ||\Gamma_1(A_-)||_1 = tr(A_+) + tr(A_-) = ||A||_1 \,.$$

Therefore Γ_2 is continuous for $||\ ||_1$, and so also its restriction $\Gamma : \mathfrak{S}(\mathsf{H}) \to \mathfrak{S}(\mathsf{H})$ is. □

Altogether we have proved the existence of U unitary, or antiunitary, satisfying $\gamma(\rho) = U\rho U^{-1}$ for any $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$. This finishes part (a).
(c) Since $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}(\rho) = U\rho U^{-1}$, we conclude $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}$ is a Wigner symmetry. In particular, the operator U satisfying (a) (in this theorem) satisfies (a) in Wigner's theorem for $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}$. By Wigner's theorem the character (unitary or antiunitary) of U from part (a) is fixed by $\gamma\restriction_{\mathfrak{S}_p(\mathsf{H}_{Sk})}$.
(d) If δ is a Wigner symmetry, by Wigner's theorem U exists, unitary or antiunitary, such that $\delta(\rho) = U\rho U^{-1}$ for any pure state. U defines the Kadison symmetry $\gamma^{(\delta)}(\rho) = U\rho U^{-1}$ extending δ to the whole space of states. Let us prove uniqueness. If two Kadison symmetries γ, γ', associated to the unitary or antiunitary U, U', coincide on $\mathfrak{S}_p(\mathsf{H}_{Sk})$, then the Wigner symmetries $\delta^{(U)} = U \cdot U^{-1}$, $\delta^{(U')} = U' \cdot U'^{-1}$ are the same. By Wigner's theorem U and U' are both either unitary or antiunitary, and $U = \chi U'$ with $|\chi| = 1$. Therefore

$$\gamma(\rho) = U\rho U^{-1} = \chi U' \rho U'^{-1} \chi^{-1} = \chi\chi^{-1} U' \rho U'^{-1} = U' \rho U'^{-1} = \gamma'(\rho)$$

for every $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$, so $\gamma = \gamma'$. The proof of Kadison's theorem is finished. □

Form the last part of the proof of part (a) we can extract yet another fact, interesting by its own means.

Proposition 12.15. *Let γ be a Wigner (or Kadison) automorphism of the complex Hilbert space H, and denote by $\mathfrak{B}_1(\mathsf{H})_{\mathbb{R}} \subset \mathfrak{B}_1(\mathsf{H})$ the real space of trace-class self-adjoint operators.*
There exist a unique linear operator $\gamma_2 : \mathfrak{B}_1(\mathsf{H})_{\mathbb{R}} \to \mathfrak{B}_1(\mathsf{H})_{\mathbb{R}}$, continuous for the norm $||\ ||_1$ of $\mathfrak{B}_1(\mathsf{H})$, that restricts to γ on $\mathfrak{S}_p(\mathsf{H})$ (or on $\mathfrak{S}(\mathsf{H})$, respectively). More precisely

$$||\gamma_2(A)||_1 \leq ||A||_1 \quad \textit{for every } A \in \mathfrak{B}_1(\mathsf{H})_{\mathbb{R}}.$$

Proof. The existence of a Kadison automorphisms is contained in Lemma 12.14, where we proved the existence of Γ_2 (above called γ_2) given Γ (γ above). Uniqueness follows from the construction, in Lemma 12.14, that led from Γ to Γ_2. For Wigner automorphisms the proof follows from the Kadison case, by statement (d) in Kadison's theorem. □

12.1.6 The dual action of symmetries on observables

The theorems of Wigner and Kadison enable us to define in a very elementary manner a (dual) *action* of a symmetry *on the observables* of the physical system.

Consider a physical system S described on the (separable, complex) Hilbert space H_S. For simplicity we shall consider the case of one sector only, as the generalisation to several coherent sectors is immediate. We know the set $\mathfrak{P}(\mathsf{H}_S)$ of elementary observables on S is described by orthogonal projectors on H. Observables on S are PVMs built with these projectors, i.e. self-adjoint (in general unbounded) operators associated to the PVMs.

Suppose $\gamma : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$ is a symmetry associated with the (anti)unitary operator U, up to phase. We define its **dual action** $\gamma^* : \mathfrak{P}(\mathsf{H}_S) \to \mathfrak{P}(\mathsf{H}_S)$ on the lattice of projectors by:

$$\gamma^*(P) := U^{-1}PU\,, \quad P \in \mathfrak{P}(\mathsf{H}_S) \tag{12.15}$$

(the arbitrary phase affecting U being irrelevant). A duality identity holds:

$$tr(\rho\gamma^*(P)) = tr(\gamma(\rho)P)\,. \tag{12.16}$$

This follows $\gamma(\rho) = U\rho U^{-1}$ by Kadison's theorem and the fact that (when computing traces if U is antiunitary) antiunitary operators preserve bases.

The mapping $\gamma^* : \mathfrak{P}(\mathsf{H}_S) \to \mathfrak{P}(\mathsf{H}_S)$ not only transforms orthogonal projectors into orthogonal projectors, but also preserves orthocomplemented, σ-complete bounded lattices. For example, the orthogonal projectors P, Q of $\mathfrak{P}(\mathsf{H}_S)$ commute if and only if $\gamma^*(P)$ commutes with $\gamma^*(Q)$. In that case $\gamma^*(P \vee Q) = \gamma^*(P) \vee \gamma^*(Q)$, and so on. If $A : D(A) \to \mathsf{H}$ is self-adjoint on H with spectral measure $P^{(A)} \subset \mathfrak{P}(\mathsf{H}_S)$, then $U^{-1}AU : U^{-1}D(A) \to \mathsf{H}_S$ is self-adjoint with spectral measure $\gamma^*\left(P^{(A)}\right)$ (see Exercises 9.1 if unitary, 12.6 if antiunitary). This fact allows to extend the action of γ^* to *every observable* in agreement with the spectral decomposition: just define, for a self-adjoint $A : D(A) \to \mathsf{H}_S$ representing an observable of S:

$$\gamma^*(A) := U^{-1}AU\,. \tag{12.17}$$

The physical meaning of $\gamma^*(A)$ is the following. When we define a Kadison symmetry γ, we are prescribing an experimental procedure under which *the system S* should be transformed. Mathematically speaking the action on states is described precisely by $\gamma : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$. The action γ^* on observables, instead, represents operative procedures *on measuring instruments* which, intuitively, correspond and generalise *passive transformations of the coordinates*. Better said, *the procedure is such that if we act by γ act on the system, or by γ^* on the instrument, we obtain the same result (expectation values, variances, outcome frequencies) as when we take the measurements.*

For instance, $\langle\gamma^*(A)\rangle_\rho$ and $\langle A\rangle_{\gamma(\rho)}$ are equal expectation values:

$$\langle\gamma^*(A)\rangle_\rho = tr(\gamma^*(A)\rho) = tr\left(U^{-1}AU\rho\right) = tr\left(AU\rho U^{-1}\right) = tr(A\gamma(\rho)) = \langle A\rangle_{\gamma(\rho)}\,.$$

This is, in practice, the content of the duality equation (12.16). The result is equivalent to saying that *the action of γ on the system can be neutralised, concerning measurement readings on the system, by the simultaneous action of* $(\gamma^*)^{-1}$ *on the instruments*.

Remarks 12.16. From the experimental point of view it is not obvious that a transformation acting on the system can be cancelled by a simultaneous action on the measuring instrument. Symmetries, *à la* Kadison or Wigner, have this property. ■

Examples 12.17. (1) Consider a spin-zero quantum particle described on $\mathbb{R}^3$, thought of as rest space of an inertial frame system with given positively-oriented orthonormal coordinates. From Chapter 11 we know the particle's Hilbert space is $L^2(\mathbb{R}^3, dx)$. Pure states are thus determined, up to arbitrary phases, by *wavefunctions*, i.e. by vectors $\psi \in L^2(\mathbb{R}^3, dx)$ such that $\int_{\mathbb{R}^3} |\psi(\mathbf{x})|^2 dx = 1$.

We want to explain how the isometries of $\mathbb{R}^3$ determine Wigner symmetries (hence Kadison symmetries), because under them *Lebesgue's measure is invariant.*

The notions of group theory used in the sequel will be summarised at a later stage (elementary facts are present in the book's appendix). Denote by $IO(3)$ the **isometry group of** $\mathbb{R}^3$, which is the *semidirect product* (see the appendix) of $O(3)$ with the *Abelian group of translations* $\mathbb{R}^3$. In practice, every element $\Gamma \in IO(3)$ is a pair $\Gamma = (R, \mathbf{t})$ acting on $\mathbb{R}^3$ as follows: $\Gamma(\mathbf{x}) := \mathbf{t} + R\mathbf{x}$. The composition law of $IO(3)$ is:

$$(\mathbf{t}', R') \circ (\mathbf{t}, R) = (\mathbf{t}' + R'\mathbf{t}, R'R) \quad \text{hence} \quad (\mathbf{t}, R)^{-1} = (-R^{-1}\mathbf{t}, R^{-1})\,.$$

Let $\Gamma : \mathbb{R}^3 \to \mathbb{R}^3$ belong to $IO(3)$, so Γ could in particular be: a translation along an axis $\mathbf{t}$, $\Gamma : \mathbb{R}^3 \ni \mathbf{x} \mapsto \mathbf{x} + \mathbf{t}$, a rotation of $O(3)$ about the origin $\mathbb{R}^3 \ni \mathbf{x} \mapsto R\mathbf{x}$ (including rotations with negative determinant in $O(3)$), or a combinations of the two. So we can define a transformation of square-integrable maps:

$$(U_\Gamma \psi)(\mathbf{x}) := \psi\left(\Gamma^{-1}\mathbf{x}\right), \quad \psi \in L^2(\mathbb{R}^3, dx). \tag{12.18}$$

The operator U is clearly linear, surjective (every isometry Γ of $\mathbb{R}^3$ is bijective) and isometric, as the Jacobian matrix J of an isometry has determinant ± 1:

$$||U_\Gamma \psi||^2 = \int_{\mathbb{R}^3} \left|\psi\left(\Gamma^{-1}\mathbf{x}\right)\right|^2 dx = \int_{\mathbb{R}^3} \left|\psi\left(\mathbf{x}'\right)\right|^2 |detJ| dx' = \int_{\mathbb{R}^3} \left|\psi\left(\mathbf{x}'\right)\right|^2 dx' = ||\psi||^2\,.$$

The transformation γ_Γ induced by the unitary operator U_Γ on states (pure or mixed) is a symmetry (Wigner or Kadison, respectively), which naturally represents the action of the isometry Γ on S given by the particle examined.

The map $IO(3) \ni \Gamma \mapsto U_\Gamma$ satisfies

$$U_{id} = I\,, \quad U_\Gamma U_{\Gamma'} = U_{\Gamma \circ \Gamma'}\,, \quad \Gamma, \Gamma' \in IO(3)$$

where *id* is the identity of $IO(3)$, because of (12.18). Thus $IO(3) \ni \Gamma \mapsto U_\Gamma$ preserves the group structure (in particular $U_{\Gamma^{-1}} = (U_\Gamma)^{-1}$); as such it is *a representation* of the group $IO(3)$ by unitary operators. We will discuss these representations in the next section.

Take now a PVM on $\mathbb{R}^3$, denoted $P^{(X)}$, that coincides with the *joint spectral measure* (see Theorem 9.16) of the three position operators:

$$(P_E^{(X)}\psi)(\mathbf{x}) = \chi_E(\mathbf{x})\psi(\mathbf{x})\,, \quad \psi \in L^2(\mathbb{R}^3, dx).$$

It is easy to prove the position operators arise by integrating the corresponding functions in this PVM:

$$X_i = \int_{\mathbb{R}^3} x_i dP^{(X)}(x) \quad i = 1,2,3.$$

Directly from definition (12.18) the **imprimitivity condition** holds:

$$U_\Gamma P_E^{(X)} U_\Gamma^{-1} = P_{\Gamma(E)}^{(X)}\,. \tag{12.19}$$

In fact for a generic map $\psi \in L^2(\mathbb{R}^3, dx)$:

$$\left(U_\Gamma P_E^{(X)} U_\Gamma^{-1}\psi\right)(\mathbf{x}) = \chi_E\left(\Gamma^{-1}(\mathbf{x})\right)\psi\left(\left(\Gamma\left(\Gamma^{-1}(\mathbf{x})\right)\right)\right) = \chi_{\Gamma(E)}(\mathbf{x})\psi(\mathbf{x})$$

$$= \left(P_{\Gamma(E)}^{(X)}\psi\right)(\mathbf{x})\,.$$

Equation (12.19) follows since ψ is arbitrary. Note that the imprimitivity equation can be written equivalently in terms of the dual action of the Kadison symmetry:

$$\gamma_\Gamma^*\left(P_E^{(X)}\right) = P_{\Gamma^{-1}(E)}^{(X)}\,.$$

In general a **system of imprimitivity** on X according to Mackey is given by: (i) a PVM P on the separable complex Hilbert space H for the Borel σ-algebra of the metrisable space X (that admits a metric making it complete and separable), (ii) a second-countable, locally compact group G acting on X so that the action[2] $\mathsf{G} \times X \ni (g,x) \mapsto gx \in X$ is measurable, (iii) a unitary representation $\mathsf{G} \ni g \mapsto V_g \in \mathfrak{B}(\mathsf{H})$ that is strongly continuous and satisfies the imprimitivity condition:

$$V_g P_E V_g^{-1} = P_{g(E)} \quad \text{for any } E \in \mathscr{B}(X),\, g \in \mathsf{G}\,.$$

The imprimitivity system is said **transitive** when the action of G on X is transitive, i.e. such that any pair $x_1, x_2 \in X$ can be transformed into one another $x_2 = gx_1$ by some $g \in \mathsf{G}$. Unitary representations of G for any imprimitivity system, up to unitary equivalence, are all determined by the famous **Imprimitivity theorem of Mackey**, which we shall not be concerned with (see for instance [Jau73]).

We have verified that $P^{(X)}$, $IO(3)$, U form a *transitive imprimitivity system on* $\mathbb{R}^3$ (we did not check the topological requests, which hold if we endow $IO(3)$ with the natural structure of a matrix subgroup of the Lie group $GL(4)$). Transitivity is obvious from elementary geometry.

[2] The map $(g,x) \mapsto gx$ is customarily taken so that $g'(gx) = (g'g)x$ and $ex = x$ for every $g, g' \in \mathsf{G}$, $x \in X$, where $e \in \mathsf{G}$ is the neutral element.

The action of γ^*_Γ on position operators can be obtained by direct computation, in analogy to the imprimitivity condition, or using the latter to integrate the spectral measure. Let $\mathbf{X} = (X_1, X_2, X_3)$ be the column vector of the X_1, X_2, X_3 restricted to the common invariant Schwartz domain $\mathscr{S}(\mathbb{R}^3)$, where the operators are essentially self-adjoint. Then

$$\gamma^*_\Gamma(\mathbf{X}) = U_\Gamma^{-1}\mathbf{X}U_\Gamma = R\mathbf{X} + \mathbf{t}I\,, \tag{12.20}$$

and in particular, considering pure translations:

$$\gamma^*_{(\mathbf{t},I)}(\mathbf{X}) = U^{-1}_{(\mathbf{t},I)}\mathbf{X}U_{(\mathbf{t},I)} = \mathbf{X} + \mathbf{t}I\,, \tag{12.21}$$

and pure rotations:

$$\gamma^*_{(\mathbf{0},R)}(\mathbf{X}) = U^{-1}_{(\mathbf{0},R)}\mathbf{X}U_{(\mathbf{0},R)} = R\mathbf{X}\,. \tag{12.22}$$

The element $(\mathbf{0}, -I) \in IO(3)$ defines the reflection about the origin. The unitary representation $\mathscr{P} := U_{(\mathbf{0},-I)}$, and the associated Wigner (Kadison) symmetry $\gamma_{\mathscr{P}}$, are called **parity inversion**. Not so precisely one often calls $(\mathbf{0}, -I)$ parity inversion. Easily, $\mathscr{P}^* = \mathscr{P}$ (so $\mathscr{P}\mathscr{P} = I$ as $\mathscr{P}^{-1} = \mathscr{P}^*$). Therefore the inversion of parity admits an associated observable, called **parity**, with two possible eigenvalues ± 1. We must emphasise that the unitary operator representing $(\mathbf{0}, -I)$ is actually defined, as usual, up to phase, so the observable $\mathscr{P}$ associated to the parity symmetry corresponds to a specific choice of phase. There are two possibilities for the phase, since also $-\mathscr{P}$ is an observable representing the inversion of parity.

(2) Consider the system of the previous example, and let us study it via the *momentum representation*. Using the Fourier-Plancherel transform, in other terms, we identify H and $L^2(\mathbb{R}^3, dk)$, so that the three momentum observables (the components of momentum in the orthonormal Cartesian coordinates of the inertial frame) are represented by the multiplication operators:

$$(P_i\widetilde{\psi})(\mathbf{k}) = \hbar k_i \widetilde{\psi}(\mathbf{k})\,,$$

as we saw in Chapter 11. We indicate by $\widetilde{\psi} = \widehat{\mathscr{F}}(\psi)$ the Fourier-Plancherel transform of $\psi \in L^2(\mathbb{R}^3, dx)$. An extremely interesting symmetry in physics is the **time reversal** $\gamma_{\mathscr{T}}$, described by antiunitary operators (later we will see why). This symmetry corresponds to flipping the sign of time, but also changing sign to particles' velocities and thus to their momentum. The antiunitary operator $\widetilde{\mathscr{T}}$ describing time reversal can be chosen (uniquely, up to phase) thus:

$$\left(\widetilde{\mathscr{T}}\widetilde{\psi}\right)(\mathbf{k}) := \overline{\widetilde{\psi}(-\mathbf{k})}\,, \quad \widetilde{\psi} \in L^2(\mathbb{R}^3, dk). \tag{12.23}$$

In contrast to $\mathscr{P}$ in the previous example, any chosen phase for $\widetilde{\mathscr{T}}$ maintains $\widetilde{\mathscr{T}}\widetilde{\mathscr{T}} = I$ because $\widetilde{\mathscr{T}}$ is antiunitary. Nonetheless, $\widetilde{\mathscr{T}}$ is not an observable since the operator is not linear. Reverting to the *position representation* with the chosen phase it can be proved that the symmetry $\gamma_{\mathscr{T}}$ is associated to an antiunitary operator

$$\mathscr{T} := \widehat{\mathscr{F}}^{-1}\widetilde{\mathscr{T}}\widehat{\mathscr{F}}$$

($\widehat{\mathscr{F}}$ is the Fourier-Plancherel transform as in Chapter 11), such that

$$(\mathscr{T}\psi)(\mathbf{x}) := \overline{\psi(\mathbf{x})}, \quad \psi \in L^2(\mathbb{R}^3, dx). \tag{12.24}$$

(3) Consider a particle having electric charge represented by the observable Q with discrete spectrum made by eigenvalues ± 1. Fix an inertial system of reference $\mathscr{I}$, with orthonormal Cartesian coordinates for which the rest space is $\mathbb{R}^3$. Then the system's Hilbert space is

$$\mathsf{H} = \mathbb{C}^2 \otimes L^2(\mathbb{R}^3, dx) \equiv L^2(\mathbb{R}^3, dx) \oplus (\mathbb{R}^3, dx),$$

where $\oplus$ denotes orthogonal sum. The canonical isomorphism between the above spaces (cf. Example 10.27(2) as well) descends from the fact that any $\Psi \in \mathbb{C}^2 \otimes L^2(\mathbb{R}^3, dx)$ can be written:

$$\Psi = |+\rangle \otimes \psi_+ \; + \; |-\rangle \otimes \psi_- ,$$

where $\{|+\rangle, |-\rangle\}$ is the canonical basis of $\mathbb{C}^2$ made by eigenvectors of the Pauli matrix σ_3 (cf. (12.10)) with eigenvalues $+1$ and -1 respectively. Hence the isomorphism reads:

$$L^2(\mathbb{R}^3, dx) \oplus (\mathbb{R}^3, dx) \ni (\psi_+, \psi_-) \mapsto |+\rangle \otimes \psi_+ \; + \; |-\rangle \otimes \psi_- \in \mathbb{C}^2 \otimes L^2(\mathbb{R}^3, dx).$$

It preserves the Hilbert structure (the inner product) if we view $L^2(\mathbb{R}^3, dx) \oplus (\mathbb{R}^3, dx)$ as an orthogonal sum. The charge observable can be thought of as the Pauli matrix σ_3 in $\mathbb{C}^2$, so on the complete space

$$Q = \sigma_3 \otimes I,$$

where I is the identity on $L^2(\mathbb{R}^3, dx)$. The superselection rule of the charge, in this simple situation, requires the space split in two coherent sectors $\mathsf{H} = \mathsf{H}_+ \oplus \mathsf{H}_-$, where $\mathsf{H}_\pm$ are the ± 1-eigenspaces of Q. By construction, the coherent decomposition coincides exactly with the natural:

$$\mathsf{H} = L^2(\mathbb{R}^3, dx) \oplus L^2(\mathbb{R}^3, dx).$$

Referring to the latter, admissible pure states are only those determined by vectors $(\psi, 0)$ or $(0, \psi)$, with $\psi \in L^2(\mathbb{R}^3, dx)$. Therefore the symmetry $\gamma_{\mathscr{C}_+}$, called **conjugation of the charge from the sector** H_+ **to the sector** H_- is represented by the unitary operator $\mathscr{C} : \mathsf{H}_+ \to \mathsf{H}_-$:

$$\mathscr{C}_+ : (\psi, 0) \mapsto (0, \psi), \quad \psi \in L^2(\mathbb{R}^3, dx). \tag{12.25}$$

The symmetry $\gamma_{\mathscr{C}_-}$, called **conjugation of the charge from the sector** H_- **to the sector** H_+ is similar:

$$\mathscr{C}_- : (0, \phi) \mapsto (\phi, 0), \quad \phi \in L^2(\mathbb{R}^3, dx). \tag{12.26}$$

Notice that $\mathscr{C}_-$ is the inverse of $\mathscr{C}_+$. Eventually, we define the Wigner symmetry called **conjugation of the charge**, that acts on the entire Hilbert space (respecting sectors) and restricts to the two above on the relative coherent spaces:

$$\mathscr{C} := \mathscr{C}_+ \oplus \mathscr{C}_- .$$

By construction $\mathscr{C}\mathscr{C} = I$, so $\mathscr{C} = \mathscr{C}^*$ and $\mathscr{I}$ is self-adjoint. Moreover

$$\mathscr{C}^* Q \mathscr{C} = -Q \quad . \tag{12.27}$$

■

12.2 Introduction to symmetry groups

This section is devoted to elementary topics from the theory of projective representations, applied to quantum symmetry groups. Given the vastitude and relevance of the material, the reader is encouraged to refer to the exhaustive treatise [BaRa86] for details.

12.2.1 Projective and projective unitary representations

Suppose we look at a group G (with product $\cdot$ and neutral element e) as a group of transformations acting on a physical system S, described on the Hilbert space H_S. For simplicity we assume H_S is the only coherent sector. Suppose, further, each transformation $g \in \mathsf{G}$ is associated to a symmetry γ_g, which we can then see as a Kadison (or Wigner) automorphism. We have already met this setup in Example 12.5(1), where G was the isometry group of the three-dimensional rest space of an inertial frame and S was the particle with no charge nor spin. Kadison automorphisms from $\mathfrak{S}(\mathsf{H}_S)$ to itself clearly form a group under the composition of maps. Hence the idea takes shape that there is a *representation of* G *in terms of Kadison automorphisms*: these should describe the group action of G on the quantum states of S. In other words we can suppose the map $\mathsf{G} \ni g \mapsto \gamma_g$ is a *group homomorphism* from G to the group of invertible maps on $\mathfrak{S}(\mathsf{H}_S)$:

$$\gamma_{g \cdot g'} = \gamma_g \circ \gamma_{g'} , \quad \gamma_e = id , \quad \gamma_{g^{-1}} = \gamma_g^{-1} , \quad g, g' \in \mathsf{G},$$

where id is the identity automorphism. Actually, the last condition is unnecessary because it follows from the former two by uniqueness of inverses. We also expect, as happens in the majority of concrete physical cases, the representation $\mathsf{G} \ni g \mapsto \gamma_g$ to be *faithful*, which means the homomorphism $\mathsf{G} \ni g \mapsto \gamma_g$ is injective. This is very often the case in physics.

Definition 12.18. *Consider a quantum system S described on the Hilbert space H_S. Let G be a group with an injective homomorphism (a faithful representation) $\mathsf{G} \ni g \mapsto \gamma_g$ defined by Wigner automorphisms $\gamma_g : \mathfrak{S}_p(\mathsf{H}_S) \to \mathfrak{S}_p(\mathsf{H}_S)$. Then G is called a* **symmetry group** *of S, and $\mathsf{G} \ni g \mapsto \gamma_g$ is its* **projective representation on** $\mathfrak{S}_p(\mathsf{H}_S)$.

Remark 12.19. **(1)** Referring only to Wigner symmetries is not restrictive since Kadison's theorem (in our formulation) warrants every Wigner automorphism γ_g extends, uniquely, to a Kadison automorphism $\gamma'_g : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$. It is straightforward that $\mathsf{G} \ni g \mapsto \gamma'_g$ is an injective homomorphism, i.e. a faithful representation of G by Kadison automorphisms. Conversely, every faithful G-representation by Kadison automorphisms determines a unique faithful G-representation by Wigner automorphisms, by restriction to $\mathfrak{S}_p(\mathsf{H}_S)$. In the sequel, despite mentioning Wigner symmetries most of the times, we will think the representation $\mathsf{G} \ni g \mapsto \gamma_g$ as given by Wigner or Kadison automorphisms, according to what will suit us best.
(2) The name *projective representation* is appropriate because $\mathfrak{S}_p(\mathsf{H}_s)$ is a *projective space*, as we saw in Chapter 7, and the map $\gamma_g : \mathfrak{S}_p(\mathsf{H}_S) \to \mathfrak{S}_p(\mathsf{H}_S)$ is well defined.
(3) Since the homomorphism $\mathsf{G} \ni g \mapsto \gamma_g$ is explicitly wanted injective, we can equivalently take, as group of symmetries, the set of automorphisms γ_g, with $g \in \mathsf{G}$, equipped with the natural group structure coming from map composition. Indeed, this group is isomorphic to G by construction. ■

Now here is an interesting issue. Suppose we have a symmetry group and a projective representation $\mathsf{G} \ni g \mapsto \gamma_g$. The map $\mathsf{G} \mapsto \gamma_g$ is certainly a representation, but *not a linear representation*, because the $\gamma_g : \mathfrak{S}_p(\mathsf{H}_S) \to \mathfrak{S}_p(\mathsf{H}_S)$ are not linear maps. Yet since to every automorphism γ_g corresponds a (linear) unitary or antiunitary operator $U_g : \mathsf{H}_S \to \mathsf{H}_S$ that satisfies $\gamma_g(\rho) = U_g \rho U_g^{-1}$ for any $\rho \in \mathfrak{S}_p(\mathsf{H}_S)$, a natural question arises: can $\mathsf{G} \ni g \mapsto U_g$ be an *(anti)linear representation* of G? Can it be given, in other terms, by *(anti)linear* (unitary and/or antiunitary) operators in $\mathfrak{B}(\mathsf{H})$? We are equivalently asking whether $\mathsf{G} \ni g \mapsto U_g$ is a *group homomorphism*, i.e. if it preserves the group structure:

$$U_{g\cdot g'} = U_g U_{g'}\,, \quad U_e = I\,, \quad U_{g^{-1}} = U_g^{-1} \quad \text{for any } g, g' \in \mathsf{G}, \tag{12.28}$$

where $I : \mathsf{H}_S \to \mathsf{H}_S$ is the identity operator. The matter is relevant from a technical point to view, too: the profusion of results available on linear representations over (Hilbert) spaces can be used to study symmetry groups of quantum systems. The answer to the preceding questions is typically *negative*, because the condition $U_{g\cdot g'} = U_g U_{g'}$ in general fails. Namely, as $\gamma_g \circ \gamma_{g'} = \gamma_{gg'}$, we have

$$U_g U_{g'} \rho (U_g U_{g'})^{-1} = U_{g\cdot g'} \rho U_{g\cdot g'}^{-1} \quad \text{for any } \rho \in \mathfrak{S}(\mathsf{H}_S).$$

Consequently:

$$(U_{g\cdot g'})^{-1} U_g U_{g'} \rho (U_g U_{g'})^{-1} U_{g\cdot g'} = \rho\,, \quad \rho \in \mathfrak{S}_p(\mathsf{H}_S).$$

This means that if $\rho = \psi(\psi|\ )$ then $(U_{g\cdot g'})^{-1} U_g U_{g'} \psi$ and ψ can differ by a phase factor, at most. The phase cannot depend upon ψ (the proof is the same as the uniqueness part in Wigner's theorem), but it may still depend on g and g'. This result is sharp – the best possible – because the U_g themselves are defined up to phase. Overall, if the (unitary or antiunitary) U_g are associated to a projective representation of a certain symmetry group, the condition $U_{g\cdot g'} = U_g U_{g'}$ weakens, in the general case, to

$$U_g U_{g'} = \omega(g, g') U_{g\cdot g'}\,, \quad g, g' \in \mathsf{G},$$

where $\omega(g,g') \in \mathbb{C}$, $|\omega(g,g')| = 1$, are complex numbers depending on how the U_g are associated to the automorphisms γ_g, but in any case respecting the theorems of Wigner and Kadison. Therefore if $U(1)$ is the group of unit complex numbers, $\omega(g,g') \in U(1)$. In particular, setting $g = g' = e$, the above implicit definition of $\omega(g,g')$ tells

$$U_e = \omega(e,e)I\,,$$

hence $U_e \rho U_e^{-1} = \rho$ as it should be.

It is not at all obvious that one can redefine phases so to obtain $\omega(g,g') = 1$ for every $g, g' \in \mathsf{G}$.

Remarks 12.20. Henceforth we will work with unitary operators, and drop the anti-unitary case. The explanation is put off until the end of the section. ■

The functions $\mathsf{G}\times\mathsf{G} \ni (g,g') \mapsto \omega(g,g') \in U(1)$ are not totally arbitrary, because associativity holds:

$$(U_g U_{g'})U_{g''} = U_g(U_{g'}U_{g''})\,.$$

A computation shows that the above is equivalent to:

$$\omega(g,g')\omega(g\cdot g', g'') = \omega(g, g'\cdot g'')\omega(g',g'')\,. \tag{12.29}$$

In turn, the latter implies:

$$\omega(g,e) = \omega(e,g)\,, \quad \omega(g,e) = \omega(g_1,e)\,, \quad \omega(g,g^{-1}) = \omega(g^{-1},g)\,, \quad g, g_1 \in \mathsf{G}\,, \tag{12.30}$$

(e being the neutral element of G, so $U_e = \omega(e,e)I$). The next definition transcends the physical meaning of the objects involved.

Definition 12.21. *Let* G *be a group and* H *a (complex) Hilbert space.*

(a) *A* **projective unitary representation of** G **on** H *is a map*

$$\mathsf{G} \ni g \mapsto U_g \in \mathfrak{B}(\mathsf{H}) \tag{12.31}$$

such that: U_g *are unitary operators, and the* **multiplier (operators)** *of the representation*

$$\omega(g,g') := U^{-1}_{g\cdot g'}U_g U_{g'}\,, \quad g,g' \in \mathsf{G}, \tag{12.32}$$

belong to $U(1)$ *for any* $g, g' \in \mathsf{G}$ *(hence (12.29) holds).*

The projective representation on $\mathfrak{S}_p(\mathsf{H})$ *given by (with obvious notation)*

$$\mathsf{G} \ni g \mapsto U_g \cdot U_g^*$$

is **induced** *by the projective unitary representation (12.31).*

The projective unitary representation (12.31) is called **irreducible** *if there is no proper closed subspace* $H_0 \subset \mathsf{H}$ *such that* $U_g(H_0) \subset H_0$ *for every* $g \in \mathsf{G}$.

Given H, H' *Hilbert spaces (possibly equal), two projective unitary representations*

$$\mathsf{G} \ni g \mapsto U_g \in \mathfrak{B}(\mathsf{H}) \textit{ and } \mathsf{G} \ni g \mapsto U'_g \in \mathfrak{B}(\mathsf{H}')$$

are said **(unitarily) equivalent** *if there exist a unitary operator* $S : \mathsf{H} \to \mathsf{H}'$ *and a map* $\chi : \mathsf{G} \ni g \mapsto \chi(g) \in U(1)$ *satisfying:*

$$\chi(g) S U_g S^{-1} = U'_g\,, \quad g \in \mathsf{G}. \tag{12.33}$$

(b) *A group homomorphism*

$$\mathsf{G} \ni g \mapsto U_g \in \mathfrak{B}(\mathsf{H}) \tag{12.34}$$

mapping elements of G *to unitary operators on* H *is a* **(proper) unitary representation of** G **on** H. *(That is to say, a unitary representation is a projective unitary representation whose multipliers equal* 1.*)*

The unitary representation (12.34) is **irreducible** *if no proper closed subspace* $H_0 \subset \mathsf{H}$ *exists such that* $U_g(H_0) \subset H_0$ *for every* $g \in \mathsf{G}$.

Given H, H' *Hilbert spaces (possibly equal), two unitary representations*

$$\mathsf{G} \ni g \mapsto U_g \in \mathfrak{B}(\mathsf{H}) \text{ and } \mathsf{G} \ni g \mapsto U'_g \in \mathfrak{B}(\mathsf{H}')$$

are **(unitarily) equivalent** *if there is a unitary operator* $S : \mathsf{H} \to \mathsf{H}'$ *such that*

$$S U_g S^{-1} = U'_g \quad \text{for every } g \in \mathsf{G}. \tag{12.35}$$

Important remark. The reader should now be able to see the difference between *projective* representations, *projective unitary* representations and *unitary* representations. The first type act on $\mathfrak{S}_p(\mathsf{H}_S)$ or $\mathfrak{S}(\mathsf{H}_S)$ representing symmetry groups, and do not involve choices without physical meaning. The other two kinds act on H_S, induce projective representations, but are affected by arbitrary choices on the phases of the unitary operators by which they act. ∎

Remark 12.22. (1) The notion of *equivalence* of two projective unitary representations is transitive, symmetric and reflexive, so it is an honest *equivalence relation* on the space of projective unitary representations of a given group on a given Hilbert space. If G is a symmetry group for the physical system S, described on the Hilbert space H_S, projective representations of G on $\mathfrak{S}_p(\mathsf{H}_S)$ are patently in one-to-one correspondence with equivalence classes of projective unitary representations of G.
(2) The property that a projective unitary representation $\mathsf{G} \ni g \mapsto U_g$ be *equivalent* to a unitary representation is actually a property of the coset of the projective unitary representation: it means that the equivalence class contains a unitary representative. When talking about symmetry groups of a quantum system, that is a feature of the projective representation on $\mathfrak{S}(\mathsf{H}_S)$ corresponding to the class.
(3) The property that a projective unitary representation $\mathsf{G} \ni g \mapsto U_g$ be *irreducible* is a property of the coset of the projective unitary representation: if one representative in the equivalence class is irreducible, all other elements are irreducible, as is clear from the definitions. Irreducible representations are important in that every representation can be constructed as direct sum or *direct integral* of irreducible representations [BaRa86].
(4) Given a symmetry group G with projective representation $\mathsf{G} \ni g \mapsto \gamma_g$, a function is automatically defined, namely $\mathsf{G} \ni g \mapsto \gamma_g^*$, that represents the group action on

observables, in the sense of Chapter 12.1.6. The definition holds beyond the particular choice of projective unitary representation of the theory on the system's Hilbert space: the phases we have to fix to pass from γ_g to the U_g cancel out when we transfer the action to observables: $\gamma_g^*(A) = U_g^{-1} A U_g$. Note that $\mathsf{G} \ni g \mapsto \gamma_g^*$ *does not define a left* G*-representation*; it is easy to see, from the definition of γ_g^*, that:

$$\gamma_g^* \gamma_{g'}^* = \gamma_{g' \cdot g}^*$$

by construction, and not $\gamma_g^* \gamma_{g'}^* = \gamma_{g \cdot g'}^*$. Furthermore, $\gamma_e^* = id$ and $\gamma_{g^{-1}}^* = (\gamma_g^*)^{-1}$. The function $\mathsf{G} \ni g \mapsto \gamma_g^*$ is a *right representation* of G, provided we endow observables with the structure of a vector space.

Definition 12.23. *Let* G *have neutral element e. A (linear)* **right representation** *of* G *on a vector space V is a map* $\mathsf{G} \ni g \mapsto \alpha_g \in GL(V)$ *such that*

$$\alpha_a \alpha_b = \alpha_{b \cdot a}\,, \quad \alpha_e = id\,, \quad (\alpha_c)^{-1} = \alpha_{c^{-1}}$$

for any $a, b, c \in \mathsf{G}$. ■

Here is a more concrete way of asking whether a projective representation $\mathsf{G} \ni g \mapsto \gamma_g$ of a symmetry group G can be described, on H_S, by a unitary representation of G. Inside the equivalence class of projective unitary representations associated to $\mathsf{G} \ni g \mapsto \gamma_g$ we fix an arbitrary representative (the ensuing discussion does not depend on this element, by remark (2) above) and consider its multipliers.

Thus we reduce to decide whether there might be a map $\chi : \mathsf{G} \ni g \mapsto \chi(g) \in \mathbb{C}$ such that $|\chi(g)| = 1$ and:

$$\omega(g, g') = \frac{\chi(g \cdot g')}{\chi(g)\chi(g')} \quad \text{for any } g, g' \in \mathsf{G}\,. \tag{12.36}$$

Proof: if said map χ exists, inserting it on the left in (12.33) renders the multipliers of $\mathsf{G} \ni g \mapsto U'_g$ trivial by (12.36). Conversely, if the multipliers of $\mathsf{G} \ni g \mapsto U'_g$ are trivial, the χ in (12.33) solves (12.36).

There are many strategies to tackle and solve the existence problem of χ [BaRa86], and one can see there exist groups, e.g. the *Lorentz group* and the *Poincaré group*, whose projective representations are described by unitary representations on the Hilbert space of the system. At the same time there are groups, like the *Galilean group*, whose (non-trivial) projective representations cannot be given by unitary representations, but only by projective unitary representations, and for which the multipliers cannot be suppressed by smart choices of the phases.

There is a colossal literature on the topic, and irreducible projective unitary representations of the groups of interest in physics (especially Lie groups) have been studied and classified [BaRa86].

12.2.2 Projective unitary representations are unitary or antiunitary

Let us return to the unitary vs. antiunitary issue of the operators U_g. Suppose to have a symmetry group with projective representation $\mathsf{G} \ni g \mapsto \gamma_g$. To an automorphism γ_g corresponds either a unitary operator or an antiunitary one $U_g : \mathsf{H}_S \to \mathsf{H}_S$ satisfying $\gamma_g(\rho) = U_g \rho U_g^{-1}$ for every $\rho \in \mathfrak{S}_p(\mathsf{H}_S)$, by Wigner's theorem. Are there criteria to decide whether the U_g are all unitary, all antiunitary, or maybe both depending on $g \in \mathsf{G}$? If U_g and $U_{g'}$ were antiunitary, the constraint $U_g U_{g'} = \chi(g,g')U_{g\cdot g'}$ would force $U_{g\cdot g'}$ to be unitary. Therefore representations of groups with more than two elements, all made by antiunitary operators (identity apart, which is always unitary) cannot exist. The hybrid case when a certain number of antiunitary operators (more than one) are present is anyway non-trivial, due to constraints such as the aforementioned one. In this respect the next proposition shows that the group G might impose the operators be all unitary.

Proposition 12.24. *Let H be a complex Hilbert space and G a group. Suppose each $g \in \mathsf{G}$ is the product of certain $g_1, g_2, \ldots, g_n \in \mathsf{G}$ (dependent on g, with n dependent on g) that admit a square root (there exist $r_k \in \mathsf{G}$ such that $g_k = r_k \cdot r_k$ for every $k = 1, \ldots, n$). Then for every projective representation $\mathsf{G} \ni g \mapsto \gamma_g$, the elements γ_g can only be associated to unitary operators by Wigner's theorem (or Kadison's).*

Proof. The proof is obvious, for $U_{r_k}U_{r_k}$ is linear even when U_{r_k} is antilinear, and from $U_{g_k} = \chi(r_k, r_k)U_{r_k}U_{r_k}$ follows U_{g_k} is linear, so also U_g must be linear. □

The following result is important in the applications, especially the case $n = 1$.

Proposition 12.25. *In relationship to Proposition 12.24, the projective representations of the additive group $\mathsf{G} = \mathbb{R}^n$ are associated to unitary operators only.*

Proof. If $\mathbf{t} \in \mathbb{R}^n$ then $\mathbf{t} = \mathbf{t}/2 + \mathbf{t}/2$, and the rest is a corollary of propostion 12.24. □

We shall see later that Proposition 12.24 is automatic when we assume G is a connected Lie group, so antiunitary operators appear only with discrete groups or when changing connected components. So in the sequel we will deal with the case in which the U_g are all unitary.

12.2.3 Central extensions and quantum group associated to a symmetry group

The approach in [BaRa86] allows to study all possible projective unitary representations of a certain group, by looking at them as restrictions of unitary representations of a larger group, a *central extension* of the starting one. The recipe, apparently overcomplicated, is technically useful (also to detect possible unitary representations of G) in that it lets us use the specific toolbox of the much developed theory of unitary representations (of the extension). Let us briefly explain the basic idea of the procedure, postponing the fundamental example where G is the Galilean group; the reader might skip this section at first and return to it when needed.

Take any group G and one projective unitary representation $\mathsf{G} \ni g \mapsto U_g$ on a Hilbert space H with multipliers ω. Define another group $\widehat{\mathsf{G}}_\omega$ consisting of pairs $(\chi, g) \in U(1) \times \mathsf{G}$ with product:

$$(\chi,g) \circ (\chi',g') = \left(\chi\chi'\omega(g,g') \,,\, g \cdot g'\right) \,, \quad (\chi,g),(\chi',g') \in U(1) \times \mathsf{G}.$$

The reader can check the definition is well posed, just owing to the fact ω satisfies (12.29), and that it produces a group with neutral element $(\omega(e,e)^{-1}, e)$, e being the neutral element of G (remember (12.30)) and $(\chi,g)^{-1} = (\chi^{-1}\omega(e,e)^{-1}\omega(g,g^{-1})^{-1}, g^{-1})$ the inverse. The following definition disregards the origin of the function ω, and only requires equation (12.29) to be valid.

Definition 12.26. *Let* G *be a group and* $\omega : \mathsf{G} \times \mathsf{G} \to U(1)$ *any function satisfying (12.29). The group* $\widehat{\mathsf{G}}_\omega = U(1) \times \mathsf{G}$ *with product*

$$(\chi,g) \circ (\chi',g) = \left(\chi\chi'\omega(g,g') \,,\, g \cdot g'\right) \,, \quad (\chi,g),(\chi',g') \in U(1) \times \mathsf{G},$$

is a **central extension** *of* G *by* $U(1)$ *with* **multiplier function** ω. *The injective homomorphism* $U(1) \ni \chi \mapsto (\chi, e) \in \widehat{\mathsf{G}}_\omega$ *and the surjective homomorphism* $\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto g \in \mathsf{G}$ *are respectively called* **canonical injection** *and* **canonical projection of the central extension**.

The names (see the appendix at the end of the book for a minimal dictionary of group theory) come about as follows: the canonical projection $\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto g \in \mathsf{G}$ is a surjective homomorphism, whose null space is the normal subgroup $\mathscr{N}$ (range of the canonical injection and isomorphic to $U(1)$) of pairs (χ, e) with $\chi \in U(1)$. $\mathscr{N}$ is contained in the *centre* of $\widehat{\mathsf{G}}_\omega$, as its elements commute with the whole $\widehat{\mathsf{G}}_\omega$ (in fact $\omega(e,g) = \omega(g,e)$). In practice the group G has been extended, to $\widehat{\mathsf{G}}_\omega$, by adding the kernel of the surjection $\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto g \in \mathsf{G}$), which is central. Notice that G is naturally identified with the quotient group $\widehat{\mathsf{G}}_\omega/\mathscr{N}$.

The procedure for obtaining all projective unitary representations $\mathsf{G} \ni g \mapsto U_g$ of G relies on three important points.

(1) If $\mathsf{G} \ni g \mapsto U_g$ has multiplier function ω, the map

$$\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto V_{(\chi,g)} := \chi U_g \,,$$

is a *unitary* $\widehat{\mathsf{G}}_\omega$-representation on H. In fact the operators $V_{(\chi,g)} : \mathsf{H} \to \mathsf{H}$ are unitary, so $V_{(\omega(e,e)^{-1},e)} = I$ and

$$V_{(\chi,g)}V_{(\chi',g')} = \chi U_g \chi' U_{g'} = \chi\chi'\omega(g,g')U_{g\cdot g'} = V_{(\chi,g)\circ(\chi',g')} \,.$$

(2) The initial representation arises from the unitary representation $\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto V_{(\chi,g)}$ *by restriction*: i.e., restricting the domain of V to elements $(1, g)$, $g \in \mathsf{G}$. We will say that the unitary representation V restricts to G in this case.

(3) Given any unitary representation

$$\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto V_{(\chi,g)}$$

of a central extension, the restriction to $\{1\} \times \mathsf{G}$, say $U_g := V_{(1,g)}$, is a projective unitary representation if and only if:

$$V_{(\chi,e)} = \chi\omega(e,e)I \quad \text{for every } \chi \in U(1). \tag{12.37}$$

In fact, $V_{(\chi,g)} = \chi U_g$ implies $V_{(\chi,e)} = \chi U_e = \chi\omega(e,e)I$ (for any projective unitary representation $\omega(e,e) := U_{ee}^{-1}U_eU_e = U_e$). From $(\chi, g) = (\chi\omega(e,e)^{-1}, e)(1,g)$, conversely, if (12.37) holds we can write

$$V(\chi,g) = V(\chi\omega(e,e)^{-1},e)V(1,g) = \chi V(1,g) =: \chi U_g\,.$$

So we have this proposition.

Proposition 12.27. *Every projective unitary representation of a group* G *is the restriction of a unitary representation of a suitable central extension* $\widehat{\mathsf{G}}_\omega$ *whose multiplier function satisfies (12.37).*

The extension procedure, especially when G is a Lie group, is extremely powerful. Using cohomology techniques it enables to catalogue *all* projective unitary representations that are continuous in some topology (and all unitary representations of a simply connected Lie group) starting from the Lie algebra of G [BaRa86]. We will return here at a later stage.

As a matter of fact we need not know *all* central extensions of G to classify projective unitary representations. It suffices to know central extensions whose multipliers *are non-equivalent*. Two multiplier functions on the same group, $\mathsf{G} \times \mathsf{G} \ni (g,g') \mapsto \omega(g,g') \in U(1)$ and $\mathsf{G} \times \mathsf{G} \ni (g,g') \mapsto \omega'(g,g') \in U(1)$, are called **equivalent** if there is a map $\chi : G \to U(1)$ such that

$$\omega(g,g') = \frac{\chi(g \cdot g')}{\chi(g)\chi(g')}\,\omega'(g,g')\,, \quad g,g' \in G.$$

If two projective unitary representations U, U' of G are equivalent, they are restrictions of unitary representations of central extensions $\widehat{\mathsf{G}}_\omega$, $\widehat{\mathsf{G}}_{\omega'}$ with equivalent multiplier functions ω, ω'. Hence, by knowing central extensions of G whose multipliers *are not equivalent* and their unitary representations, we actually know the equivalence classes of projective unitary representations of G, and so all projective unitary representations of G.

Further, if $\omega(e,e) \neq 1$ for a certain ω, using an equivalence transformation by a constant function χ we can reduce to the case $\omega(e,e) = 1$. Multipliers such that $\omega(e,e) = 1$ (whence $\omega(e,g) = \omega(g,e) = \omega(e,e) = 1$) are **normalised**. The central extension has neutral element $(1,e)$, and (12.37) reads

$$V_{(\chi,e)} = \chi I\,, \quad \chi \in U(1)\,. \tag{12.38}$$

Projective unitary representations arising thus satisfy $U_e = I$.

To finish, we make a few physical considerations on the meaning of $\widehat{\mathsf{G}}_\omega$, when there are no unitary representations of G, only projective unitary representations. Suppose having a symmetry group $\mathsf{G} \ni g \mapsto \gamma_g$ for the physical system S, hence a projective representation on $\mathfrak{S}(\mathsf{H}_S)$, that is *not* describable by means of a unitary representation. We can anyway choose phases arbitrarily and extend G to $\widehat{\mathsf{G}}_\omega$ using the multipliers found, and then take $\widehat{\mathsf{G}}_\omega$ as the *true* symmetry group of S. The latter admits in this way two representations. One from G itself:

$$\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto g \in \mathsf{G}\,,$$

that captures the *classical* action of the group. But there is also a *quantum* and *unitary* one:

$$\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto \chi U_g\,,$$

representing the group action on the states of the system (actually on the system's Hilbert space, and so on states, too.

In this light the group $\widehat{\mathsf{G}}_\omega$ is sometimes called the *quantum group* associated to G. Note, however, that a specific central extension $\widehat{\mathsf{G}}_\omega$ cannot be selected using the construction seen above, for which only projective representations given by automorphisms of Wigner or Kadison have a physical meaning. In order to choose among the various central extensions it is necessary to give a physical meaning to the single projective unitary representations of G, or to the unitary representations of the possible extensions $\widehat{\mathsf{G}}_\omega$. This can be done, by enriching G and turning it into a Lie group, as we will see. For the projective unitary representations of the Galilean group, multipliers have a straightforward meaning, for they are related to the *mass* of the physical system. This will be all the more clear after discussing Lie groups of symmetries.

12.2.4 Topological symmetry groups

We turn to *topological* symmetry groups and *Lie* groups of symmetries. Lie groups are a subclass of topological groups. The majority of quantum symmetry groups, with the notable exclusion of discrete symmetries (parity inversion and time reversal) in particular, are Lie groups. We will study in depth the additive Lie group $\mathbb{R}$, whose importance should not go amiss, both physically and technically.

Definition 12.28. *A* **topological group** *is a group* G *and a topological space at the same time, whose operations of product* $\mathsf{G} \times \mathsf{G} \ni (f, g) \mapsto f \cdot g \in \mathsf{G}$*, and inversion* $\mathsf{G} \ni g \mapsto g^{-1}$*, are continuous in the product topology of* $\mathsf{G} \times \mathsf{G}$ *and the topology of* G*, respectively.*

The theory of topological groups and their representations occupies a huge chapter of mathematics [NaSt82], and we shall be just concerned with a few very elementary results that befit our physical models. We present below some examples and properties of topological groups, with an eye to the *Haar measure*.

Examples 12.29. (1) The real **general linear group** $GL(n,\mathbb{R})$ and the complex general linear group $GL(n,\mathbb{C})$ of nonsingular $n \times n$ real and complex matrices, are (evident) topological groups, if we equip them with the topology induced by $\mathbb{R}^{n^2}$ and $\mathbb{C}^{n^2}$.

(2) Using the standard topology any subgroup of the above two is a topological group. For instance ($\eta := diag(-1,1,1,1)$):

the **unitary group** $U(n) = \{U \in GL(n,\mathbb{C}) \mid UU^* = I\}$;
the **special unitary group** $SU(n) := \{U \in U(n) \mid detU = 1\}$;[3]
the **orthogonal group** $O(n) := \{R \in GL(n,\mathbb{R}) \mid RR^t = I\}$;
the **special orthogonal group** $SO(n) := \{R \in O(n) \mid detR = 1\}$;
the **special linear group** $SL(n,\mathbb{R}) := \{A \in GL(n) \mid detA = 1\}$;
the **Lorentz group** $O(1,3) := \{\Lambda \in GL(4) \mid \Lambda\eta\Lambda^t = \eta\}$;
the **orthochronous Lorentz group** $O(1,3)\Uparrow := \{\Lambda \in O(1,3) \mid \Lambda_{11} > 0\}$;
the **special orthochronous Lorentz group** $SO(1,3)\Uparrow := \{\Lambda \in O(1,3)\Uparrow \mid det\Lambda > 0\}$.

The list (including $GL(n,\mathbb{R})$ and $GL(n,\mathbb{C})$) is made of *closed* subsets in $\mathbb{R}^{n^2}$, or $\mathbb{C}^{n^2}$ if matrices are complex. This from the definitions: just notice that by continuity every sequence in one of those groups converges in $\mathbb{R}^{n^2}$, or $\mathbb{C}^{n^2}$, to an element of the group.

The groups $O(n), SO(n), U(n), SU(n)$ (not the others) are *bounded*, and therefore *compact groups*. Boundedness follows from the definition and the Cauchy-Schwarz inequality. For example, for $U \in U(n)$ we have $\sum_{k=1}^n U_{ik}\overline{U_{jk}} = \delta_{ij}$ by definition of $U(n)$. Hence $\sum_{i,k=1}^n U_{ik}\overline{U_{ik}} = n$, so $\sum_{i,k=1}^n |U_{ik}|^2 = n$ and $U(n)$ is contained in the closed ball of radius $\sqrt{n}$ in $\mathbb{C}^{n^2}$.

(3) Some groups do not look like matrix groups, like the additive group $\mathbb{R}$. But it, too, just like the **isometry group of** $\mathbb{R}^n$, $IO(n)$, built as in Example 12.17(1) replacing $O(3)$ with $O(n)$, can be realised by matrices. For $IO(n)$, one representation is by real $(n+1)\times(n+1)$ matrices

$$M((R,\mathbf{t})) := \begin{bmatrix} 1 & \mathbf{0}^t \\ \mathbf{t} & R \end{bmatrix}, \quad \mathbf{t} \in \mathbb{R}^n, R \in O(n) \tag{12.39}$$

(a subgroup of the topological group $GL(n+1,\mathbb{R})$ with induced topology). $IO(n) \ni (R,\mathbf{t}) \mapsto M((R,\mathbf{t}))$ is an isomorphism. The additive group $\mathbb{R}^n$ arises by restriction, via the homeomorphism $\mathbb{R}^n \ni \mathbf{t} \mapsto M((I,\mathbf{t}))$ ($\mathbb{R}^n$ with standard structure).

The *Galilean group* (Chapter 12.3.3) and the *Poincaré* group are topological groups, built analogously via matrices.

(4) Yet there exist topological groups (even Lie groups) that cannot be viewed as matrix groups, an example being the *universal covering* (Definition 12.44) of $SL(2,\mathbb{R})$.

(5) *Locally compact Hausdorff* groups, like $\mathbb{R}^n$ (an Abelian group for the sum of $\mathbb{R}^n$), $GL(n,\mathbb{R})$, $GL(n,\mathbb{C})$ and subgroups thereof, admit a special *regular* Borel measure, called the *Haar measure*. The Haar measure is defined up to a factor and is translation-invariant under the group.

Its definition is contaned in the following classical theorem, proved by Weil in full generality [Hal69, BaRa86], of which we provide no proof. Recall that if G has product $\circ$, the **left** and **right orbits of** $B \subset \mathsf{G}$ **under** $g \in \mathsf{G}$ are:

$$gB := \{g \circ b \mid b \in B\} \quad \text{and} \quad Bg := \{b \circ g \mid b \in B\}$$

[3] The word *special*, for matrix groups, indicates *determinant equal* 1, and is often denoted by putting an S before the group's name.

respectively. A positive σ-additive measure μ on the Borel σ-algebra $\mathscr{B}(\mathsf{G})$ of the locally compact Hausdorff group G is called **left-invariant** if

$$\mu(gB) = \mu(B) \quad \text{for any } B \in \mathscr{B}(\mathsf{G}),\, g \in \mathsf{G},$$

and **right-invariant** if

$$\mu(Bg) = \mu(B) \quad \text{for any } B \in \mathscr{B}(\mathsf{G}),\, g \in \mathsf{G}.$$

Note that $\mu(gB)$, $\mu(Bg)$ are well defined. Since the multiplication by $h \in \mathsf{G}$ on the left, $\mathsf{G} \ni b \mapsto f_h(b) := h \circ b$, is continuous, and since $gB = (f_{g^{-1}})^{-1}(B)$, we have $gB \in \mathscr{B}(\mathsf{G})$ if $B \in \mathsf{G}$. Similarly $Bg \in \mathscr{B}(\mathsf{G})$ if $B \in \mathsf{G}$.

Theorem 12.30. *Let* G *be a locally compact Hausdorff group. Up to a positive factor, there exists a unique positive* σ*-additive measure* μ_G *– the* **left-invariant Haar measure of** G *– on the Borel* σ*-algebra* $\mathscr{B}(\mathsf{G})$*, that is regular (*$\mu_\mathsf{G}(B) = \inf\{\mu_\mathsf{G}(U) \,|\, B \subset U, U \text{ open}\}$ *and* $\mu_\mathsf{G}(B) = \sup\{\mu_\mathsf{G}(K) \,|\, K \subset B, K \text{ compact}\}$*) and such that:*

(i) μ_G *is left-invariant;*
(ii) $\mu_\mathsf{G}(B) > 0$ *if* $B \in \mathscr{B}(\mathsf{G}) \setminus \{\varnothing\}$ *is open,* $\mu(K) < +\infty$ *if* $K \in \mathscr{B}(\mathsf{G})$ *is compact*[4].

Furthermore, if G *is compact,* μ_G *is also right-invariant because* $\mu_\mathsf{G}(E) = \mu_\mathsf{G}(E^{-1})$*, where* $E^{-1} := \{g^{-1} \,|\, g \in E\}$ *for any* $E \in \mathscr{B}(\mathsf{G})$*.*

A similar result for right-invariant measures defines, up to the usual positive factor, the **right-invariant Haar measure** ν_G. This in general is different (factor apart) from the (left-invariant) Haar measure μ_G: they coincide in case G is compact, by the theorem, because $\nu(E) := \mu(E^{-1})$ is right-invariant on $\mathscr{B}(\mathsf{G})$ if μ is left-invariant on $\mathscr{B}(\mathsf{G})$. If so, one speaks of the **bi-invariant Haar measure**.

The Abelian group $(\mathbb{R}, +)$ has the Lebesgue measure as Haar measure: the left- and right-invariant Haar measures coincide. The group $GL(n,\mathbb{R})$ (and its subgroups of (2)) has Haar measure:

$$\mu_{GL(n,\mathbb{R})}(B) := \int_B |\det g(x_{11}, \ldots, x_{nn})|^{-n} dx;$$

where $g \in GL(n,\mathbb{R})$ has entries x_{ij} seen as coordinates of $\mathbb{R}^{n^2}$, and dx is the Lebesgue measure on $\mathbb{R}^{n^2}$. ■

At this point we want to specialise the notion of symmetry group to topological groups, which entails imposing topological constraints on the associated projective representation.

Suppose we have a symmetry group $\mathsf{G} \ni g \mapsto \gamma_g$ for the quantum system S described on the Hilbert space H_S. If G is a topological group, we expect the homomorphism $g \mapsto \gamma_g$ to be continuous in some sense. This means choosing a topology on the space of maps γ_g, which we may think of as either Kadison automorphisms or Wigner automorphisms. In the sequel we adopt Wigner's point of view. We give first the definition and then explain it mathematically and physically.

[4] NB: some authors require the last condition in the definition of regular Borel measure.

Definition 12.31. *Consider a quantum system S described on the Hilbert space* H_S*. Let* G *be a topological group with a projective representation* $\mathsf{G} \ni g \mapsto \gamma_g$ *on* H*, such that*

$$\lim_{g \to g_0} tr(\rho_1 \gamma_g(\rho_2)) = tr\left(\rho_1 \gamma_{g_0}(\rho_2)\right), \quad g_0 \in \mathsf{G},\ \rho_1, \rho_2 \in \mathfrak{S}_p(\mathsf{H}_S).$$

Then G *is a* **topological** *group of symmetries for S, and* $\mathsf{G} \ni g \mapsto \gamma_g$ *is a* **continuous** *projective representation on* $\mathfrak{S}_p(\mathsf{H}_S)$*.*

Physically this is reasonable, for it says the transition probability between two pure states, one of which is the image of the action of the symmetry group, is a *continuous* map for the action. In Wigner's quantum-symmetry setup, this is more than sound.

But the definition is also natural in mathematical terms, as we explain now. Let $\mathfrak{B}_1(\mathsf{H}_S)_\mathbb{R}$ be the real vector space of self-adjoint, trace-class operators with norm $|| \ ||_1$. By Proposition 12.15 every Wigner automorphism γ_g is the restriction to $\mathfrak{S}_p(\mathsf{H}_S)$ of a *linear operator* $(\gamma_2)_g : \mathfrak{B}_1(\mathsf{H}_S)_\mathbb{R} \to \mathfrak{B}_1(\mathsf{H}_S)_\mathbb{R}$, determined by γ_g and continuous for $|| \ ||_1$. Consider then the mapping $\Gamma : \mathsf{G} \ni g \mapsto (\gamma_2)_g$. Putting the strong topology on $\mathfrak{B}_1(\mathsf{H}_S)_\mathbb{R}$ and the standard one on the domain, we will say Γ is continuous if for any $\rho \in \mathfrak{B}_1(\mathsf{H}_S)$, $g_0 \in \mathsf{G}$:

$$\lim_{g \to g_0} ||(\gamma_2)_g(\rho) - (\gamma_2)_{g_0}(\rho)||_1 = 0\,.$$

Now restrict to $\mathfrak{S}_p(\mathsf{H}_S)$ with the induced topology, thus reverting to the representation $\mathsf{G} \ni g \mapsto \gamma_g$ in terms of Wigner automorphisms. Then $\mathsf{G} \ni g \mapsto \gamma_g$ is continuous if, for any $\rho \in \mathfrak{S}_p(\mathsf{H}_S)$, $g_0 \in \mathsf{G}$:

$$\lim_{g \to g_0} ||\gamma_g(\rho) - \gamma_{g_0}(\rho)||_1 = 0\,.$$

This notion of continuity is, apparently, different from that of Definition 12.31. The next proposition tells they are indeed the same.

Proposition 12.32. *Let* H *be a complex Hilbert space and* $||\rho||_1 = tr(|\rho|)$ *the norm on trace-class operators* $\mathfrak{S}(\mathsf{H}_S)$*. Then restricting to pure states:*

$$||\rho - \rho'||_1 = 2\sqrt{1 - (tr(\rho\rho'))^2}\,, \quad \rho, \rho' \in \mathfrak{S}_p(\mathsf{H}). \tag{12.40}$$

Equivalently:

$$||\psi(\psi|\ ) - \psi'(\psi'|\ )||_1 = 2\sqrt{1 - |(\psi|\psi')|^2}\,, \quad \psi, \psi' \in \mathsf{H},\ ||\psi|| = ||\psi'|| = 1. \tag{12.41}$$

Therefore $\mathfrak{S}_p(\mathsf{H})$ *is a metric space with distance function:*

$$d(\rho, \rho') := 2\sqrt{1 - (tr(\rho\rho'))^2}\,, \quad \rho, \rho' \in \mathfrak{S}_p(\mathsf{H}).$$

Proof. The first assertion is a trivial transcription of the second one, and the third is obvious once the first two are proven, by general properties of norms. To prove the second statement it suffices to construct an orthonormal basis ψ_1, ψ_2 of the span of ψ, ψ', assuming $\psi_1 = \psi$ and decomposing ψ' in that basis. Then with $b := (\psi'|\psi_2)$:

$$\psi(\psi|\ ) - \psi'(\psi'|\ ) = -|b|\psi_1(\psi_1|\ ) + |b|\psi_2(\psi_2|\ ) .$$

This is the spectral decomposition of $\rho - \rho'$, so

$$|\rho' - \rho| = |b|\psi_1(\psi_1|\ ) + |b|\psi_2(\psi_2|\ ) = |b|I ,$$

and

$$||\psi(\psi|\ ) - \psi'(\psi'|\ )||_1 = tr(|b|I) = 2|b| = 2\sqrt{1 - |(\psi'|\psi_1)|^2} = 2\sqrt{1 - |(\psi'|\psi)|^2}$$

because $1 = ||\psi'||^2 = |(\psi'|\psi_1)|^2 + |(\psi'|\psi_2)|^2$. □

Remarks 12.33. The last claim of the proposition is quite interesting, for $\mathfrak{S}_p(\mathsf{H})$ is not a normed space, not even being a vector space. Nonetheless, it is a metric space (Definition 2.78) and the distance has a meaning: it is related to the probability amplitude. ■

Mathematics and physics eventually meet in the next result.

Proposition 12.34. *Consider a quantum system S described on the Hilbert space* H_S*. Let* G *be a topological group. A projective representation* $\mathsf{G} \ni g \mapsto \gamma_g$ *on* H *is continuous (Definition 12.31), so* G *is a topological symmetry group for S, if and only if it is continuous with respect to:*

(i) *the topology of* G *on the domain;*
(ii) *the strong topology on the codomain, restricted to* $\mathfrak{S}_p(\mathsf{H}_S)$*.*

That is:

$$\lim_{g \to g_0} ||\gamma_g(\rho) - \gamma_{g_0}(\rho)||_1 = 0 , \quad \rho \in \mathfrak{S}_p(\mathsf{H}_S),\ g_0 \in \mathsf{G}. \tag{12.42}$$

Proof. Equation (12.40) implies

$$||\gamma_g(\rho) - \gamma_{g_0}(\rho)||_1 = 2\sqrt{1 - tr\left(\gamma_g(\rho)\gamma_{g_0}(\rho)\right)} .$$

If $\mathsf{G} \ni g \mapsto \gamma_g$ is continuous for Definition 12.31 then $\lim_{g\to g_0} tr\left(\gamma_g(\rho)\gamma_{g_0}(\rho)\right) = tr\left(\gamma_{g_0}(\rho)\gamma_{g_0}(\rho)\right) = 1$. Substituting above yields (12.42). Conversely, from (12.40), the trace's invariance under cyclic permutations gives

$$tr\left(\gamma_{g_0}(\rho)\gamma_g(\rho)\right) = 1 - \frac{1}{4}||\gamma_g(\rho) - \gamma_{g_0}(\rho)||_1^2 .$$

Set $\rho_1 := \gamma_{g_0}(\rho)$ (without loss of generality, as γ_{g_0} is onto) and $\rho_2 := \rho$. Then:

$$\lim_{g \to g_0} tr(\rho_1 \gamma_g(\rho_2)) = 1 - \frac{1}{4}\lim_{g \to g_0} ||\gamma_{g_0}(\rho) - \gamma_g(\rho)||_1^2 = 1 - \frac{1}{4}||\gamma_{g_0}(\rho) - \gamma_{g_0}(\rho)||_1^2$$
$$= tr\left(\rho_1 \gamma_{g_0}(\rho_2)\right) .$$

Hence (12.42) implies continuity for Definition 12.31. □

12.2.5 Strongly continuous projective unitary representations

Consider a physical system S described on the Hilbert space H_S, a topological symmetry group G and a projective representation $\mathsf{G} \ni g \mapsto \gamma_g$. Let us associate to G a projective unitary representation $\mathsf{G} \ni g \mapsto U_g$, in the sense $\gamma_g(\rho) = U_g \rho U_g^{-1}$, for every pure state $\rho \in \mathfrak{S}_p(\mathsf{H}_S)$ of the system and every element $g \in \mathsf{G}$. Clearly if $\mathsf{G} \ni g \mapsto U_g$ is strongly continuous, then $\mathsf{G} \ni g \mapsto \gamma_g$ is a continuous projective representation: Definition 12.31 holds, in fact, since if $\rho_i = \psi_i(\psi_i|\ )$, $i = 1,2$:

$$tr\left(\rho_1 U_g \rho_2 U_g^*\right) = |(\psi_1|U_g\psi_2)| \to |(\psi_1|U_{g_0}\psi_2)| = tr\left(\rho_1 U_{g_0} \rho_2 U_{g_0}^*\right) \text{ as } g \to g_0\,.$$

Here is an interesting problem: knowing $\mathsf{G} \ni g \mapsto \gamma_g$ is continuous, establish if the phases of the unitary operators U_g can be fixed so to obtain a projective unitary representation that is *strongly continuous*. I.e. we would like

$$U_g\psi \to U_{g_0}\psi \quad \text{as } g \to g_0 \text{ for any } \psi \in \mathsf{H}.$$

In its general form the question is very hard, although Wigner gave a *local* answer. We will show that given a topological symmetry group G and a continuous projective representation $\mathsf{G} \ni g \mapsto \gamma_g$, it is possible to fix the multipliers ω so to make the projective unitary representation $\mathsf{G} \ni g \mapsto U_g$ become strongly continuous on a neighbourhood of the neutral element of G. Moreover, also the multipliers will be continuous on that neighbourhood. This local result is not usually global. We will prove that for $\mathsf{G} = \mathbb{R}$ the result holds everywhere on the group and multipliers can be fixed to 1, so that the representation is simultaneously *unitary* and *strongly continuous*. The consequences in physics reach far: we will be able to justify the postulate of time evolution, and explain the relationship between the existence of symmetries and the presence of preserved quantities as the system S evolves in time: a quantum formulation, in other words, of Nöther's theorem. All this later, because now we focus on mathematical aspects.

Proposition 12.35. *Consider a quantum system S described on the Hilbert space H_S, and let G be a topological group with continuous projective representation $\gamma : \mathsf{G} \ni g \mapsto \gamma_g$. There exist an open neighbourhood $A \subset \mathsf{G}$ of $e \in \mathsf{G}$ and a projective unitary representation associated to γ, $\mathsf{G} \ni g \mapsto U_g$, that is strongly continuous on A. The multipliers*

$$\omega(g,g') = \left(U_{g\cdot g'}\right)^{-1} U_g U_{g'}\,, \quad g,g' \in \mathsf{G}$$

define a continuous map on an open neighbourhood A' of e with $A' \cdot A' \subset A$.

Proof. Fix $\phi \in \mathsf{H}$, $||\phi|| = 1$. As $\mathsf{G} \ni g \mapsto tr(\phi(\phi|\ )\gamma_g(\phi(\phi|\ )))$ is continuous and equals 1 for $g = e$, there is an open neighbourhood A_0 of e where

$$tr(\phi(\phi|\ )\gamma_g(\phi(\phi|\ ))) \neq 0\,.$$

Represent γ by a projective unitary representation V, arbitrarily chosen, which we have by Wigner's theorem. Around A_0, then:

$$0 \neq tr(\phi(\phi|\ )\gamma_g(\phi(\phi|\ ))) = (\phi|V_g\phi)\,.$$

Define ($(\phi|V_g\phi) \neq 0$ guarantees it is possible):

$$\chi_g := \frac{\overline{(\phi|V_g\phi)}}{|(\phi|V_g\phi)|}$$

and pass to a new projective unitary representation U:

$$U_g := \chi_g V_g\,, \quad \text{if } g \in A_0 \text{ and } \quad U_g := V_g\,, \quad \text{if } g \notin A_0\,.$$

Then on A_0:

$$0 < \frac{|(\phi|V_g\phi)|^2}{|(\phi|V_g\phi)|} = (\phi|U_g\phi)$$

so

$$0 < (\phi|U_g\phi) = |(\phi|U_g\phi)| = tr(\phi(\phi|\ )\gamma_g(\phi(\phi|\ ))) \quad \text{for any } g \in A_0\,. \tag{12.43}$$

Equation (12.43) has two consequences on some open neighbourhood A of e, $A \subset A_0$:

$$U_e = 1\,, \quad \text{and} \quad U_{g^{-1}} = U_g^{-1}\,, \quad g \in A. \tag{12.44}$$

In fact, $U_e = \chi I$ for some $\chi \in U(1)$, so $(\phi|U_e\phi) = \chi(\phi|\phi) = \chi$. As $(\phi|U_e\phi) > 0$, we can only have $\chi = 1$. As for the second property, $U_{g^{-1}} = \chi'_g U_g^{-1}$ for some $\chi'_g \in U(1)$. Since $g \mapsto g^{-1}$ is continuous and from $e^{-1} = e$ there is an open neighbourhood of e, $A \subset A_0$, for which $g^{-1} \in A_0$ if $g \in A$. Working on A,

$$0 < (\phi|U_{g^{-1}}\phi) = \chi'_g(\phi|U_g^{-1}\phi) = \chi'_g(\phi|U_g^*\phi) = \chi'_g(U_g\phi|\phi) = \chi'_g(\phi|U_g\phi)$$

because $(\phi|U_g\phi)$ is real so $(\phi|U_g\phi) = (U_g\phi|\phi)$. Since $(\phi|U_g\phi) > 0$, necessarily $\chi'_g = 1$. This proves (12.44).

Fix a unit vector $\psi \in \mathsf{H}$, possibly distinct from the above ϕ. By continuity of γ, as in Definition 12.31 with $\rho_1 = U_s\psi(U_s\psi|\ )$ and $\rho_2 = \psi(\psi|\ )$, we find

$$\lim_{r\to s} |(U_r\psi|U_s\psi)| = |(U_s\psi|U_s\psi)| = 1\,. \tag{12.45}$$

Choosing $\rho_1 = \phi(\phi|\ )$, $\rho_2 = \psi(\psi|\ )$ gives

$$\lim_{r\to s} |(\phi|U_r\psi)| = |(\phi|U_s\psi)|\,. \tag{12.46}$$

Substituting in the general identities produces

$$||U_s\psi - (U_r\psi|U_s\psi)U_r\psi||^2 = 1 - |(U_r\psi|U_s\psi)|^2\,, \tag{12.47}$$

so

$$\lim_{r\to s} (U_r\psi|U_s\psi)U_r\psi = U_s\psi \tag{12.48}$$

and in particular, for $\psi = \phi$:

$$\lim_{r\to s} (U_r\phi|U_s\phi)(\phi|U_r\phi) = (\phi|U_s\phi)\,. \tag{12.49}$$

On the other hand, our choice of ϕ and of the phase in U implies

$$\lim_{r\to s}(\phi|U_r\phi) = \lim_{r\to s}|(\phi|U_r\phi)| = |(\phi|U_s\phi)| = (\phi|U_s\phi)\,, \tag{12.50}$$

and so using (12.50) in (12.49) tells

$$\lim_{r\to s}(U_r\phi|U_s\phi) = 1\,. \tag{12.51}$$

Now, U_t is unitary, and for any $\psi \in \mathsf{H}$ (any $\psi = \phi$) we have

$$||U_r\psi - U_s\psi||^2 = 2 - 2Re(U_r\psi|U_s\psi)\,, \tag{12.52}$$

so (12.51) implies, for $r \in A$, that the map $r \mapsto U_r\phi$ is continuous, with the chosen ϕ. Therefore $r \mapsto (U_r)^{-1}\phi$ is continuous, since (12.49) holds when r is replaced by r^{-1} and s by s^{-1} ($g \mapsto g^{-1}$ is continuous, and $(U_r)^{-1} = U_{r^{-1}}$ by (12.44)). From (12.48) follows

$$\lim_{r\to s}(U_r\psi|U_s\psi)(\phi|U_r\psi) = (\phi|U_s\psi) \quad \text{i.e.} \quad \lim_{r\to s}(U_r\psi|U_s\psi)((U_r)^{-1}\phi|\psi) = (\phi|U_s\psi)\,.$$

As $((U_r)^{-1}\phi|\psi) \to ((U_s)^{-1}\phi|\psi) = (\phi|U_s\psi)$ by the continuity of $(U_s)^{-1}\phi$, necessarily $(U_r\psi|U_s\psi) \to 1$ as $r \to s$. Use this result in (12.52):

$$\lim_{r\to s}||U_r\psi - U_s\psi|| = 0\,. \tag{12.53}$$

But ψ is arbitrary, so $A \ni g \mapsto U_g$ is strongly continuous.

Now the second claim. From $U(e) = 1$ and $U_{g^{-1}} = U_g^{-1}$, on A we have

$$\omega(g,e) = \omega(e,g) = 1\,. \tag{12.54}$$

From

$$(U_r^{-1}\phi|U_s\phi) = \omega(r,s)^{-1}(\phi|U_{r\cdot s}\phi) \tag{12.55}$$

and $(\phi|U_{r\cdot s}\phi) > 0$ if $r\cdot s \in A$, we infer $(r,s) \mapsto \omega(r,s)^{-1}$ is continuous for $r,s,r\cdot s \in A$. Since the product of G is continuous if $e\cdot e = e$, there is a neighbourhood $A' \subset A$ of e where $r,s \in A'$ implies $r\cdot s \in A$. Taking A' small enough renders $A'\times A' \ni (r,s) \mapsto \omega(r,s) = \overline{\omega(r,s)^{-1}}$ continuous. □

12.2.6 A special case: the topological group $\mathbb{R}$

We prove in this section a very important theorem about continuous representations of the *additive group* $\mathbb{R}$ equipped with the standard topology. The result is crucial in physics, as we will have time to explain.

Theorem 12.36. *Let $\mathbb{R} \ni r \mapsto \gamma_r$ be a continuous projective representation of $\mathbb{R}$ on the Hilbert space H.*

(a) *There exists a strongly continuous one-parameter unitary group (Definition 9.22) $\mathbb{R} \ni r \mapsto W_r$ such that*

$$\gamma_r(\rho) = W_r\rho W_r^{-1} \quad \textit{for any } r \in \mathbb{R},\ \rho \in \mathfrak{S}_p(\mathsf{H}). \tag{12.56}$$

(b) *A second strongly continuous one-parameter unitary group $\mathbb{R} \ni r \mapsto U_r$ satisfies (12.56) (with U_r replacing W_r) if and only if there exists $c \in \mathbb{R}$ such that*

$$U_r = e^{-icr} W_r \quad \textit{for any } r \in \mathbb{R}.$$

(c) *There exists a self-adjoint operator $A : D(A) \to \mathsf{H}$ on H, unique up to additive constants, such that:*

$$\gamma_r(\rho) = e^{-irA} \rho e^{irA} \quad \textit{for any } r \in \mathbb{R},\ \rho \in \mathfrak{S}_p(\mathsf{H}).$$

Proof. (a) Let $[-b,b] \subset A$, $b > 0$, be an interval in the open neighbourhood of 0, say $A \subset \mathbb{R}$, satisfying Proposition 12.35 for $\mathsf{G} = \mathbb{R}$. Decompose $\mathbb{R}$ into the disjoint union of intervals $(na,(n+1)a]$, $n \in \mathbb{Z}$, with $a = b/2$. Any $r \in \mathbb{R}$ belongs to one interval only, so $r = n_r a + t_r$ for unique $t_r \in (0,a]$ and $n_r \in \mathbb{Z}$. Since $\gamma_x \gamma_y = \gamma_{x+y}$:

$$\gamma_r = \gamma_{n_r a + t_r} = (\gamma_a)^{n_r} \gamma_{t_r}\,.$$

Hence if $\mathbb{R} \ni r \mapsto U_r$ is the projective unitary representation of Proposition 12.35:

$$\gamma_r(\rho) = ((U_a)^{n_r} U_{t_r})\, \rho\, ((U_a)^{n_r} U_{t_r})^{-1}\,,$$

for every $\rho \in \mathfrak{S}_p(\mathsf{H}_S)$. For every $t \in (-a-\varepsilon, a+\varepsilon)$ and some $\varepsilon > 0$ the map $t \mapsto U_t$ is strongly continuous, so

$$\mathbb{R} \ni r \mapsto V_r$$

with $V_r := (U_a)^{n_r} U_{t_r}$, $n_r \in \mathbb{Z}$ and $t_r \in (0,a]$ as above, is strongly continuous too. The only discontinuities can occur at the endpoints. Consider then $r \in (na,(n+1)a]$ and let us verify V_r is continuous at na. With $r_- < na$, $r_+ > na$ we have

$$V_{r_-}\psi = (U_a)^{(n-1)} U_{t_{r_-}}\psi \quad \text{and} \quad V_{r_+}\psi = (U_a)^n U_{t_{r_+}}\psi\,.$$

As $(-a,a) \ni t \mapsto U_t \psi$ is continuous, by definition of V:

$$\lim_{r_- \to na^-} V_{r_-}\psi = V_{na}\psi\,.$$

Now we need to see the right and left limits coincide, i.e. that the limit of $(U_a)^{(n-1)} U_{t_{r_-}}\psi$, as $t_{r_-} \to a^-$, coincides with the limit of $(U_a)^n U_{t_{r_+}}\psi$ as $t_{r_+} \to 0^-$. We have

$$\lim_{t \to a^-} (U_a)^{n-1} U_t \psi = \lim_{t-a \to 0^-} (U_a)^{n-1} \omega(a,t-a)^{-1} U_a U_{t-a}\psi$$

$$= \lim_{t-a \to 0^-} \omega(a,t-a)^{-1} (U_a)^n U_{t-a}\psi = \lim_{\tau \to 0^-} \omega(a,\tau)^{-1} (U_a)^n U_\tau \psi\,.$$

By the previous proof $(0,a] \ni \tau \mapsto \omega(a,\tau)^{-1}$ is continuous, since $a, \tau, a+\tau \in A$ by construction. Moreover, $\chi(a,0) = 1$ from (12.54). We also know $(0,a] \ni t \mapsto U_t \psi$ is continuous, so:

$$\lim_{t \to a^-} (U_a)^{n-1} U_t \psi = \lim_{\tau \to 0^-} \omega(a,\tau)^{-1} (U_a)^n U_\tau \psi = \lim_{\tau \to 0^+} \omega(a,\tau)^{-1} (U_a)^n U_\tau \psi$$

$$= \lim_{t \to 0^+} (U_a)^n U_t \psi\,.$$

We have proved

$$V_{na}\psi = \lim_{r_- \to na^-} V_{r_-}\psi = \lim_{t_{r_-} \to a^-} (U_a)^{(n-1)} U_{t_{r_-}}\psi = \lim_{t_{r_+} \to 0^+} (U_a)^n U_{t_{r_+}}\psi = \lim_{r_+ \to na^+} V_{r_+}\psi ,$$

as required. Note $(V_r)^{-1} = (U_{t_r})^{-1}(U_a)^{-n_r} = U_{-t_r}(U_a)^{-n_r}$, where the second identity in (12.44) was used. In analogy to the proof for V_r, also $\mathbb{R} \ni r \mapsto (V_r)^{-1}$ is continuous in the strong topology.

We claim the multipliers of V can be set to 1: for this, first we will show they give a continuous map $\mathbb{R}^2 \ni (r,s) \mapsto \omega(r,s) \in U(1)$, using that $\mathbb{R} \ni t \mapsto V_t\psi$ and $\mathbb{R} \ni t \mapsto (V_t)^{-1}\psi$ are continuous. Then we will prove the latter function is equivalent to the constant map 1. By definition

$$\omega(r,s)V_{r+s} = V_r V_s .$$

Fix $r_0, s_0 \in \mathbb{R}$. There must exist $\psi, \phi \in \mathsf{H} \setminus \{0\}$ so that $(\psi|V_{r_0+s_0}\phi) \neq 0$, otherwise $V_{r_0+s_0}\phi = 0$ for every ϕ, which is impossible by hypothesis as V_t is unitary. By continuity there is a neighbourhood B of (r_0, s_0) such that $(r,s) \in B$ implies $(\psi|V_{r+s}\phi) \neq 0$. Then

$$\omega(r,s) = \frac{((V_r)^{-1}\psi|V_s\phi)}{(\psi|V_{r+s}\phi)} .$$

Hence $\mathbb{R}^2 \ni (r,s) \mapsto \omega(r,s) \in U(1)$ is continuous around (r_0,s_0), and so continuous on $\mathbb{R}^2$. We may write $\omega(r,s) = e^{-if(r,s)}$ for some function $f : \mathbb{R}^2 \to \mathbb{R}$. The continuous map ω can be thought of as valued in the unit circle $\mathbb{S}^1$, homeomorphic to $U(1)$. The map f can be chosen continuous (the fundamental group of $\mathbb{R}^2$ is trivial, so the lifting property of covering spaces holds, cf. [Ser94II, theorem 18.2]). Equation (12.29) now reads

$$f(s,t) - f(r+s,t) + f(r,s+t) - f(r,s) = 2\pi k_{r,s,t} \quad \text{for } k_{r,s,t} \in \mathbb{Z} .$$

Continuous functions map connected sets ($\mathbb{R}^3$) to connected sets (a subset of $2\pi\mathbb{Z}$ with induced standard topology), so the right-hand side is constant. But the left-hand side is zero for $r = s = t = 0$, so:

$$f(s,t) - f(r+s,t) + f(r,s+t) - f(r,s) = 0 \quad \text{for every } r,s,t \in \mathbb{R} . \tag{12.57}$$

Fix a C^1 map $g : \mathbb{R} \to \mathbb{R}$ with compact support such that:

$$\int_{\mathbb{R}} g(x)dx = 1 \quad \text{and} \quad \int_{\mathbb{R}} \frac{dg}{dx}\, dx = 0 . \tag{12.58}$$

Define the continuous function:

$$\chi(r) := e^{-ih(r)} \quad \text{where} \quad h(r) := -\int_0^r du \int_{\mathbb{R}} f(u,t)\, \frac{dg}{dt}\, dt - \int_{\mathbb{R}} f(r,t)g(t)dt .$$

The new representation $W_r := \chi(r)V_r$ has multiplier $\omega'(r,s) = \omega(r,s)\frac{\chi(r)\chi(s)}{\chi(r+s)}$, so

$$\omega(r,s)' = e^{-if'(r,s)} \quad \text{where} \quad f'(r,s) = f(r,s) - h(r+s) + h(r) + h(s) .$$

A moderately-involved computation on the right side, using h, (12.57), (12.58), and the easily proved

$$\int_0^{r+s} du\, F(u) - \int_0^r du\, F(u) - \int_0^s du\, F(u) = \int_0^s du\, (F(u+r) - F(u))$$

eventually gives $f'(r,s) = 0$, i.e. $\chi'(r,s) = 1$ for any $(r,s) \in \mathbb{R}^2$. This makes the projective unitary representation $\mathbb{R} \ni r \mapsto W_r$ actually unitary. Since $\mathbb{R} \ni x \mapsto \chi(x)$ is continuous by construction and $\mathbb{R} \ni r \mapsto V_r$ is strongly continuous, also $W = \chi V$ is strongly continuous. That is to say, $\mathbb{R} \ni r \mapsto W_r$ is a strongly continuous one-parameter unitary group satisfying (12.56), thus ending (a).

(b) If there is another strongly continuous one-parameter unitary group U representing γ:

$$U_{-r} W_r \psi = \chi(r)\psi\,, \quad \psi \in \mathsf{H}. \tag{12.59}$$

(We have already proved $\chi(r)$ and ψ are independent in similar situations.) Consequently $W_r = \chi(r)U_r$. Multiply by $W_s = \chi(s)U_s$, and use the additivity of W and U in the parameter:

$$W_{r+s} = \chi(r)\chi(s)U_{r+s} \quad \text{so} \quad U_{-(r+s)}W_{r+s} = \chi(r)\chi(s)I.$$

Comparing with $U_{-(r+s)}W_{r+s} = \chi(r+s)I$, produces

$$\chi(r+s) = \chi(r)\chi(s)\,. \tag{12.60}$$

Equation (12.59) has another corollary:

$$(U_r\phi|W_r\psi) = \chi(r)(\phi|\psi)\,.$$

By Stone's theorem (Theorem 9.29) we can write $U_t = e^{-itB}$, $W_t = e^{-itA}$ for self-adjoint operators defined on dense domains $D(A), D(B)$. Choose $\phi \in D(B)$, $\psi \in D(A)$ so that $(\phi|\psi) \neq 0$ (always possible by density). By Stone the first derivative of $\mathbb{R} \ni t \mapsto \chi(r)$ has to satisfy

$$\frac{d}{dt}(U_r\phi|W_r\psi) = \left(\frac{d}{dt}U_r\phi\middle| W_r\psi\right) + \left(U_r\phi\middle|\frac{d}{dt}\,W_r\psi\right),$$

hence it exists and equals

$$(-iBU_r\phi|W_r\psi) + (U_r\phi| - iAW_r\psi)\,.$$

Since the derivative of χ exists, and (12.60) holds:

$$\frac{d}{dx}\chi(x) = \lim_{h\to 0}\frac{1}{h}\,(\chi(x+h) - \chi(x)) = \chi(x)\lim_{h\to 0}\frac{1}{h}\,(\chi(h) - \chi(0)) = \chi(x)c\,.$$

Hence $\chi(x) = e^{icx}$ for some $c \in \mathbb{R}$ and then

$$W_x = e^{icx}U_x\,.$$

Conversely let W be as in (a) and fix $c \in \mathbb{R}$. A direct computation shows $U_x := e^{-icx}W_x$ is a strongly continuous one-parameter unitary group that represents γ.

(c) The strongly continuous one-parameter unitary group $\mathbb{R} \ni r \mapsto W_r$, built in (a), represents γ and has a self-adjoint generator A, by Stone's theorem. Therefore $W_r = e^{-irA}$. If $B : D(B) \to \mathsf{H}$ is another self-adjoint operator representing γ, its one-parameter group $U_t = e^{-itB}$ fulfills (b). Then there is $c \in \mathbb{R}$ such that $e^{-itA} = e^{-itc}e^{-itB}$. By Stone's theorem the left-hand side admits strong derivative at $t = 0$ on $D(A)$, and the derivative is $-iA$. Similarly, the right-hand side admits strong derivative at $t = 0$, at least on $D(B)$, which equals $-icI - iB$. Consequently $D(A) \subset D(B)$ and $A = (cI + B)\restriction_{D(A)}$. Note $cI + B$ is self-adjoint on $D(B)$. Since A is self-adjoint it does not have proper self-adjoint extensions, and then $D(A) = D(B)$ and $A = B + cI$. $\square$

Examples 12.37. **(1)** Consider Example 12.17(1). The physical system is a quantum particle with no spin, described on the Hilbert space $L^2(\mathbb{R}^3, dx)$ if we fix an inertial reference system and identify $\mathbb{R}^3$ with the rest space via orthonormal Cartesian coordinates.

The subgroup $ISO(3)$ of isometries of $\mathbb{R}^3$ consists of functions:

$$(\mathbf{t}, R) : \mathbb{R}^3 \ni \mathbf{x} \mapsto \mathbf{t} + R\mathbf{x}\,, \tag{12.61}$$

with $\mathbf{t} \in \mathbb{R}^3$, $R \in SO(3)$. Taking $R \in SO(3)$, as opposed to $R \in O(3)$ explains the 'S' in $ISO(3)$. As mentioned in Example 12.29(3) (about $IO(n)$ there, but the argument is the same), $ISO(3)$ is a matrix group. Consider 4×4 real matrices:

$$g_{(\mathbf{t},R)} := \begin{bmatrix} 1 & \mathbf{0}^t \\ \mathbf{t} & R \end{bmatrix}, \quad \mathbf{t} \in \mathbb{R}^n, R \in SO(3). \tag{12.62}$$

The topology is inherited from $GL(4, \mathbb{R})$ i.e. $\mathbb{R}^{16}$. The matrices $g_{(\mathbf{t},R)}$ correspond one-to-one to elements of $ISO(3)$, and $ISO(3) \ni (\mathbf{t}, R) \mapsto g_{(\mathbf{t},R)}$ is an isomorphism, beside a linear representation of $ISO(3)$. In order to make the action of $ISO(3)$ explicit on points in $\mathbb{R}^3$, let us write points as column vectors $(1, x_1, x_2, x_3)^t$ of $\mathbb{R}^4$, where x_1, x_2, x_3 are the Cartesian coordinates of $\mathbf{x} \in \mathbb{R}^3$. In this way we recover the action of $g_{(\mathbf{t},R)}$ on $\mathbb{R}^3$ described by (12.61). We can indifferently see $ISO(3)$ as the group of maps (12.61) or the matrix group (12.62). In either case it will be a topological group from now on. Similarly we may imagine $IO(3)$ as a matrix group, simply allowing R to vary in the whole $O(3)$. With the given topologies, the construction makes $ISO(3)$ a topological subgroup of $IO(3)$ and its connected component at the identity $(\mathbf{0}, I)$.

The linear unitary $ISO(3)$-representation on $L^2(\mathbb{R}^3, dx)$ seen in Example 12.17(1):

$$(U_\Gamma \psi)(\mathbf{x}) := \psi(\Gamma^{-1}\mathbf{x})\,, \quad \Gamma \in ISO(3),\ \psi \in L^2(\mathbb{R}^3, dx)$$

is strongly continuous, since

$$||U_\Gamma \psi - U_{\Gamma_0}\psi|| = ||U_{\Gamma_0^{-1}\circ\Gamma}\psi - \psi|| \to 0 \quad \text{as } \Gamma \to \Gamma_0. \tag{12.63}$$

Now look at U_Γ acting on pure states of $\mathsf{H} = L^2(\mathbb{R}^3, dx)$:

$$\gamma_\Gamma\,(\psi(\psi|\ )) := U_\Gamma \psi\,(\psi|\ )\,U_\Gamma^{-1}\,.$$

The strongly continuous unitary representation $ISO(3) \ni \Gamma \mapsto U_\Gamma$ renders $ISO(3)$ a *topological group of symmetries* for the spin-zero quantum particle.

(2) Let P_i be the self-adjoint operator of the momentum observable along the axis x_i, and $\mathbf{P}$ the column vector $(P_1, P_2, P_3)^t$. With an eye on the previous example, let us focus on the subgroup of translations along an axis $\mathbf{t} \in \mathbb{R}^3$. Such subgroup is the strongly continuous one-parameter unitary group $\mathbb{R} \ni r \mapsto U_r^{(\mathbf{t})}$, with

$$\left(U_r^{(\mathbf{t})}\psi\right)(\mathbf{x}) := \psi(\mathbf{x} - r\mathbf{t}), \quad t \in \mathbb{R},\ \psi \in L^2(\mathbb{R}^3, dx).$$

It is easy to prove the symmetric operator $\mathbf{t} \cdot \mathbf{P}\restriction_{\mathscr{S}(\mathbb{R}^3)}$ is essentially self-adjoint, so (cf. Lemma 11.11)

$$\left(e^{-i\frac{r}{\hbar}\overline{\mathbf{t}\cdot\mathbf{P}\restriction_{\mathscr{S}(\mathbb{R}^3)}}}\psi\right)(\mathbf{x}) = \psi(\mathbf{x} - r\mathbf{t}), \quad \psi \in L^2(\mathbb{R}^3, dx). \tag{12.64}$$

Therefore:

The self-adjoint operator, which exists by Theorem 12.36(c), generating the strongly continuous one-parameter unitary group of translations along $\mathbf{t}$ *is the momentum operator along* $\mathbf{t}$, *i.e. the only self-adjoint extension of* $\frac{1}{\hbar}\mathbf{t} \cdot \mathbf{P}\restriction_{\mathscr{S}(\mathbb{R}^3)}$ *(up to the constant* $\hbar^{-1}$*).*

Observe that the generator can be modified by adding constants. ■

12.2.7 Round-up on Lie groups and algebras

In this last section we assume the reader is familiar with differentiable manifolds, including real-analytic ones (the basic notions are summarised in the appendix with some detail). We recall fundamental results [NaSt82, War75, Kir74] in the theory of Lie groups and provide a few examples, all without proofs.

Definition 12.38. *A* **real Lie group of dimension** *n is a real-analytic n-manifold* G *equipped with two analytic maps:*

$$\mathsf{G} \ni g \mapsto g^{-1} \in \mathsf{G} \quad \textit{and} \quad \mathsf{G} \times \mathsf{G} \ni (g,h) \mapsto g \cdot h \in \mathsf{G}$$

(where $\mathsf{G} \times \mathsf{G}$ *has the analytic product structure), that make* G *a group with neutral element e.*
The **dimension** *of the Lie group* G *is the dimension n the manifold* G.

Remarks 12.39. Analiticity in Definition 12.38 can be watered down to having G just a topological manifold with continuous operations in the manifold topology (i.e. a topological group that is Hausdorff, paracompact, and locally homeomorphic to $\mathbb{R}^n$). In fact, a famous 1952 theorem of Gleason, Montgomery and Zippin – solving *Hilbert's fifth problem* – proves the following.

Theorem 12.40 (Gleason, Montgomery, Zippin). *Every topological group with the structure of* C^0 *manifold also admits an analytic substructure for which the group operations are analytic. This is completely determined by the* C^0 *structure and the operations.*

Therefore any Lie group can be thought of, in a unique fashion, as an analytic Lie group, even when defined topologically and the operations are only assumed continuous. ■

Definition 12.41. *Consider Lie groups* G, G', *with neutral elements* e, e' *and operations* $\cdot$, $\circ$.
A **Lie group homomorphism** *is an analytic map* $f : \mathsf{G} \to \mathsf{G}'$ *that is also a group homomorphism.*
If the homomorphism $f : \mathsf{G} \to \mathsf{G}'$ *is invertible and* f^{-1} *is a homomorphism,* f *is called* **Lie group isomorphism** *and* G, G' *are* **isomorphic** *(under* f*).*
A **local homomorphism of Lie groups** *is an analytic map* $h : \mathscr{O}_e \to \mathsf{G}'$, *where* $\mathscr{O}_e \subset \mathsf{G}$ *is an open neighbourhood of* e *and* $h(g_1 \cdot g_2) = h(g_1) \circ h(g_2)$ *provided* $g_1 \cdot g_2 \in \mathscr{O}_e$. *(This forces* $h(e) = e'$[5] *and* $h(g^{-1}) = h(g)^{-1}$ *for* $g, g^{-1} \in \mathscr{O}_e$.*)*
If the local homomorphism h *is an analytic diffeomorphism on its range (given by an open neighbourhood* $\mathscr{O}_{e'}$ *of* e'*), and the inverse* $f^{-1} : \mathscr{O}_{e'} \to \mathsf{G}$ *is a local homomorphism, then* h *is a* **local isomorphism of Lie groups**. *The Lie groups* G, G' *are* **locally isomorphic** *(under* h*).*

In the same spirit of the previous remark we may weaken the assumptions of differentiability when defining local homomorphisms [NaSt82].

Proposition 12.42. *Let* G, G' *be Lie groups,* $\mathscr{O}_e \subset \mathsf{G}$ *an open neighbourhood of the identity* $e \in \mathsf{G}$.
If $h : \mathscr{O}_e \to \mathsf{G}'$ *is continuous and* $h(g_1 \cdot g_2) = h(g_1) \circ h(g_2)$ *provided* $g_1 \cdot g_2 \in \mathscr{O}_e$, *then* h *is analytic and thus a local homomorphism of Lie groups.*

Two important concepts for our purposes are *one-parameter subgroups* and *Lie algebras*, which we now recall.

Let G be a Lie group with neutral element e and product $\cdot$. The tangent space at a point $g \in \mathsf{G}$ is denoted $\mathsf{T}_g\mathsf{G}$. Every $g \in \mathsf{G}$ defines an (analytic) map $L_g : \mathsf{G} \ni h \mapsto g \cdot h$, and let us write $dL_g : \mathsf{T}_h\mathsf{G} \to \mathsf{T}_{g\cdot h}\mathsf{G}$ for its differential. Given $A \in \mathsf{T}_e\mathsf{G}$, we consider the first-order *Cauchy problem* on G: find a differentiable $f : (-\alpha, \beta) \to \mathsf{G}$, $\alpha, \beta > 0$ such that

$$\frac{df}{dt} = dL_{f(t)}A \quad \text{with } f(0) = e.$$

The maximal solution is always complete, i.e. with largest-possible domain $(-\alpha, \beta) = \mathbb{R}$. We will indicate the maximal solution with

$$\mathbb{R} \ni t \mapsto \exp(tA)$$

and we will call it the **one-parameter subgroup generated by** A. If $T \in T_e\mathsf{G}$ it can be proved that

$$\exp(tT)\exp(t'T) = \exp((t+t')T)\,, \quad (\exp(tT))^{-1} = \exp(-tT)\,,\ \forall t, t' \in \mathbb{R}.$$

[5] In fact $h(e) = h(e \cdot e) = h(e) \circ h(e)$, so applying $h(e)^{-1}$ we get $e' = h(e)$.

Consider now a given $T \in \mathsf{T}_e\mathsf{G}$ and the collection of maps parametrised by $t \in \mathbb{R}$:

$$F_{t,T} : \mathsf{G} \ni g \mapsto \exp(tT)\, g\, \exp(-tT)\,.$$

As $F_{t,T}(e) = e$, the differential $dF_{t,T}|_e$ maps $\mathsf{T}_e\mathsf{G}$ to itself, and is the **adjoint** of $F_{t,T}$

$$Ad\, F_{t,T} : \mathsf{T}_e\mathsf{G} \to \mathsf{T}_e\mathsf{G}\,.$$

The **commutator** [War75] is the map from $\mathsf{T}_e\mathsf{G} \times \mathsf{T}_e\mathsf{G}$ to $\mathsf{T}_e\mathsf{G}$:

$$[T,Z] := \frac{d}{dt}|_{t=0}(Ad\, F_{t,T})\, Z\,, \quad T,Z \in \mathsf{T}_e\mathsf{G}\,.$$

The commutator has three properties:

linearity:	$[aA+bB,C] = a[A,C]+b[B,C]\,,$
skew-symmetry:	$[A,B] = -[B,A]\,,$
Jacobi identity:	$[A,[B,C]]+[B,[C,A]]+[C,[A,B]] = 0$

holding for any $a,b \in \mathbb{R}$ and $A,B,C \in \mathsf{T}_e\mathsf{G}$. The first two actually imply bilinearity. The third property is a consequence of the associativity of the group law.

Let us fix a local coordinate system $x_1,\dots,x_n$ compatible with the (analytic) structure of G over an open neighbourhood U of e, so that the neutral element becomes the origin. In these coordinates we can expand the group law on $U \times U$ in Taylor series up to the second order

$$\psi(X,X') = X + X' + B(X,X') + O\left(\left(|X|^2 + |X'^2|\right)^{3/2}\right)\,, \tag{12.65}$$

where $X,X' \in \mathbb{R}^n$ are the column vectors of the coordinates of elements $g,g' \in U$ whose product $g \cdot g'$ belongs to U. The mapping $B : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is bilinear, and it is easy to see that the commutator, in the coordinate basis of $\mathsf{T}_e\mathsf{G}$, becomes:

$$[T,T'] = B(T,T') - B(T',T)\,, \tag{12.66}$$

where T, T' are (column) vectors in $\mathsf{T}_e\mathsf{G}$.

Definition 12.43. *A vector space* V *endowed with a bilinear, skew-symmetric map* $[\,,\,] : \mathsf{V} \times \mathsf{V} \to \mathsf{V}$, *called* **Lie bracket**, *that satisfies the Jacobi identity is said a* **Lie algebra**.
Given Lie algebras $(\mathsf{V},[\,,\,])$, $(\mathsf{V}',[\,,\,]')$, *a linear mapping* $\phi : \mathsf{V} \to \mathsf{V}'$ *is a* **Lie algebra homomorphism** *if* $[\phi(A),\phi(B)]' = \phi[A,B]$ *for any* $A,B \in \mathsf{V}$. *If* ϕ *is also bijective one calls it* **Lie algebra isomorphism**.
Given a Lie group G, *the tangent space* $\mathsf{T}_e\mathsf{G}$ *with the Lie bracket* $[\,,\,]$ *given by the commutator is the* **Lie algebra of the Lie group** G.
A **Lie subalgebra** V' *in a Lie algebra* $(\mathsf{V},[\,,\,])$ *is a closed subspace under the Lie bracket,* $[A,B] \in \mathsf{V}'$ *for* $A,B \in \mathsf{V}'$. *An* **ideal** J *in a Lie algebra* $(\mathsf{V},[\,,\,])$ *is a Lie subalgebra such that*

$$[A,B] \in J \quad \text{for any } A \in J,\ B \in \mathsf{V}.$$

A (non-Abelian) Lie algebra V *is said* **simple** *if it contains no proper ideals, and* **semisimple** *if direct sum of simple Lie algebras.*

A crucial feature of Lie groups for physics is that *the Lie algebra of a Lie group determines the group almost entirely*, as the classical, and famous, next results show [NaSt82]. First, though, a topological reminder.

Definition 12.44. *Let* X, R *be topological spaces.* R *is a* **covering space** *of* X *if there is a continuous onto map* $\pi : \mathsf{R} \to \mathsf{X}$*, called* **covering map***, as follows:*

(i) *for any* $x \in \mathsf{X}$ *there exists an open set* $U \ni x$ *such that* $\pi^{-1}(U) = \cup_{j\in J} A_j$*, with* $A_j \subset \mathsf{R}$ *open,* $A_j \cap A_i = \varnothing$ *if* $i \neq j$*,* $i, j \in J$*;*
(ii) $\pi\restriction_{A_j} : A_j \to U$ *is a homeomorphism for every* $j \in J$*.*

A covering R *of* X *is a* **universal covering** *if it is simply connected (Definition 1.28).*

Two universal coverings R, R' of X are homeomorphic under the map $f : \mathsf{R} \to \mathsf{R}'$ such that $\Pi = f \circ \Pi'$, if $\Pi : \mathsf{R} \to \mathsf{X}$, $\Pi' : \mathsf{R}' \to \mathsf{X}$ are the covering maps. Similarly, if X has universal covering R and a covering R', with covering maps $\Pi : \mathsf{R} \to \mathsf{X}$, $\pi : \mathsf{R}' \to \mathsf{X}$, then there is a covering map $p : \mathsf{R} \to \mathsf{R}'$ with $\pi \circ p = \Pi$ [Ser94II].

The first result, called *third Lie theorem* in the literature, says the following.

Theorem 12.45. *Let* V *be a finite-dimensional (real) Lie algebra.*

(a) *There exists a connected and simply connected (real) Lie group* G_V *with Lie algebra* V*.*
(b) G_V *is, up to isomorphisms, the universal covering of any Lie group having* V *as Lie algebra, and the covering map is a Lie group homomorphism.*
(c) *If a Lie group* G *has* V *as Lie algebra, it is isomorphic to a quotient* $\mathsf{G}_\mathsf{V}/H_\mathsf{G}$*, where* $H_\mathsf{G} \subset \mathsf{G}_\mathsf{V}$ *is a discrete normal subgroup (hence contained in the centre of* G_V*).*

Theorem 12.46 (Lie). *Let* G, G' *be (real) Lie groups with Lie algebras* V, V'*.*

(a) $f : \mathsf{V} \to \mathsf{V}'$ *is a Lie algebra homomorphism if and only if there is a local Lie group homomorphism* $h : \mathsf{G} \to \mathsf{G}'$ *such that* $dh|_e = f$*. Moreover:*

(i) *h is determined completely by f;*
(ii) *f is an isomorphism* $\Leftrightarrow$ *h is a local isomorphism.*

(b) *If* G, G' *are connected and* G *also simply connected, then* $f : \mathsf{V} \to \mathsf{V}'$ *is a homomorphism if and only if there is a homomorphism* $h : \mathsf{G} \to \mathsf{G}'$ *such that* $dh|_e = f$*. Moreover:*

(i) *h is determined completely by f;*
(ii) *f isomorphism* $\Rightarrow$ *h onto;*
(iii) *f isomorphism and* G' *simply connected* $\Rightarrow$ *h isomorphism.*

Definition 12.47. *An (embedded) submanifold* $\mathsf{G}' \subset \mathsf{G}$ *in a Lie group that is also a subgroup inherits a Lie group structure from* G*. In such case* G' *is a* **Lie subgroup** *in* G*. (*G' *and* G *have the same identity element* $\{e\}$ *as the inclusion is a homomorphism.) The subgroup* G' *is discrete when the set* $\{e\}$ *is open in the induced topology.*

Notice that G' is discrete iff every singlet $\{g\} \subset \mathsf{G}'$ is open. This is because the translation $\mathsf{G}' \to \mathsf{G}' : h \mapsto gh$ is a homeomorphism. Equivalently the subgroup G' is discrete iff it is made of isolated points. That is, for every $g \in \mathsf{G}'$ there is an open subset A_g of G such that $g \in A_g$ but $h \notin A_g$ if $\mathsf{G}' \ni h \neq g$.

The Lie algebra of a subgroup G' is a **Lie subalgebra** of the Lie algebra of G, meaning $\mathsf{T}_e\mathsf{G}'$ is a subspace of $\mathsf{T}_e\mathsf{G}$ and the bracket on $\mathsf{T}_e\mathsf{G}'$ is the restriction of the one on $\mathsf{T}_e\mathsf{G}$ [NaSt82].

Theorem 12.48 (Cartan). *If* $\mathsf{G}' \subset \mathsf{G}$ *is a closed subgroup of the Lie group* G*, then it is a Lie subgroup of* G.

Remark 12.49. (1) In principle an abstract Lie algebra can have infinite dimension as vector space. The dimension of the Lie algebra of a Lie group G, instead, is always finite for it coincides with the dimension of the manifold G.
(2) Theorem 12.48 clearly subsumes discrete subgroups as special cases. Then the manifold underlying the Lie subgroup has dimension zero.
(3) Let G be a Lie group of dimension n and $\{T_1, \dots T_n\}$ a basis of the Lie algebra $\mathsf{T}_e\mathsf{G}$. As the Lie bracket is bilinear it can be written in components

$$[T_i, T_j] = \sum_{k=1}^{dim \mathsf{T}_e\mathsf{G}} C_{ijk} T_k \,.$$

The coefficients C_{ijk} are the **structure constants** of the Lie group[6]. The Jacobi identity is equivalent to the following equation (of obvious proof):

$$\sum_{s=1}^{n} \left(C_{ijs}C_{skr} + C_{jks}C_{sir} + C_{kis}C_{sjr} \right) = 0\,, \quad r = 1, \dots, n. \tag{12.67}$$

If two Lie groups have the same structure constants with respect to some bases of their Lie algebras, they are *locally isomorphic* in the sense of Theorems 12.45–12.46. (If the structure constants are equal, the linear map identifying bases is an isomorphism.) Conversely, the structure constants of locally isomorphic Lie groups are the same in bases related by the pullback of the local isomorphism. ■

Given a Lie group G, the **exponential mapping** is the analytic function

$$\exp : \mathsf{T}_e\mathsf{G} \ni T \mapsto \exp(tT)|_{t=1} \,.$$

The exponential mapping has an important property, sanctioned by the next result [NaSt82].

Theorem 12.50. *Let* G *be a Lie group with neutral element e and exponential map* exp.

(a) *There exist open neighbourhoods* U *of* $\mathbf{0} \in \mathsf{T}_e\mathsf{G}$ *and* V *of* $e \in \mathsf{G}$ *such that*

$$\exp\restriction_U : U \to \mathsf{V}$$

is an analytic diffeomorphism (bijective, analytic, with analytic inverse).

[6] The structure constants are the components of a *tensor*, called the *structure tensor* of the Lie group.

(b) *If* G *is compact then* $\exp(\mathsf{T}_e\mathsf{G}) = \mathsf{G}$.
(c) *If* G' *is a Lie group with exponential map* $\exp'$ *and* $h : \mathsf{G} \to \mathsf{G}'$ *a homomorphism:*

$$h \circ \exp = \exp' \circ dh|_e .$$

Property (a) has a useful corollary. Fix a basis $T_1, \ldots, T_n$ on the Lie algebra of G, Then the inverse to

$$F : (x_1, \ldots, x_n) \mapsto \exp\left(\sum_{k=1}^{n} x_n T_n\right)$$

defines a local chart, compatible with the analytic structure, around the neutral element. This is called **normal coordinate system**. Normal coordinates, in general, do not cover G. In normal coordinates a vector $T \in \mathsf{T}_e\mathsf{G} \equiv \mathbb{R}^n$ fixes a point of G only around e. Hence group multiplication in G becomes a map $\psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. Expanding the latter with Taylor around the origin of $\mathbb{R}^n \times \mathbb{R}^n$ gives

$$\psi(T, T') = T + T' + \frac{1}{2}[T, T'] + O\left(\left(|T|^2 + |T'^2|\right)^{3/2}\right) , \tag{12.68}$$

where $[T, T'] : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is the commutator in the basis of $\mathsf{T}_e\mathsf{G} \times \mathsf{T}_e\mathsf{G}$ associated to normal coordinates. The proof is left to the reader. Property (a) has also another consequence, whose proof is an exercise.

Proposition 12.51. *Let* G *be a Lie group with neutral element e and product* $\cdot$.

(a) *There exists an open set* $A \ni e$ *in* $\mathsf{T}_e\mathsf{G}$ *such that, for any* $g \in A$, $g = \exp(tT)$ *for some* $t \in \mathbb{R}$ *and some* $T \in \mathsf{T}_e\mathsf{G}$.
(b) *If* G *is connected and* $g \notin A$, *there are finitely many elements* $g_1, g_2, \ldots, g_n \in A$ *such that* $g = g_1 \cdot \cdots \cdot g_n$.

The fundamental **Baker–Campbell–Hausdorff formula** [NaSt82]:

$$\exp(X)\exp(Y) = \exp(Z(X,Y)) \tag{12.69}$$

holds on any connected and simply connected Lie group G, with X, Y in the open neighbourhood U of the origin where exp is a local diffeomorphism onto the open neighbourhood $\exp(U) \subset \mathsf{G}$ of the neutral element. In (12.69) the term $Z(X,Y)$ is defined by the series:

$$Z(X,Y) = \sum_{\mathbb{N} \ni n > 0} \frac{(-1)^{n-1}}{n} \sum_{r_i + s_i > 0, 1 \le i \le n} \frac{\left(\sum_{i=1}^{n}(r_i + s_i)\right)^{-1}}{r_1! s_1! \cdots r_n! s_n!} [X^{r_1} Y^{s_1} X^{r_2} Y^{s_2} \ldots X^{r_n} Y^{s_n}] \tag{12.70}$$

$$[X^{r_1} Y^{s_1} \ldots X^{r_n} Y^{s_n}] := \underbrace{[X, [X, \cdots [X,}_{r_1 \text{ times}} \underbrace{[Y, [Y, \cdots [Y,}_{s_1 \text{ times}} \ldots \underbrace{[X, [X, \cdots [X,}_{r_n \text{ times}} \underbrace{[Y, [Y, \cdots Y]}_{s_n \text{ times}}] \cdots]] \tag{12.71}$$

and the right-hand side is taken to be zero if $s_n > 1$ or $s_n = 0$ and $r_n > 1$.

Examples 12.52. (1) $M(n,\mathbb{R})$ will denote from now on the set of real $n \times n$ matrices, and $M(n,\mathbb{C})$ the same over the complex numbers.

The group $GL(n,\mathbb{R})$ of invertible real $n \times n$ matrices is a n^2-dimensional Lie group with analytic structure induced by $\mathbb{R}^{n^2}$. Its Lie algebra is the set of real $n \times n$ matrices $M(n,\mathbb{R})$ and the Lie bracket is the usual commutator $[A,B] := AB - BA$, $A,B \in M(n,\mathbb{R})$.

An important feature of $GL(n,\mathbb{R})$ is that its one-parameter subgroups have this form:

$$\mathbb{R} \ni t \mapsto e^{tA} := \sum_{k=0}^{+\infty} \frac{t^k}{k!} A^k ,$$

for any $A \in M(n,\mathbb{R})$, and the convergence is in any equivalent Banach norm of $\mathbb{R}^{n^2}$ (or $\mathbb{C}^{n^2}$) (Chapter 2.5).

(2) Any closed subgroup of $GL(n,\mathbb{R})$ we have met as topological group, like $O(n)$, $SO(n), IO(n), ISO(n), SL(n,\mathbb{R})$, the Galilean, Lorentz and Poincaré groups, are therefore Lie groups. As $GL(n,\mathbb{C})$ can be seen as a subgroup in $GL(2n,\mathbb{R})$ (decomposing every matrix element in real and imaginary part), complex matrix groups like $U(n)$ and $SU(n)$, too, are real Lie groups. We must emphasise that working with matrix Lie groups is not great a restriction, for [War75] every compact Lie group is isomorphic to a matrix group. For non-compact Lie groups the story is completely different, a counterexample being the universal covering of $SL(2,\mathbb{R})$.

(3) The exponential of matrices $A,B \in M(n,\mathbb{C})$ has interesting characteristics. First, $e^{A+B} = e^A e^B = e^B e^A$ if $AB = BA$. The proof is similar to the number case that uses Taylor's expansion. There is, though, another useful fact: $A \in M(n,\mathbb{C})$ satisfies, for any $t \in \mathbb{C}$,

$$\det e^{tA} = e^{t \operatorname{tr} A} , \quad \text{in particular} \quad \det e^A = e^{\operatorname{tr} A} .$$

Let us prove this identity. We want to differentiate $\mathbb{C} \ni t \mapsto \det e^{tA}$, i.e. find

$$\lim_{h \to 0} \frac{\det e^{(t+h)A} - \det e^{tA}}{h} = \lim_{h \to 0} \frac{\det (e^{tA} e^{hA}) - \det e^{tA}}{h} = \det e^{tA} \lim_{h \to 0} \frac{\det e^{hA} - 1}{h}$$

as long as the last limit exists. Since $e^{hA} = I + hA + ho(h)$, with $o(h) \to 0$ as $h \to 0$ in the standard topology of $\mathbb{C}^{n^2}$, it follows

$$\lim_{h \to 0} \frac{\det e^{(t+h)A} - \det e^{tA}}{h} = \det e^{tA} \lim_{h \to 0} \frac{\det (I + hA + ho(h)) - 1}{h} .$$

There are many ways to see that $\det(I + hA + ho(h)) = 1 + h\sum_{i=1}^n A_{ii} + h0(h)$, and substituting above we find

$$\frac{d \det e^{tA}}{dt} = \det e^{tA} \operatorname{tr} A .$$

That also proves the function is smooth. Hence $f_A : \mathbb{C} \ni t \mapsto \det e^{tA}$ solves the differential equation:

$$\frac{d f_A(t)}{dt} = (\operatorname{tr} A) f_A(t) .$$

Also $g_A : \mathbb{C} \ni t \mapsto e^{t\,\mathrm{tr}A}$ solves the equation. And both functions satisfy the initial condition $f_A(0) = g_A(0) = 1$, so by uniqueness of maximal solutions of first-order equations we obtain $\det e^{tA} = e^{t\,\mathrm{tr}A}$, any $t \in \mathbb{R}$.

(4) The group of rotations $O(n) := \{R \in M(n,\mathbb{R}) \mid RR^t = I\}$ of $\mathbb{R}^n$ is an important Lie group in physics. That it is a subgroup of $GL(n,\mathbb{R})$ is evident because $\{R \in M(n,\mathbb{R}) \mid RR^t = I\}$ is closed in the Euclidean topology. (Clearly $O(n)$ contains its limit points: $A_k \in O(n)$ and $A_k \to A \in \mathbb{R}^{n^2}$ as $k \to \infty$ imply $A_k^t \to A^t$ and $I = A_k A_k^t \to AA^t$.) The Lie algebra of $O(n)$, denoted $o(n)$, is the vector space of real, skew-symmetric $n \times n$ matrices, and has dimension $n(n-1)/2 = \dim O(n)$. The proof is that Lie algebra vectors are tangent vectors $\dot{R}(0)$ at the identity of the group (the identity matrix) to curves $R = R(u)$ such that $R(u)R(u)^t = I$, $R(0) = I$. By definition, then, they satisfy $\dot{R}(0)R(0)^t + R(0)\dot{R}(0)^t = 0$, i.e. $\dot{R}(0) + \dot{R}(0)^t = 0$. But this defines real skew-symmetric $n \times n$ matrices, a space of dimension $n(n-1)/2$. On the other hand, if A is a real skew-symmetric $n \times n$ matrix, $R(t) = e^{tA} \in O(n)$ as follows form the elementary properties of the exponential function, and $\dot{R}(0) = A$. We conclude that the Lie algebra of $O(n)$ consists of the whole class of real skew-symmetric $n \times n$ matrices.

Eventually note that $O(n)$ is compact, since closed and bounded as we saw earlier. Boundedness is explained in analogy to $U(n)$:

$$||R||^2 = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} R_{ij} R_{ij} \right) = \sum_{i=1}^{n} \delta_{ii} = n \,, \quad \text{for any } R \in O(n)\,.$$

The three-dimensional Lie group $O(3)$ has two connected components: the compact (connected) group $SO(3) := \{R \in O(3) \mid \det R = 1\}$ and the compact set (not a subgroup) $\mathscr{P}SO(3) := \{\mathscr{P}R \in O(3) \mid R \in SO(3)\}$, where $\mathscr{P} := -I$ is the parity transformation.

(5) We will explain how *the whole group* $SO(3)$ *is covered by its exponential map.* Define a special basis of $so(3)$ given by matrices $(T_i)_{jk} = -\varepsilon_{ijk}$ where $\varepsilon_{ijk} = 1$ if i,j,k is a cyclic permutation of $1,2,3$, $\varepsilon_{ijk} = -1$ if i,j,k is a non-cyclic permutation, $\varepsilon_{ijk} = 0$ otherwise. More explicitly

$$T_1 := \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad T_2 := \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \quad T_3 := \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{12.72}$$

All are skew-symmetric so they belong in $so(3)$, and they are clearly linearly independent, hence a basis of $so(3)$. Structure constants are simple in this basis:

$$[T_i, T_j] = \sum_{k=1}^{3} \varepsilon_{ijk} T_k \,. \tag{12.73}$$

The exponential representation of $SO(3)$ is as follows: $R \in SO(3)$ if and only if there exist a unit vector $\mathbf{n} \in \mathbb{R}^3$ and a number $\theta \in \mathbb{R}$ such that

$$R = e^{\theta \mathbf{n}\cdot\mathbf{T}}, \quad \text{where} \quad \mathbf{n}\cdot\mathbf{T} := \sum_{i=1}^{3} n_i T_i \,.$$

(6) The compact group $SU(2)$, seen as *real* Lie group, has Lie algebra the real vector space of skew-Hermitian matrices with zero trace (traceless because the determinant in the group equals 1). Consequenly the Lie algebra of $SU(2)$ has a basis formed by $-\frac{i}{2}\sigma_j$, $j=1,2,3$, where σ_k are the Pauli matrices σ_i of (12.10)-(12.10). The factor $1/2$ is present so to satisfy the commutation relations:

$$\left[-\frac{i\sigma_i}{2}\,,\,-\frac{i\sigma_j}{2}\right]=\sum_{k=1}^{3}\varepsilon_{ijk}\left(-\frac{i\sigma_k}{2}\right). \tag{12.74}$$

By the remark ensuing Theorem 12.48 the Lie algebras of $SU(2)$ and $SO(3)$ are isomorphic. Hence by Theorems 12.45–12.46 the Lie groups are locally isomorphic. As $SU(2)$ is connected and simply connected (it is homeomorphic to the boundary $\mathbb{S}^3$ of the unit ball in $\mathbb{R}^4$), whereas $SO(3)$ is not simply connected, $SU(2)$ must be the universal covering of $SO(3)$. The Lie algebra isomorphism should arise from differentiating a surjective homomorphism from $SU(2)$ to $SO(3)$. The latter is actually well known (Exercise 12.16), so let us recall it briefly. The exponential map of $SU(2)$ covers the entire group by compactness. In practice every matrix $U\in SU(2)$ can be written

$$U=e^{-i\theta\mathbf{n}\cdot\frac{\sigma}{2}}$$

where $\theta\in\mathbb{R}$ and $\mathbf{n}$ is a unit vector in $\mathbb{R}^3$. The aforementioned sujective morphism is the onto map

$$R:SU(2)\ni e^{-i\theta\mathbf{n}\cdot\frac{\sigma}{2}}\mapsto e^{\theta\mathbf{n}\cdot\mathbf{T}}\in SO(3)\,.$$

Clearly this is not invertible, because the right-hand side is invariant under translations $\theta\to\theta+2\pi$, while the left-hand side changes sign (take the unit $\mathbf{n}=\mathbf{e}_3$ along the axis x_3). In fact it is easy to see that the kernel of h consists of two points $\pm I\in SU(2)$. ■

12.2.8 Symmetry Lie groups, theorems of Bargmann, Gårding, Nelson, FS[3]

To conclude our treatise on symmetry groups we deal with connected Lie groups G. Any projective G-representation must be representable by unitary operators and never by antiunitary ones. We have in fact the following result.

Proposition 12.53. *Let* G *be a connected Lie group. For any projective representation* $\mathsf{G}\ni g\mapsto\gamma_g$ *the images* γ_g *can be associated to unitary operators only, according to Wigner's theorem (or Kadison's).*

Proof. By Proposition 12.51, every $g\in\mathsf{G}$ is the product of a finite number of elements $h=\exp(tT)$. Then $h=r\cdot r$ with $r=\exp(tT/2)$. Using Proposition 12.24 the claim follows. □

At this point we deal with a number of general results on strongly continuous unitary representations of Lie groups.

It will be useful in the sequel to observe, first of all, that any projective representation of a topological group G may be seen as projective representation of its universal covering group $\widetilde{\mathsf{G}}$.

In fact if $\pi : \widetilde{\mathsf{G}} \to \mathsf{G}$ is the covering map (a continuous homomorphism of topological groups [NaSt82]), and $\gamma : \mathsf{G} \ni g \mapsto \gamma_g$ is a continuous projective G-representation on the Hilbert space H, then $\gamma \circ \pi : \widetilde{\mathsf{G}} \ni h \mapsto \gamma_{\pi(h)}$ is a continuous projective $\widetilde{\mathsf{G}}$-representation; note that it does not distinguish elements $h, h' \in \widetilde{\mathsf{G}}$ if $\pi(h) = \pi(h')$. Put equivalently, if $h \cdot h'^{-1} \in Ker(\pi)$ then $\gamma \circ \pi(h) = \gamma \circ \pi(h')$ i.e. $(\gamma \circ \pi)(Ker(\pi)) = id$, or $Ker(\pi) \subset Ker(\gamma \circ \pi)$. This proves the following.

Proposition 12.54. *Let* G *be a topological group and* $\pi : \widetilde{\mathsf{G}} \to \mathsf{G}$ *its universal covering. Every continuous projective representation* $\gamma : \mathsf{G} \ni g \mapsto \gamma_g$ *of* G *on the Hilbert space* H *arises from the continuous projective representation* $\gamma' : \widetilde{\mathsf{G}} \ni g \mapsto \gamma'_g$ *on* H *such that* $Ker(\pi) \subset Ker(\gamma')$*, induced by* $\mathsf{G} \equiv \widetilde{\mathsf{G}}/Ker(\pi)$.

Remarks 12.55. When needed, henceforth, we will use projective unitary representations of $\widetilde{\mathsf{G}}$ instead of G, because the latter are determined by the former. ■

We will prepare the ground for an important theorem due to Bargmann [Bar54], that provides sufficient conditions for a continuous projective representation to be given by a unitary representation. The preliminary idea, presented in Chapter 12.2.4, is that a projective unitary representation

$$\mathsf{G} \ni g \mapsto U_g$$

of a group G is the restriction of a unitary representation

$$\widehat{\mathsf{G}}_\omega \ni g \mapsto V_g$$

of a suitable central extension $\widehat{\mathsf{G}}_\omega$ of G. This is always possible by virtue of Proposition 12.27. Assume G is a Lie group, and the projective unitary representation $\mathsf{G} \ni g \mapsto U_g$ induces a continuous projective representation. We can choose the phases of the U_g so that the representation $\mathsf{G} \ni g \mapsto U_g$ is continuous around the identity of G by Proposition 12.35. This cannot be extended to the entire G, in general. But using a representation of $\widehat{\mathsf{G}}_\omega$ and the Lie structure of G allows to do so: the next technical result, cited without proof [Kir74], explains how.

Theorem 12.56. *Let* G *be a connected Lie group and* $\mathsf{G} \ni g \mapsto \gamma_g$ *a continuous projective representation on the Hilbert space* H. *There exist a central extension* $\widehat{\mathsf{G}}_\omega$ *and a strongly continuous unitary representation*

$$\widehat{\mathsf{G}}_\omega \ni (\chi, g) \mapsto V_{(\chi,g)}$$

with $\omega(e,e) = 1$, $V_{(\chi,e)} = \chi I$ *for any* $\chi \in U(1)$. *Moreover:*

(a) $\widehat{\mathsf{G}}_\omega$ *is a connected Lie group (whose global differentiable structure is not the product structure of* $U(1)$ *and* G*, in general); the canonical inclusion* $U(1) \to \widehat{\mathsf{G}}_\omega$ *and canonical projection* $\widehat{\mathsf{G}}_\omega \to \mathsf{G}$ *are Lie group homomorphisms.*
(b) *As differentiable manifold* $\widehat{\mathsf{G}}_\omega$ *is, around the identity, the local product of the standard* $U(1)$ *and* G*, and the map* $\omega : \mathsf{G} \times \mathsf{G} \to U(1)$ *arises from Proposition 12.35 by an equivalence, making it* C^∞ *around* (e,e).

(c) *the map* $(1,g) \mapsto V_{(1,g)}$ *is a strongly continuous projective unitary representation that induces* $\mathsf{G} \ni g \mapsto \gamma_g$*:*

$$\gamma_g(\rho) = V_{(1,g)} \rho V_{(1,g)}^{-1} \quad \textit{for any } g \in \mathsf{G},\ \rho \in \mathfrak{S}_p(\mathsf{H}).$$

So let us assume, by the above theorem, any strongly continuous projective representation of a Lie group G is obtainable as strongly continuous projective unitary representation of a central extension of G, itself a Lie group.

This allows to prove the aforementioned Bargmann theorem.

Let us go through the proof's idea, heuristically. Take a Lie group G (connected and simply connected in the theorem) and its central $U(1)$-extensions $\widehat{\mathsf{G}}_\omega$. Projective unitary representations of G are honest unitary representations of $U(1)$-extensions of G. The question is when are continuous unitary representations of $\widehat{\mathsf{G}}_\omega$ reducible to continuous unitary representations of G. The Lie algebra of $\widehat{\mathsf{G}}_\omega$ is the vector space $\mathbb{R} \oplus \mathsf{T}_e\mathsf{G}$ with bracket

$$\left[r \oplus T, r' \oplus T'\right] = \alpha(T,T') \oplus [T,T'],$$

where $r \oplus T$ is the generic element in $\mathbb{R} \oplus \mathsf{T}_e\mathsf{G}$ and $\alpha : \mathsf{T}_e\mathsf{G} \times \mathsf{T}_e\mathsf{G} \to \mathsf{T}_e\mathsf{G}$ a bilinear skew-symmetric map. An common alternative way to write this is to fix a basis of $\mathsf{G}_e T$ and set

$$[T_i, T_j] = \alpha_{ij} I + \sum_{k=1}^{n} C_{ijk} T_k\,, \tag{12.75}$$

where $r = r' = 0$, $\alpha_{ij} := \alpha(T_i, T_j)$ are by construction skew-symmetric, $\alpha_{ij} = \alpha_{ji}$, and in consequence of Jacobi's identity (and corresponding to (12.79) under Bargmanns' theorem):

$$\alpha_{ij} = \alpha_{ji}\,, \tag{12.76}$$

$$0 = \sum_{s=1}^{n} \left(C_{ijs}\alpha_{sk} + C_{jks}\alpha_{si} + C_{kis}\alpha_{sj}\right)\,. \tag{12.77}$$

The numbers α_{ij} are often called **central charges**. The key idea behind Bargmann's theorem is to redefine the generators

$$T_k \to T'_k := \beta_k I + T_k$$

so that the β_k absorb central charges, allowing to write the bracket relations of the Lie algebra of G as:

$$[T'_i, T'_j] = \sum_{k=1}^{n} C_{ijk} T'_k\,.$$

If this is possible, we expect to view a unitary $\widehat{\mathsf{G}}_\omega$-representation as a G-representation. Referring to (12.75), we understand that that is true when the β_k solve

$$\alpha_{ij} = \sum_{k=1}^{n} C_{ijk} \beta_k \tag{12.78}$$

(note C_{ijk} and α_{ij} are given, once $\widehat{\mathsf{G}}_\omega$ is known). The hypothesis of Bargmann's theorem, formulated by (12.80), is just condition (12.78), as the proof will explain. The linear function β of the statement, in fact, is completely determined by the coefficients β_k if we set $\beta(T_k) := \beta_k$.

Theorem 12.57 (Bargmann). *Let* G *be a connected, simply connected Lie group. Every continuous projective* G*-representation on the Hilbert space* H *is induced by a strongly continuous unitary representation on* H *provided the following condition holds. For any skew-symmetric bilinear map* $\alpha : \mathsf{T}_e\mathsf{G} \times \mathsf{T}_e\mathsf{G} \to \mathbb{R}$ *satisfying*

$$\alpha\left([T,T'],T''\right) + \alpha\left([T',T''],T\right) + \alpha\left([T'',T],T'\right) = 0\,, \quad T,T',T'' \in \mathsf{T}_e\mathsf{G}, \quad (12.79)$$

there exists a linear map $\beta : \mathsf{T}_e\mathsf{G} \to \mathbb{R}$ *with*

$$\alpha(T,T') = \beta\left([T,T']\right)\,, \quad T,T' \in \mathsf{T}_e\mathsf{G}. \quad (12.80)$$

Proof. Consider a continuous projective representation $\gamma : \mathsf{G} \ni g \mapsto \gamma_g$ on the Hilbert space H. By Theorem 12.56, suitably choosing the multiplier function allows to define a central $U(1)$-extension $\widehat{\mathsf{G}}_\omega$ of G, and a projective unitary representation $V : \mathsf{G} \ni g \mapsto V_g$ su H that is strongly continuous and induces γ. The canonical inclusion and projection homomorphisms are Lie morphisms. Moreover, around the origin $\widehat{\mathsf{G}}_\omega$ is the product $U(1) \times \mathsf{G}$ and the map ω is differentiable for the structure of $\mathsf{G} \times \mathsf{G}$, locally. The multiplier function is *normalised* so that $\omega(e,e) = \omega(e,g) = \omega(g,e) = 1$, hence the neutral element of $\widehat{\mathsf{G}}_\omega$ is $(1,e)$. (As we know, we can always reduce to this case via an equivalence transformation by a constant map.) The real vector space underlying the Lie algebra of $\widehat{\mathsf{G}}_\omega$ is $\mathbb{R} \oplus \mathsf{T}_e\mathsf{G}$, where $\oplus$ is the direct sum (not orthogonal, as there is no inner product around). We will denote $r \oplus T$ the elements, where $r \in \mathbb{R}$ and $T \in \mathsf{T}_e\mathsf{G}$. By the definition of Lie bracket $[\ ,\ ]$ of $\mathsf{T}_e\mathsf{G}$, a few computations involving (12.66) say that the bracket $[\ ,\ ]_\omega$ of $\mathsf{T}_{1\oplus e}\widehat{\mathsf{G}}_\omega$ has the form:

$$\left[r \oplus T, r' \oplus T'\right]_\omega = \alpha(T,T') \oplus [T,T'] \quad (12.81)$$

where $\alpha : \mathsf{T}_e\mathsf{G} \times \mathsf{T}_e\mathsf{G} \to \mathbb{R}$ is a bilinear skew-symmetric map satisfying (12.79), owing to the Jacobi identity of $[\ ,\ ]_\omega$. Now we show the universal covering of $\widehat{\mathsf{G}}_\omega$ is the Lie group $\mathbb{R} \otimes \mathsf{G}$, where $\otimes$ is the *direct product* of the two ($\mathbb{R}$ is an additive Lie group). The direct product of two Lie groups is a Lie group with the analytic product structure. The topological space underlying $\mathbb{R} \otimes \mathsf{G}$ is the product $\mathbb{R} \times \mathsf{G}$, simply connected as the factors are. By Theorem 12.45 $\mathbb{R} \otimes \mathsf{G}$ is the unique simply connected Lie group, up to isomorphisms, having that Lie algebra, and hence is the universal covering of all Lie groups with the Lie algebra of $\mathbb{R} \otimes \mathsf{G}$. We will show $\widehat{\mathsf{G}}_\omega$ is one of those. The Lie algebra of $\mathbb{R} \otimes \mathsf{G}$ is $\mathbb{R} \oplus \mathsf{T}_e\mathsf{G}$ with bracket:

$$\left[r \oplus T, r' \oplus T'\right]_\otimes = 0 \oplus [T,T']\,. \quad (12.82)$$

To prove the claim it suffices to exhibit an isomorphism mapping the Lie algebra of $\mathbb{R} \otimes \mathsf{G}$ to the Lie algebra of $\widehat{\mathsf{G}}_\omega$, when there is $\beta : \mathsf{T}_e\mathsf{G} \to \mathbb{R}$ satisfying (12.80). Let us construct the isomorphism. Fix a basis $T_1,\dots,T_n$ in the Lie algebra of G, and a corresponding basis

$$1\oplus 0,\ 0\oplus T_1,\dots,\ 0\oplus T_n\ \in \mathsf{T}_{(0,e)}\mathbb{R}\otimes\mathsf{G}$$

in the Lie algebra of $\mathbb{R}\otimes\mathsf{G}$. Consider the new basis in the Lie algebra of $\widehat{\mathsf{G}_\omega}$:

$$1\oplus 0,\ \beta(T_1)\oplus T_1,\dots,\ \beta(T_n)\oplus T_n\ \in \mathsf{T}_{(1,e)}\widehat{\mathsf{G}_\omega}\,.$$

This is clearly a basis because the vectors are linearly independent if $T_1,\dots,T_n$ form a basis. Consider the unique linear bijection $f:\mathsf{T}_{(0,e)}\mathbb{R}\otimes\mathsf{G}\to\mathsf{T}_{(1,e)}\widehat{\mathsf{G}}_\omega$ such that:

$$f(1\oplus 0):=1\oplus 0\,,\quad f(0\oplus T_k):=\beta(T_k)\oplus T_k\quad \text{for } k=1,2,\dots,n.$$

We claim it preserves brackets:

$$[f(r\oplus T)\,,f(r'\oplus T')]_\omega = f([r\oplus T\,,r'\oplus T]_\otimes)\,,$$

and hence is an isomorphism. As f is linear and brackets are bilinear and skew, it is enough to prove the claim on pairs of distinct basis elements. Evidently $[f(1\oplus 0)\,,f(0\oplus T_k)]_\omega = 0 = f([1\oplus 0\,,0\oplus T_k]_\otimes)$. As for the remaining non-trivial commutators,

$$[f(0\oplus T_h)\,,f(0\oplus T_k)]_\omega = [\beta(T_h)\oplus T_h\,,\beta(T_k)\oplus T_k]_\omega = \alpha(T_h,T_k)[T_h,T_k]$$

$$=\beta([T_h,T_k])\oplus[T_h,T_k] = \beta\left(\sum_{s=1}^n C_{hks}T_s\right)\oplus\sum_{s=1}^n C_{hks}T_s = \sum_{s=1}^n C_{hks}\left(\beta(T_s)\oplus T_s\right)$$

$$=\sum_{s=1}^n C_{hks}f\left(0\oplus T_s\right) = f\left(\sum_{s=1}^n C_{hks}0\oplus T_s\right) = f\left([0,\oplus T_h\,,0\oplus T_s]_\otimes\right)\,.$$

where C_{hks} are the structure constants of G in the basis $T_1,\dots,T_n$. Therefore the universal covering of $\widehat{\mathsf{G}}_\omega$ is $\mathbb{R}\otimes\mathsf{G}$, and there is a surjective Lie homomorphism

$$\Pi:\mathbb{R}\otimes\mathsf{G}\ni(r,g)\mapsto(\chi(r,g),h(r,g))\in\widehat{\mathsf{G}}_\omega\,,$$

such that

$$d\Pi|_{(0,g)} = f \tag{12.83}$$

(the latter determines the map uniquely, by Theorem 12.46). Now let us study the homomorphism Π, keeping exploiting that $\widehat{\mathsf{G}}_\omega$ is a central $U(1)$-extension of G. Easily $h(r,e)=e$ for any $r\in\mathbb{R}$. Consider in fact the one-parameter group of $\mathbb{R}\otimes\mathsf{G}$

$$\mathbb{R}\ni r\mapsto(r,e)=\exp\{r(1\oplus 0)\}\,;$$

Π maps it, by Theorem 12.45(c), to the one-parameter subgroup of $\widehat{\mathsf{G}}_\omega$:

$$\mathbb{R}\ni r\mapsto\exp\{rf(1\oplus 0)\} = (\chi(r,e),h(r,e)) = \exp\{r(1\oplus 0)\} = \exp\{(r\oplus 0)\}\,.$$

The Baker–Campbell–Hausdorff formula (12.69) and the relations (12.81) give, for any $r \in \mathbb{R}$ around 0:

$$(\chi(r,e), h(r,e)) = \exp\{(r \oplus 0)\} = (\chi(r,e), e).$$

As $h(r,e)h(s,e) = h(r+s,e)$ by the properties of one-parameter subgroups, the identity found extends to any $r \in \mathbb{R}$, so $h(r,e) = e$ for every $r \in \mathbb{R}$. Define $\chi(r) := \chi(r,e)$. Then

$$\Pi : (r,e) \mapsto (\chi(r), e) \quad \text{and} \quad \chi(r)\chi(r') = \chi(r+r') \quad \text{for any } r, r' \in \mathbb{R}.$$

The second equation follows because $r \mapsto \exp\{rf(1 \oplus 0)\} = (\chi(r,e), h(r,e))$ is a one-parameter subgroup. Setting $h(g) := h(0,g)$ and $\phi(g) := \chi(0,g)$, we can write

$$\Pi : \mathbb{R} \otimes \mathsf{G} \ni (r,g) \mapsto (\chi(r)\phi(g), h(g)) \in \widehat{\mathsf{G}}_\omega . \tag{12.84}$$

Let us study the map $h : (0,g) \mapsto g$ and prove it is an isomorphism. As Π is a group homomorphism it maps the product $(r,g) \cdot (r',g')$ to the images' product, so

$$(\chi(r), h(g)) \cdot (\chi(r), h(g')) = (\chi(r+r')\phi(g)\phi(g')\omega(h(g),h(g')) \,,\, h(gg')) .$$

This implies $h : \mathsf{G} \ni g \equiv (0,g) \mapsto h(g) \in \mathsf{G}$ – the domain G being a Lie subgroup in $\mathbb{R} \otimes \mathsf{G}$ – is a group homomorphism. But Π is onto, so h is too. The map $\widehat{\mathsf{G}}_\omega(\chi, s) \mapsto s \in \mathsf{G}$ is an onto Lie homomorphism by definition of central extension, so we conclude $h : \mathsf{G} \ni g \mapsto h(g) \in \mathsf{G}$ is a surjective Lie homomorphism. By (12.83), it is easy to see $dh : 0 \oplus T_k \to T_k$. Consequently, by (iii) of Theorem 12.46(b) dh is the differential at the identity of a unique Lie isomorphism from G (subgroup of $\mathbb{R} \oplus \mathsf{G}$) to G. By construction it must coincide with h.

To finish take the multiplier function ω and $\phi : \mathsf{G} \to U(1)$. Then $\phi(e) = 1$, because $\Pi : (0,e) \mapsto (1,e)$. Since $\Phi : (0,g) \mapsto (\phi(g), h(g))$ is a Lie homomorphism and the analytic structure of $\widehat{\mathsf{G}}_\omega$ is the product around the identity, there ϕ is differentiable. The projection Π maps $(0,g) \cdot (0,g')$ to the product of the images. Therefore

$$(\phi(g)\phi(g')\omega(h(g),h(g')) \,,\, h(gg')) = (\phi(gg'), h(gg')) ,$$

so

$$\phi(g)\phi(g')\omega(h(g),h(g')) = \phi(gg') , \quad g, g' \in \mathsf{G}. \tag{12.85}$$

There remains to find a continuous unitary representation

$$U : \mathsf{G} \ni g \mapsto U_g$$

inducing the projective representation γ. Since $h : \mathsf{G} \to \mathsf{G}$ is an isomorphism, define

$$U_g := \phi(h^{-1}(g))V_g , \quad g \in \mathsf{G}.$$

By construction this projective unitary representation induces γ, since $\phi(h^{-1}(g)) \in U(1)$. At the same time, by (12.85):

$$\begin{aligned} U_g U'_g &= \phi(h^{-1}(g))\phi(h^{-1}(g'))V_g V_{g'} \\ &= \omega(g,g')\phi(h^{-1}(g))\phi(h^{-1}(g'))V_{gg'} = \phi(gg')V_{gg'} = U_{gg'} . \end{aligned}$$

Hence U is a proper unitary representation. To finish we show U is continuous. As V is continuous, h^{-1} is continuous, $h^{-1}(e) = e$ and ϕ is continuous around e, $g \mapsto U_g = \phi(h^{-1}(g))V_g$ is certainly continuous on a neighbourhood A of the neutral element e of G. That U is a representation of unitary operators implies its continuity (in the strong topology) at every point. In fact, if $\psi \in \mathsf{H}$:

$$||U_g\psi - U_{g_0}\psi|| = ||U_{g_0^{-1}}(U_g\psi - U_{g_0}\psi)|| = ||U_{g_0^{-1}g}\psi - \psi|| \to 0 \quad \text{as } g \to g_0.$$

We used the fact that $g_0^{-1}g \in A$ if g is sufficiently close to g_0, as G is a topological group. □

Remark 12.58. **(1)** By a previous remark, Bargmann's theorem provides informations also in case the connected Lie group is not simply connected, by looking at its projective representations as representations of the (simply connected) universal covering.
(2) An alternative, and more sophisticated, way to state Bargmann's theorem relies on the *cohomology theory of Lie groups*. The existence of a linear map β for any bilinear skew α satisfying (12.79) is equivalent to imposing that *the second real cohomology group* $H^2(\mathsf{T}_e\mathsf{G}, \mathbb{R})$ of the Lie algebra is trivial [BaRa86, Kir74]. An important result relying on group cohomology techniques is that Bargmann's theorem holds for simply connected Lie groups G whose Lie algebra is simple or semisimple. Physically important cases are $SL(2,\mathbb{C})$ (the universal covering of the Lorentz group) and the universal covering of the Poincaré group, since the Lie algebras of those groups are semisimple. Therefore, dealing with relativistic quantum theories one can always take advantage of Bargmann's theorem dealing with spacetime symmetries. Conversely, the treatment of spacetime symmetries in Galilean quantum mechanics is much more complicated as we shall see soon.
(3) Remember that Bargmann's theorem gives sufficient, not necessary, conditions. ■

Examples 12.59. **(1)** The simples instance, yet far from trivial, for which Bargmann's theorem applies it the Abelian Lie group $\mathbb{R}$. The assumptions are automatic, for the Lie algebra is $\mathbb{R}$ with zero bracket, and the only skew functional $\alpha : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the null map. However, the result is not obvious, as confirmed by the fact we proved it not effortlessly using Theorem 12.36 (using only the topological group structure, actually
(2) Consider the simply connected Lie group $SU(2)$, and indirectly $SO(3)$, which has $SU(2)$ as universal covering (Examples 12.52(5) and (6)). We want to prove all continuous projective unitary $SU(2)$-representations (hence of $SO(3)$ by Proposition 12.54) are induced by corresponding strongly continuous unitary $SU(2)$-representations, because the latter's Lie algebra befits Bargmann's theorem.

The Lie algebra $su(2)$ of $SU(2)$ (Example 12.52(6)) has a basis made of $-i\sigma_k/2$, where $\sigma_1, \sigma_2, \sigma_3$ are the *Pauli matrices* seen several times. Identify $su(2)$ with $\mathbb{R}^3$ by the vector space isomorphism that sends the basis of $su(2)$ to the canonical basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ of $\mathbb{R}^3$. Every linear skew functional $\alpha : su(2) \times su(2) \to \mathbb{R}$ is a real skew-symmetric matrix A, in the sense there is a unique real skew-symmetric 3×3 matrix

A such that $\alpha(u,v) = \sum_{i,j=1}^{3} u_i A_{ij} v_k$, $u,v \in \mathbb{R}^3$ (the proof is left to the reader). By (12.74), condition (12.80) reads, in terms of the A associated to the functional α,

$$\sum_{i,j=1}^{3} u_i A_{ij} v_k = \beta\left(\sum_{r,s,k=1}^{3} \varepsilon_{rsk} u_r v_s \mathbf{e}_k\right),$$

for any $u,v \in \mathbb{R}^3$ (i.e. $su(2)$) and a given linear functional $\beta : \mathbb{R}^3 \to \mathbb{R}$ (to be determined). By the latter's linearity, we can rephrase:

$$\sum_{i,j=1}^{3} u_i A_{ij} v_k = \sum_{r,s,k=1}^{3} \varepsilon_{rsk} u_r v_s b_k\,,$$

for any $u,v \in \mathbb{R}^3$ and some $b \in \mathbb{R}^3$ whose components $b_k = \beta(\mathbf{e}_k)$ determine the functional β. Observe that the vector b, i.e. the functional β satisfying (12.80), exists, since every real skew matrix A acting on $\mathbb{R}^3$ corresponds one-to-one to some $b \in \mathbb{R}^3$: $A_{ij} = \sum_{k=1}^{3} \varepsilon_{ijk} b_k$ (inverting $b_k = \frac{1}{2}\sum_{i,j=1}^{3} \varepsilon_{ijk} A_{ij}$), as is well known and as one proves with ease.

Therefore condition (12.80) holds for any linear skew functional, and so Bargmann's theorem applies. Note that we did not have to assume (12.79) for the skew functional $\alpha : su(2) \times su(2) \to \mathbb{R}$, for that is granted: using $\alpha(u,v) = \sum_{i,j,k=1}^{3} \varepsilon_{ijk} u_i v_j b_k$, where $b \in \mathbb{R}^3$ determines α, a direct computation shows (12.79) is valid, because of the known formula

$$\sum_{k=1}^{3} \varepsilon_{ijk} \varepsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}\,.$$

■

Now we will discuss the converse problem: construct continuous projective representations that give a Lie group of symmetries. We already know it suffices to build continous unitary representations of the group's central extensions, so we concentrate on the problem of manifacturing strongly continuous unitary representations of a given Lie group. The idea is to start from a Lie algebra representation in terms of self-adjoint operators, reminiscent of the exponentiation of the generators of a Lie group. Physically, the procedure is appealing because generators have a precise meaning. In the next chapter we will see that the generators (self-adjoint operators) represent preseved quantities during motion, if the time evolution is a subgroup of the symmetry group.

As first thing we deal with constructing an operator representation for the Lie algebra, in presence of a strongly continuous unitary representation of the Lie group. Consider a strongly continuous unitary representation of the Lie group G

$$\mathsf{G} \ni g \mapsto U_g$$

on the Hilbert space H. Fix a one-parameter subgroup $\mathbb{R} \ni t \mapsto \exp(tT) \in \mathsf{G}$ associated to the element $T \in \mathsf{T}_e\mathsf{G}$. Stone's Theorem 9.29 ensures

$$U_{\exp(tT)} = e^{-itA_U(T)}\,, \quad \text{for any } t \in \mathbb{R}, \tag{12.86}$$

where $A_U(T)$ is a self-adjoint operator on H, in general unbounded (the sign $-$ is conventional) with domain $D(A_U(T))$, and completely determined by $T \in \mathsf{T}_e\mathsf{G}$. We will call the self-adjoint operators $A_U(T)$, $T \in \mathsf{T}_e\mathsf{G}$, the **generators** of the representation U. From Stone's theorem thay are defined as

$$A_U(T)\psi := i\frac{d}{dt}|_{t=0}U_{\exp(tT)}\psi \quad \text{iff } \psi \in D(A_U(T)). \tag{12.87}$$

Regarding the fact the $-iA_U(T)$ define a representation of the Lie algebra of G, we can only hope they satisfy:

$$(A_U(T)A_U(T') - A_U(T')A_U(T))\psi = iA_U([T,T'])\psi\,, \tag{12.88}$$

$\psi \in \mathscr{D}$, where $\mathscr{D} \subset D(A_U(T))$ is an invariant subspace for all the $A_U(T)$. As a matter of fact it is well known [BaRa86] that such a $\mathscr{D}$ exists and is *dense* in H. A first candidate is the **Gårding space** $\mathscr{D}_G$, defined as the subspace in H containing all vectors ψ such that $\mathsf{G} \ni g \mapsto U_g\psi$ is a *smooth* map (differentiating in the Hilbert topology and with respect to any local coordinate system on G). If ψ belongs to the Gårding space, $\mathscr{D}_G$ is dense and invariant for every $A_U(T)$, plus $\mathsf{T}_e\mathsf{G} \ni T \mapsto -iA_U(T)\restriction_{\mathscr{D}_G}$ is a representation of the Lie algebra $\mathsf{T}_e\mathsf{G}$, meaning it is linear and satisfies (12.88) [BaRa86].

A technical useful result, due to Gårding [BaRa86], says that $\mathscr{D}_G$ is a core for the generators:

Theorem 12.60 (Gårding). *Let* G *be a Lie group and* $\mathsf{G} \ni g \mapsto U_g$ *a strongly continuous unitary representation on the Hilbert space* H*. Let* $\mathscr{D}_G$ *be the Gårding space and* $\mathsf{T}_e\mathsf{G} \ni T \mapsto -iA_U(T)$ *the representation of the Lie algebra on* $\mathscr{D}_G$ *as explained above. Then every operator* $A_U(T)$ *and every real polynomial* $p(A_U(T))$*,* $T \in \mathsf{T}_e\mathsf{G}$*, are essentially self-adjoint on* $\mathscr{D}_G$*.*

There is another space $\mathscr{D}_N$ with similar features to $\mathscr{D}_G$. Found by Nelson [BaRa86], it turns out to be more useful than the Gårding space to *recover* the representation U by exponentiating the Lie algebra representation.

By definition $\mathscr{D}_N$ consists of vectors $\psi \in \mathsf{H}$ such that $\mathsf{G} \ni g \mapsto U_g\psi$ is *analytic* in g, i.e. developable in power series in analytic coordinates around every point of G. The elemens of $\mathscr{D}_N$ are called **analytic vectors of the representation** U and $\mathscr{D}_N$ is the **space of analytic vectors of the representation** U. Thus [BaRa86] $\mathscr{D}_N \subset \mathscr{D}_G$, and $\mathscr{D}_N$ is invariant for the operators $A_U(T)$ and for every U_g, $g \in \mathsf{G}$. At last, $\mathsf{T}_e\mathsf{G} \ni T \mapsto -iA_U(T)\restriction_{\mathscr{D}_N}$ is a representation of $\mathsf{T}_e\mathsf{G}$, because it is linear and satisfies (12.88) [BaRa86].

An important relationship exists between analytic vectors in $\mathscr{D}_N$ and analytic vectors according to Chapter 9. Nelson proved the following important result [BaRa86], which implies that $\mathscr{D}_N$ is dense in H, as we said, because analytic vectors for a self-adjoint operator are dense by Proposition 9.21(f). (An operator is introduced, called *Nelson operator*, that sometimes has to do with the *Casimir operators* [BaRa86] of the represented group.)

Proposition 12.61. *Let* G *be a Lie group and* $\mathsf{G} \ni g \mapsto U_g$ *a strongly continuous unitary representation on the Hilbert space* H. *Take* $T_1, \ldots, T_n \in \mathsf{T}_e\mathsf{G}$ *a basis and define* **Nelson's operator** *on* $\mathscr{D}_G$ *by*

$$\Delta := \sum_{k=1}^{n} A_U(T_k)^2 ,$$

where the $A_U(T_k)$ *are restricted to the domain* $\mathscr{D}_G$. *Then*

(a) Δ *is essentially self-adjoint.*
(b) *Every analytic vector of the self-adjoint* $\overline{\Delta}$ *is analytic for the representation* U *(an element of* $\mathscr{D}_N$*).*
(c) *Every analytic vector of the self-adjoint* $\overline{\Delta}$ *is analytic for every operator* $A_U(T_k)$ *(essentially self-adjoint on* $\mathscr{D}_N$ *by Nelson's criterion)*[7].

We can finally state the famous theorem of Nelson that enables to associate representations of the only simply connected Lie group with a given Lie algebra to representations of that Lie algebra.

Theorem 12.62 (Nelson). *Consider a real n-dimensional Lie algebra V of operators* $-iS$ *(with each S symmetric on the Hilbert space* H*, defined on a common subspace* $\mathscr{D}$ *dense in H and V-invariant) with the usual commutator of operators as Lie bracket. Let* $-iS_1, \cdots, -iS_n \in \mathsf{V}$ *be a basis of* V *and define Nelson's operator with domain* $\mathscr{D}$:

$$\Delta := \sum_{k=1}^{n} S_k^2 .$$

If Δ *is essentially self-adjoint, there exists a strongly continuous unitary representation*

$$\mathsf{G}_\mathsf{V} \ni g \mapsto U_g$$

on H*, of the unique simply connected Lie group* G_V *with Lie algebra* V*, that is completely determined by*

$$\overline{S} = A_U(-iS) \quad \textit{for every } -iS \in \mathsf{V}.$$

In particular, the symmetric operators S are essentially self-adjoint on $\mathscr{D}$*, their closure being self-adjoint.*

The above assumptions were weakened by Flato, Simon, Snellman and Sternheimer [BaRa86]:

Theorem 12.63 (FS^3, Flato, Simon, Snellman, Sternheimer). *Consider a real n-dimensional Lie algebra* V *of operators* $-iS$ *(with each S symmetric on the Hilbert space* H*, defined on a common subspace* $\mathscr{D}$ *dense in H and V-invariant) with the usual commutator of operators as Lie bracket.*
Let $-iS_1, \cdots, -iS_n \in \mathsf{V}$ *be a basis. If the elements of* $\mathscr{D}$ *are analytic vectors for every* S_k*,* $k = 1, \ldots, n$*, then there is a strongly continuous unitary representation*

$$\mathsf{G}_\mathsf{V} \ni g \mapsto U_g$$

[7] Statements (a), (b) form theorem 2 in [BaRa86, Chapter 11.3]. Statement (c) follows from lemma 7 [BaRa86, Chapter 11.2] and from Proposition 9.21(c) of this book.

on H*, of the unique simply connected Lie group* $\mathsf{G_V}$ *with Lie algebra* V*, that is completely determined by:*

$$\overline{S} = A_U(-iS) \quad \text{for every } -iS \in \mathsf{V}\,.$$

In particular, the symmetric operators S are essentially self-adjoint on $\mathscr{D}$*, their closure being self-adjoint.*

Examples 12.64. (1) Consider two families of operators $\mathscr{P}_k$, $\mathscr{X}_k$, $k = 1, 2, \ldots, n$, on a dense $\mathscr{D} \subset \mathsf{H}$ in a Hilbert space, and suppose they are symmetric. Assume they satisfy, on the domain, Heisenberg's commutation relations, seen in Chapter 11 (where we set $\hbar = 1$):

$$[-i\mathscr{X}_h, -i\mathscr{P}_k] = -i\delta_{hk} I \quad k, h = 1, \ldots, n. \tag{12.89}$$

We may add $-iI$ to the generators. Then $-iI, -i\mathscr{X}_1, \ldots, -i\mathscr{X}_n, -i\mathscr{P}_1, \ldots, -i\mathscr{P}_n$ form a basis for the Lie algebra of the Heisenberg group $\mathscr{H}(n)$ on $\mathbb{R}^{2n+1}$ (see the end of Chapter 11). The Heisenberg group is simply connected. Nelson's theorem guarantees that if, on $\mathscr{D}$, the operator:

$$\Delta - I := \sum_{k=1}^{n} \mathscr{X}_k^2 + \sum_{k=1}^{n} \mathscr{P}_k^2$$

is essentially self-adjoint (we should consider Δ, but it is clear that Δ is essentially self-adjoint if and only if $\Delta - I$ is), then there is a unique unitary and strongly continuous representation $\mathscr{H}(n) \ni (\eta, \mathbf{t}, \mathbf{u}) \mapsto H((\eta, \mathbf{t}, \mathbf{u}))$ on H with I, $\overline{\mathscr{X}_h} =: X_h$ and $\overline{\mathscr{P}_h} =: P_h$, $h = 1, \ldots, n$ as (self-adjoint) generators. Therefore if this representation of the *Heisenberg group* is irreducible, by the Stone–von Neumann theorem (Theorem 11.33) there is a unitary transformation from H to $L^2(\mathbb{R}^n, dx)$ mapping X_h and P_h to the usual position and momentum operators of axiom A.5, Chapter 11 (for $n = 3$ and with the obvious generalisation for $n > 3$). An elementary example is to take $n = 1$, $\mathsf{H} = L^2(\mathbb{R}, dx)$, the operator $\mathscr{X}$ seen as multiplication by the coordinate x, $\mathscr{P} := -i\frac{\partial}{\partial x}$, and defining $\mathscr{D}$ to be the Schwartz space $\mathscr{S}(\mathbb{R})$. In this case $\overline{\Delta - I}$ coincides with the Hamiltonian of the harmonic oscillator of Chapter 9. The operator $\Delta - I$ has an eigenvector basis made by Hermite functions (belonging in $\mathscr{S}(\mathbb{R})$), which are a basis of $L^2(\mathbb{R}, dx)$ as well. Hence $\Delta - I$ (and so Δ, by Proposition 9.21) admits a set of analytic vectors (Hermite functions) whose finite combinations are dense in the Hilbert space. By Nelson's criterion $\Delta - I$ is essentially self-adjoint, and we may apply the above result.

(2) We have a result about commuting spectral measures.

Theorem 12.65. *Let* $A : D(A) \to \mathsf{H}$*,* $B : D(B) \to \mathsf{H}$ *be symmetric operators. If there is a dense space* $D \subset D(A^2 + B^2) \cap D(AB) \cap D(BA)$ *on which A and B commute, and where* $A^2 + B^2$ *is essentially self-adjoint, then A and B are essentially self-adjoint on D and the spectral measures of* $\overline{A}$ *and* $\overline{B}$ *commute.*

The proof is an easy consequence of Nelson's Theorem 12.62. ∎

12.2.9 The Peter–Weyl theorem

This section contains part of a general theorem about strongly continuous unitary representations of compact Hausdorff groups: the celebrated *Peter–Weyl theorem*. Compact Lie groups are included, of course, due to their structure of differentiable manifolds. The Peter–Weyl theorem, in its general form [BaRa86], states two remarkable facts that we will prove: strongly continuous unitary representations of compact Hausdorff groups can be split in *orthogonal sums* (even with uncountably many summands) of *irreducible* representations, and strongly continuous irreducible unitary representations are necessarily *finite-dimensional*. Both are far from obvious. In general, namely, a strongly continuous unitary representation of a topological group might be a *direct integral* of strongly continuous irreducible unitary representations (e.g., unitary representations of the Abelian Lie group $\mathbb{R}$); moreover, there could be infinite-dimensional irreducible representations (like for the Lorentz group).
Let us start with a lemma taking care of the finite-dimensional case.

Lemma 12.66. *Let* $\pi : \mathsf{G} \ni g \mapsto U_g$ *be a unitary representation (not necessarily continuous) of the group* G *(even if not topological) on the finite-dimensional Hilbert space* H. *Then* H *decomposes in an orthogonal sum* $\mathsf{H} = \bigoplus_{k=1}^{n} \mathsf{H}_k$ *where for each* $k = 1, 2, \ldots, n$*:*

(i) $U_g(\mathsf{H}_k) \subset \mathsf{H}_k$ *for every* $g \in \mathsf{G}$*;*
(ii) $\pi_k : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}_k}$ *is an irreducible unitary* G*-representation on* H_k .

If $\mathsf{G} \ni g \mapsto U_g$ *is strongly continuous, so are all maps* π_k.

Proof. If π is not irreducible it will have a non-trivial invariant subspace $\hat{\mathsf{H}}_1 \subset \mathsf{H}$, with $0 < dim(\hat{\mathsf{H}}_1) \leq dim(\mathsf{H}) - 1$. Consider the new unitary representation of $\hat{\pi}_1 : \mathsf{G} \ni g \mapsto U_g\restriction_{\hat{\mathsf{H}}_1}$. If this is not irreducible, as above we can find a non-trivial π-invariant $\hat{\mathsf{H}}_2 \subset \hat{\mathsf{H}}_1$ with $0 < dim(\hat{\mathsf{H}}_2) \leq dim(\mathsf{H}) - 2$. The iteration stops after a finite number of steps, since $dim(\mathsf{H}) < +\infty$, and produces an invariant subspace $\mathsf{H}_1 \neq \{0\}$ for which $\pi_1 : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}_1}$ is irreducible. Consider $\mathsf{H}'_2 := \mathsf{H}_1^{\perp}$. By construction $U_g(\mathsf{H}_1^{\perp}) \subset \mathsf{H}_1^{\perp}$, since $z \in \mathsf{H}_1^{\perp}$ and $x \in \mathsf{H}_1$ imply $(U_g z|x) = (z|U_g^* x) = (z|U_{g^{-1}}x) = 0$ ($U_{g^{-1}}x \in \mathsf{H}_1$ by assumption). Hence $\mathsf{H} = \mathsf{H}_1 \oplus \mathsf{H}'_2$ and $\pi'_2 : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}'_2}$ is a unitary G-representation on H'_2. If π'_2 is irreducible we finish, otherwise we iterate to obtain $\mathsf{H}'_2 = \mathsf{H}_2 \oplus \mathsf{H}'_3$, where $\pi_2 : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}_2}$ is irreducible, $\mathsf{H}_2, \mathsf{H}'_3$ are orthogonal to H_1, $\pi(\mathsf{H}'_3) = \pi_2(\mathsf{H}'_3) \subset \mathsf{H}'_3$ and $\pi'_3 : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}'_3}$ is a unitary G-representation on H'_3. By induction the algorithm is finite, and yields $\mathsf{H}'_k = \{0\}$ if $k = n+1$ for n large enough, because every H_k has dimension at least 1, so $\sum_{k=1}^{n} dim(\mathsf{H}_k) \geq n$, but we also have $\sum_{k=1}^{n} dim(\mathsf{H}_k) \leq dim(\mathsf{H}) < +\infty$.

The last claim is immediate, because everything is finite-dimensional. □

Now let us generalise the lemma to infinitely many dimensions for compact Hausdorff groups. The result, part of a more general statement due to Peter and Weyl, makes use of the *Haar measure* of Example 12.29(5) in Theorem 12.30.

Theorem 12.67 (Peter–Weyl theorem, part I). *Let* G *be a compact Hausdorff group and* $\pi : \mathsf{G} \ni g \mapsto U_g \in \mathfrak{B}(\mathsf{H})$ *a strongly continuous unitary representation.*

(a) *If* π *is irreducible then* H *is finite-dimensional:* $dim(\mathsf{H}) < +\infty$.
(b) *If* π *is reducible, it can be decomposed in a sum of strongly continuous, irreducible unitary representations of* G. *That is,* H *is the Hilbert sum of closed (mutually orthogonal) spaces* $\mathsf{H} = \bigoplus_{k\in K} \mathsf{H}_k$, *where for each* $k \in K$*:*

(i) $\mathsf{H}_k \subset \mathsf{H}$ *is finite-dimensional;*
(ii) $U_g(\mathsf{H}_k) \subset \mathsf{H}_k$ *for every* $g \in \mathsf{G}$*;*
(iii) $\pi_k : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}_k}$ *is a strongly continuous and irreducible unitary* G*-representation on* H_k.

Proof. From now on μ_{G} will be the Haar measure of G, which by Theorem 12.30 is bi-invariant and may be chosen so that $\mu_{\mathsf{G}}(\mathsf{G}) = 1$ (since G is compact). The final statement in Theorem 12.30 implies that if $f \in L^1(\mathsf{G}, \mu_{\mathsf{G}})$ and G is compact, then

$$\int_{\mathsf{G}} f(g)d\mu_{\mathsf{G}}(g) = \int_{\mathsf{G}} f(g^{-1})d\mu_{\mathsf{G}}(g)\,, \tag{12.90}$$

to be used later.
(a) For $x \in \mathsf{H}$ define the operator $K_x : \mathsf{H} \to \mathsf{H}$ by asking, for any $z, y \in \mathsf{H}$:

$$(z|K_x y) = \int_{\mathsf{G}} \left(z\left|U_g x\right.\right)(U_g x|y)\, d\mu_{\mathsf{G}}(g)\,. \tag{12.91}$$

As in the proof of Proposition 9.27, using Riesz's representation and the definition of adjoint to bounded operators, K_x is well defined and $K_x \in \mathfrak{B}(\mathsf{H})$. In particular, since U_g is isometric:

$$||K_x y||^2 \le \int_{\mathsf{G}} |(K_x y|U_g x)|\,|(U_g x|y)|d\mu_{\mathsf{G}}(g) \le \int_{\mathsf{G}} ||K_x y||\,||U_g x|||(U_g x|y)|d\mu_{\mathsf{G}}(g)$$

so

$$||K_x y|| \le \int_{\mathsf{G}} ||U_g x|||(U_g x|y)|d\mu_{\mathsf{G}}(g) \le \int_{\mathsf{G}} ||U_g x||||U_g x||||y||d\mu_{\mathsf{G}}(g)$$
$$= ||x||^2||y|| \int_{\mathsf{G}} d\mu_{\mathsf{G}}(g) = ||x||^2||y||\,,$$

and then $||K_x|| \le ||x||^2$. Moreover

$$U_g K_x = K_x U_g \quad \text{for any } x \in \mathsf{H},\ g \in \mathsf{G}\,. \tag{12.92}$$

In fact, U_g is unitary and $U_g U_{g'} = U_{gg'}$, so

$$(z|U_g K_x y) = \left(U_g^* z|K_x y\right) = \int_{\mathsf{G}} \left(U_g^* z|U_{g'} x\right)(U_{g'} x|y)\, d\mu_{\mathsf{G}}(g')$$

$$= \int_{\mathsf{G}} \left(z|U_{gg'} x\right)(U_{g'} x|y)\, d\mu_{\mathsf{G}}(g') = \int_{\mathsf{G}} \left(z|U_{gg'} x\right)(U_g U_{g'} x|U_g y)\, d\mu_{\mathsf{G}}(g')\,.$$

Now as $\mu_G(gA) = \mu_G(A)$ (it is the Haar measure), the last integral becomes

$$\int_G (z|U_{gg'}x)\,(U_{gg'}x|U_g y)\,d\mu_G(g') = \int_G (z|U_{gg'}x)\,(U_{gg'}x|U_g y)\,d\mu_G(gg')$$

$$= \int_G (z|U_s x)\,(U_s x|U_g y)\,d\mu_G(s) = (z|K_x U_g y)\ .$$

But $z \in H$ is arbitrary, so (12.92) holds. With this settled, let us begin to prove (a). If the representation $G \ni g \mapsto U_g \in \mathfrak{B}(H)$ is irreducible, by Proposition 11.16 (Schur's lemma) equation (12.92), valid for every $g \in G$, is valid only if $K_x = \chi(x)I$ for some $\chi(x) \in \mathbb{C}$. Hence

$$\int_G (y|U_g x)\,(U_g x|y)\,d\mu_G(g) = (y|K_x y) = \chi(x)||y||^2\,,$$

$x, y \in H$, and so

$$\int_G |\,(y|U_g x)\,|^2\,d\mu_G(g) = \chi(x)||y||^2\,. \tag{12.93}$$

As $U_g^* = U_{g^{-1}}$, the latter reads

$$\int_G \left|\left(U_{g^{-1}}y\middle|x\right)\right|^2\,d\mu_G(g) = \chi(x)||y||^2$$

or

$$\int_G \left|\left(x\middle|U_{g^{-1}}y\right)\right|^2\,d\mu_G(g) = \chi(x)||y||^2$$

and even, by (12.90),

$$\int_G |(x|U_{g'}y)|^2\,d\mu_G(g') = \chi(x)||y||^2\,.$$

Using (12.93) with x, y swapped allows to conclude the left-hand side equals $\chi(y)||x||^2$, so that $\chi(x)||y||^2 = \chi(y)||x||^2$ irrespective of $x, y \in H$. This means $\chi(x) = c||x||^2$ for any $x \in H$ and some constant $c \geq 0$. Set $x = y$, $||x|| = 1$, so (12.93) gives:

$$\int_G |\,(x|U_g x)\,|^2\,d\mu_G(g) = \chi(x)||x||^2 = c||x||^4 = c\,;$$

hence $c > 0$ because the continuous, non-negative $G \ni g \mapsto |(x|U_g x)|$ reaches $||x|| = 1$ at $g = e$, and non-empty open sets have non-zero Haar measure (Theorem 12.30(ii)). To finish with (a), consider n orthonormal vectors $\{z_k\}_{k=1,\ldots,n} \subset H$. Setting $x = e_k$ and $y = e_1$ in (12.93) gives:

$$\int_G |\,(e_1|U_g e_k)\,|^2\,d\mu_G(g) = \chi(e_k)||e_1||^2 = c > 0\,, \quad k = 1, 2, \ldots$$

By the orthonormality of the $U_g e_k$ and Bessel's inequality (3.17):

$$nc = \sum_{k=1}^{n} \int_G |(e_1|U_g e_k)|^2 d\mu_G(g) = \int_G \sum_{k=1}^{n} |(e_1|U_g e_k)|^2 d\mu_G(g)$$

$$\leq \int_G ||e_1||^2 d\mu_G(g) = 1\,.$$

Whichever $c > 0$, the number n cannot be arbitrarily large: it must be finite. Therefore the dimension of H is finite and not bigger than $1/c$. This concludes (a).
(b) The proof of (b) needs a lemma.

Lemma 12.68. *Consider a closed subspace* $\{0\} \neq \mathsf{H}_1 \subseteq \mathsf{H}$ *such that* $U_g(\mathsf{H}_1) \subset \mathsf{H}_1$ *for* $g \in \mathsf{G}$. *Then there exists a finite-dimensional* $\{0\} \neq \mathsf{H}_0 \subset \mathsf{H}_1$ *such that, for any* $g \in \mathsf{G}$, $U_g(\mathsf{H}_0) \subset \mathsf{H}_0$ *and* $\mathsf{G} \ni g \mapsto \pi\restriction_{\mathsf{H}_0}$ *is an irreducible* G*-representation on* H_0.

Proof of Lemma 12.68. Consider the family $\mathscr{X}$ of sets $\{\mathsf{H}_j\}_{j\in J}$ where J has arbitrary cardinality, each $\mathsf{H}_j \subset \mathsf{H}$ is finite-dimensional, non-null and such that $\pi(\mathsf{H}_j) \subset \mathsf{H}_j$, $\mathsf{H}_j \perp \mathsf{H}_{j'}$ for $j \neq j'$, and $\pi_j : \mathsf{G} \ni g \mapsto U_g\restriction_{\mathsf{H}_j}$ is an irreducible G-representation on H_j. The π_j are certainly strongly continuous since π is. Endow $\mathscr{X}$ with the order relation given by inclusion. Clearly any ordered subset $\mathscr{E} \subset \mathscr{X}$ is upper bounded by the union of elements in $\mathscr{E}$. Zorn's lemma tells we have a maximal element in $\mathscr{X}$. By construction this is $\{\mathsf{H}'_m\}_{m\in M} \in \mathscr{X}$ not properly cointained in any $\{\mathsf{H}_j\}_{j\in J} \in \mathscr{X}$. Now consider the closed Hilbert sum $\mathsf{H}' := \bigoplus_{m\in M} \mathsf{H}'_m$. By construction $U_g(\mathsf{H}') \subset \mathsf{H}'$, because every U_g is continuous. The orthogonal complement $\mathsf{H}'^{\perp}$ is π-invariant, because $x \in \mathsf{H}'^{\perp}$ and $y \in \mathsf{H}'$ imply $(U_g x|y) = (x|U_{g^{-1}}y) = 0$ since $U_{g^{-1}}y \in \mathsf{H}'$, $y \in \mathsf{H}'$. Suppose $\mathsf{H}'^{\perp} \neq \{0\}$. Then $\mathsf{H}'^{\perp}$ contains a finite-dimensional subspace $\mathsf{H}_0 \neq \{0\}$. By construction $\{\mathsf{H}'_m\}_{m\in M} \cup \{\mathsf{H}_0\}$ is in $\mathscr{X}$ and contains the maximal $\{\mathsf{H}'_m\}_{m\in M}$, a contradiction. Therefore $\mathsf{H}'^{\perp} = \{0\}$, i.e. $\mathsf{H} = \bigoplus_{m\in M} \mathsf{H}'_m$, proving (b). □

To finish the proof of part (b) it suffices to prove the next result.

Lemma 12.69. *Let* $\{0\} \neq \mathsf{H}_1 \subseteq \mathsf{H}$ *be a closed subspace such that* $U_g(\mathsf{H}_1) \subset \mathsf{H}_1$, $g \in \mathsf{G}$. *There exists a finite-dimensional* $\{0\} \neq \mathsf{H}_0 \subset \mathsf{H}_1$ *such that, for any* $g \in \mathsf{G}$, $U_g(\mathsf{H}_0) \subset \mathsf{H}_0$ *and* $\mathsf{G} \ni g \mapsto \pi\restriction_{\mathsf{H}_0}$ *is an irreducible* G*-representation on* H_0.

Proof of Lemma 12.69. From (12.91) and the inner product's elementary properties $(K_x z|y) = (z|K_x y)$ for any $x, y, z \in \mathsf{H}$. Since $K_x \in \mathfrak{B}(\mathsf{H})$, we have $K_x^* = K_x$, i.e. K_x is self-adjoint. By (12.91):

$$(x\,|K_x x) = \int_{\mathsf{G}} |(U_g x|x)|^2 \, d\mu_{\mathsf{G}}(g) \geq 0\,.$$

At the same time $|(U_g x|x)|^2 = 1$ if $g = e$, and by continuity $(x\,|K_x x) > 0$ since non-empty open sets have finite measure. Hence $K_x \neq 0$ for any $x \in \mathsf{H}$. Now we claim $K_x \in \mathfrak{B}_2(\mathsf{H})$ (K_x is a Hilbert–Schmidt operator). For this it suffices to show Definition 4.22 applies. If $\{e_k\}_{k\in S}$ indicates a basis in H, a few manipulations give

$$\sum_{k\in F} ||K_x e_k||^2 = \sum_{k\in F} \int_{\mathsf{G}} \left(\int_{\mathsf{G}} (e_k|U_h x)(U_h x|U_g x)(U_g x|e_k) d\mu_{\mathsf{G}}(h) \right) d\mu_{\mathsf{G}}(g)$$

for every finite $F \subset S$. For a given k, the iterated integral coincides with the integral in the product measure, by Fubini–Tonelli: in fact we are integrating a continuous map on a compact set ($\mathsf{G} \times \mathsf{G}$), so a bounded map, and the integration domain has finite measure (1). Swapping the integral and the (finite) sum:

$$\sum_{k\in F} ||K_x e_k||^2 = \int_{\mathsf{G}\times\mathsf{G}} (U_h x|U_g x) \sum_{k\in F} (U_g x|e_k)(e_k|U_h x)\, d\mu_{\mathsf{G}}(h)\otimes d\mu_{\mathsf{G}}(g)\,.$$

Using the obvious upper bounds, from $|(U_h x|U_g x)| \leq ||x||^2$ and Schwarz's inequality:

$$\sum_{k\in F} ||K_x e_k||^2 \leq \int_{\mathsf{G}\times\mathsf{G}} |(U_h x|U_g x)| \sum_{k\in F} |(U_g x|e_k)|\,|(e_k|U_h x)|\, d\mu_{\mathsf{G}}(h)\otimes d\mu_{\mathsf{G}}(g)$$

$$\leq ||x||^2 \int_{\mathsf{G}\times\mathsf{G}} \sqrt{\sum_{k\in F} |(e_k|U_h x)|^2}\sqrt{\sum_{k\in F} |(U_g x|e_k)|^2}\, d\mu_{\mathsf{G}}(h)\otimes d\mu_{\mathsf{G}}(g)$$

$$\leq ||x||^2 \int_{\mathsf{G}\times\mathsf{G}} \sqrt{\sum_{k\in S} |(e_k|U_h x)|^2}\sqrt{\sum_{k\in S} |(U_g x|e_k)|^2}\, d\mu_{\mathsf{G}}(h)\otimes d\mu_{\mathsf{G}}(g)$$

$$\leq ||x||^2 \int_{\mathsf{G}\times\mathsf{G}} ||U_g x||\,||U_h x||\, d\mu_{\mathsf{G}}(h)\otimes d\mu_{\mathsf{G}}(g) \leq ||x||^4 \int_{\mathsf{G}\times\mathsf{G}} 1\, d\mu_{\mathsf{G}}(h)\otimes d\mu_{\mathsf{G}}(g)$$

$$= ||x||^4 < +\infty\,.$$

But $F \subset S$ was finite but arbitary, so

$$\sum_{k\in S} ||K_x e_k||^2 \leq ||x||^4 < +\infty$$

and $K_x \in \mathfrak{B}_2(\mathsf{H})$. Since any Hilbert–Schmidt operator, like K_x, is compact, and at present $K_x = K_x^*$, we invoke Hilbert's Theorems 4.17 and 4.18 to decompose H in a Hilbert sum of eigenspaces $\mathsf{H}_\lambda^{(x)}$ of K_x:

$$\mathsf{H} = \bigoplus_{\lambda\in\sigma_p(K_x)} \mathsf{H}_\lambda^{(x)}\,.$$

Each summand, with the possible exclusion of $\mathsf{H}_0^{(x)}$, has finite dimension. Since $K_x \neq 0$, by theorem (4.18(a)) there is an eigenvalue $\lambda_1 \neq 0$. By (12.92) every eigenspace $\mathsf{H}_\lambda^{(x)}$ is π-invariant. Therefore $\mathsf{H}_0 := \mathsf{H}_{\lambda_1}^{(x)}$ satisfies the requests. □

Remark 12.70. (1) Theorem 12.67, actually, applies to a wider class of strongly continuous representations of compact Hausdorff groups. Let $\pi : \mathsf{G} \ni g \mapsto A_g \in \mathfrak{B}(\mathsf{H})$ be a strongly continuous representation of the compact Hausdorff group G, given by bounded (perhaps non-unitary) operators on the Hilbert space H. Using the Haar measure of G, we define the inner product

$$\langle u|v\rangle_{\mathsf{G}} := \int_{\mathsf{G}} (A_g u|A_g v)\,d\mu_{\mathsf{G}}(g)\,, \quad u,v \in \mathsf{H}$$

on H. This (1) is well defined, (2) renders $(\mathsf{H}, \langle\,|\,\rangle_{\mathsf{G}})$ a Hilbert space, (3) makes the norm $||\;||_{\mathsf{G}}$ associated to $\langle\,|\,\rangle_{\mathsf{G}}$ *equivalent* (Definition 2.99) to the norm $||\;||$ of $(\,|\,)$. In addition, $\pi : \mathsf{G} \ni g \mapsto A_g \in \mathfrak{B}(\mathsf{H})$ is a strongly continuous *unitary* representation on the Hilbert space $(\mathsf{H}, \langle\,|\,\rangle_{\mathsf{G}})$.

(2) We will state, and not prove, the remaining part of the Peter–Weyl theorem [BaRa86]. In the sequel, given a compact Hausdorff group G we will denote by $\{T^s\}_{s\in S}$ the family of *irreducible* representations that are not unitarily equivalent. Said better, in every equivalence class $[T^s]$ of irreducible unitary representations we select a representative T^s acting on the finite-dimensional Hilbert space H_s of dimension d_s common to the entire class $[T^s]$. We will also need to use the **right regular representation** of G, i.e. the strongly continuous unitary representation $R : \mathsf{G} \ni g \mapsto R_g$ acting on $L^2(\mathsf{G}, \mu_{\mathsf{G}})$ by

$$(R_g f)(h) := f(hg)\,, \quad g,h \in \mathsf{G},\, f \in L^2(\mathsf{G}, \mu_{\mathsf{G}}).$$

That R_g is unitary is a consequence of the fact that the Haar measure is bi-invariant, if G is compact.

Theorem 12.71 (Peter–Weyl theorem, part II). *Under the assumptions of Theorem 12.67 the following hold.*

(c) *Let $\{T^s\}_{s\in S}$ be the non-unitarily equivalent irreducible representations of G. Consider an orthonormal basis $\{\phi^s_k\}_{k=1,\dots dim(\mathsf{H}_s)}$ of H_s for every s, and the corresponding matrix elements $D^s(g)_{ij} = (\phi_i|T^s\phi_j)$. Then the functions*

$$\mathsf{G} \ni g \mapsto \sqrt{d_s} D^s(g)_{ij} \in \mathbb{C}\,, \quad s \in S,\ i,j \in \{1,2,\dots,dim(\mathsf{H}_s)\}$$

form a basis of $L^2(\mathsf{G}, \mu_{\mathsf{G}})$ and finitely span a dense space in $C(\mathsf{G})$ for $||\ ||_\infty$ (so in each $L^p(\mathsf{G}, \mu_{\mathsf{G}})$).

(d) *$L^2(\mathsf{G}, \mu_{\mathsf{G}})$ decomposes in a Hilbert sum of finite-dimensional subspaces that are invariant and irreducible under the right regular representation R of G. Furthermore:*

(i) *on each summand, the unitary matrices of every subrepresentation are the $D^s(g)$ of (c);*
(ii) *up to unitary equivalence, every irreducible unitary G-representation shows up in the R-decomposition into irreducible subrepresentations;*
(iii) *every irreducible unitary representation T^s has multiplicity d_s in the decomposition.* ■

12.3 Examples

In this section we discuss a few important examples of the theory we have developed.

12.3.1 The symmetry group $SO(3)$ and the spin

We now concentrate on unitary representations of the compact Lie group $SU(2)$, seen as the universal covering of $SO(3)$ (Example 12.29(2)). With the aid of Bargmann's theorem and Proposition 12.54 (see Example 12.59(2) as well), unitary $SU(2)$-representations will be used to define an action of the Lie group $SO(3)$ – by a continuous projective representation – on the physical system made by a particle of spin s.

From Theorem 12.67 unitary $SU(2)$-representations are direct sums of irreducible and finite-dimensional unitary representations. In the sequel we describe them.

Until now we discussed the quantum system of a particle on the Hilbert space $L^2(\mathbb{R}^3, dx)$ (fixing an inertial frame $\mathscr{I}$ that identifies $\mathbb{R}^3$ with the rest space, with a right-handed triple of Cartesian axes). Experience shows this description is not physically adequate: $L^2(\mathbb{R}^3, dx)$ is not always good enough to account for the physical structure of real particles. The latter possess a feature, called *spin*, determined by an associated constant s, just like the mass is attached to the particle; this constant may only be integer or semi-integer $s = 0, 1/2, 1, 3/2, \ldots$.

Having a spin means, physically, the particle possesses an *intrinsic angular momentum* [Mes99, CCP82], and there are observables, not representable by the fundamental position and momentum, that describe the intrinsic angular momentum. Let us summarise the mathematics involved, referring to [Mes99, CCP82] for a sweeping physics debate on this crucial topic.

If a particle has spin $s = 0$, the description is the usual for spin-zero particles. If $s = 1/2$, the particle's Hilbert space is larger and in fact is the tensor product $L^2(\mathbb{R}^3, dx) \otimes \mathbb{C}^2$, where $\mathbb{C}^2$ (seen as Hilbert space) is the **spin space**. The three **spin operators** are the Hermitian matrices (for the moment we use the constant value $\hbar$, only to set it to one subsequently for simplicity) $S_k := \frac{\hbar}{2}\sigma_k$, $k = 1, 2, 3$ and the σ_k are the *Pauli matrices* seen earlier. Thus the commutation relations:

$$[-iS_i, -iS_j] = \hbar \sum_{k=1}^{3} \varepsilon_{ijk}(-iS_k) \tag{12.94}$$

hold. The associated observables are the components of the particle's intrinsic angular momentum in the given inertial reference system. For $s = 1/2$ the possible values of each component are $-\hbar/2$ and $\hbar/2$, since the eigenvalues of a Pauli matrix are ± 1.

For generic spin s the description is similar, but the spin space is now $\mathbb{C}^{2s+1}$. There the matrices S_k of the spin operators, replacing $\frac{\hbar}{2}\sigma_k$, are Hermitian, satisfy (12.94) and have $2s+1$ eigenvalues $-\hbar s, -\hbar(s-1), \ldots, \hbar(s-1), \hbar s$ of multiplicity 1. For $m, m' = s, s-1, \ldots, -s+1, -s$, here is what they look like, explicitly:

$$(S_1)_{m'm} = \frac{\hbar}{2}\left(\sqrt{(s-m)(s+m+1)}\,\delta_{m',m+1} + \sqrt{(s+m)(s-m+1)}\,\delta_{m',m-1}\right),$$
$$(S_2)_{m'm} = \frac{\hbar}{2i}\left(\sqrt{(s-m)(s+m+1)}\,\delta_{m',m+1} - \sqrt{(s+m)(s-m+1)}\,\delta_{m',m-1}\right),$$
$$(S_3)_{m'm} = m\hbar\,\delta_{m',m}\,.$$

For the recipe to construct the S_k and a deeper analysis of the spin we suggest consulting [Mes99, CCP82]. Here we just make three comments.

(a) The operator $S^2 := \sum_{k=1}^{3} S_k^2$ satisfies

$$S^2 = \hbar^2 s(s+1) I$$

where $I : \mathbb{C}^{2s+1} \to \mathbb{C}^{2s+1}$ is the identity matrix.

(b) The space $\mathbb{C}^{2s+1}$ is *irreducible* under the $SU(2)$-representation given by exponentiating $-iS_k$:

$$V^s : SU(2) \ni e^{-i\theta \frac{\hbar}{2}\mathbf{n}\cdot\sigma} \mapsto e^{-i\theta\mathbf{n}\cdot\mathbf{S}}\,. \tag{12.95}$$

For $s = 0, 1/2, 1, 3/2, \ldots$ and up to unitary equivalence, the V^s produce *every* irreducible finite-dimensional unitary $SU(2)$-representation.
(c) The matrix S_3 is chosen so to coincide with

$$\hbar \,\mathrm{diag}(s, s-1, \ldots, -s+1, -s).$$

Typically the eigenvector basis of S_3, i.e. the canonical basis of $\mathbb{C}^{2s+1}$, is denoted $\{|s, s_3\rangle\}_{|s_3|\leq s}$. Pure states $\Psi(\Psi|\ )$ are thus determined by a collection of $2s+1$ wavefunctions ψ_{s_3} in $L^2(\mathbb{R}^3, dx)$ with unit norm, and therefore a pure state is given by a unit vector

$$\Psi = \sum_{|s_3|\leq s} \psi_{s_3} \otimes |s, s_3\rangle .$$

By this $L^2(\mathbb{R}^3, dx) \otimes \mathbb{C}^{2s+1}$ becomes naturally isomorphic to the orthogonal sum of $2s+1$ copies of $L^2(\mathbb{R}^3, dx)$, so Ψ is identified with a column vector

$$\Psi \equiv (\psi_s, \psi_{s-1}, \cdots, \psi_{-s+1}, \psi_{-s})^t$$

of wavefunctions. In QM's jargon these are called **spinors** of dimension s.

If s is an integer, the representation $SU(2) \ni e^{-i\theta \frac{1}{2}\mathbf{n}\cdot\sigma} \mapsto e^{-i\frac{\theta \mathbf{n}\cdot\mathbf{S}}{\hbar}}$ on $\mathbb{C}^{2s+1}$, associated to the spin matrices, is a faithful $SO(3)$-representation, since the kernel of the covering map $SU(2) \to SO(3)$ consists of the identity I and $-I$. If s is half an integer, instead, the above is a faithful $SU(2)$-representation.

One last important remark on the construction of the observables S_k and the relative irreducible $SU(2)$-representations, found in all QM manuals and based on the commutation relations of the S_k only, is the following. The purely algebraic construction works *because we assume the observables S_k are defined on the whole Hilbert space, and have discrete spectrum*. This is theoretically not obvious, and is merely due to the *finite-dimensional* ambient one works in, so the operators S_k are Hermitian matrices. This is guaranteed by the Peter–Weyl theorem, provided one uses *irreducible unitary representations* of a *compact* group like $SU(2)$. The same procedure would not work as well with non-compact groups such as the Lorentz group.

This is the point where we start setting $\hbar = 1$ to simplify notations. We discuss the relationship between total angular momentum and $SU(2)$, or the rotation group $SO(3)$. For a particle of spin s let

$$\mathscr{J}_k = \mathscr{L}_k \otimes I + I \otimes S_k \tag{12.96}$$

be the **(total) angular momentum operators** on $\mathsf{H} = L^2(\mathbb{R}^3, dx) \otimes \mathbb{C}^{2s+1}$. The *orbital angular momentum operators* $\mathscr{L}_k$, defined in (10.38) and discussed in Chapter 10, have as closure the observables associated to the components of the orbital angular momentum. Above, the first I denotes the identity operator on $\mathbb{C}^{2s+1}$ and the second the identity on $L^2(\mathbb{R}^3, dx)$. The domain is the linear invariant space $\mathscr{D} := \mathscr{S}(\mathbb{R}^3) \otimes \mathbb{C}^{2s+1}$. By construction these operators satisfy the bracket relations defining the Lie algebra $so(3)$:

$$[-i\mathscr{J}_i, -i\mathscr{J}_j] = \sum_{k=1}^{3} \varepsilon_{ijk}(-i\mathscr{J}_k) . \tag{12.97}$$

We wish to apply Nelson's Theorem 12.62 to the Lie algebra spanned by the operators $\mathscr{J}_k$. Consider the symmetric operator

$$\mathscr{J}^2 = \sum_{k=1}^{3} (\mathscr{L}_k \otimes I + I \otimes S_k)^2$$

defined on $\mathscr{D}$. It admits an eigenvector basis obtained from the basis of H of products

$$|l,m,s_z,n\rangle := Y^l_m \psi_n \otimes |s,s_z\rangle \in \mathscr{D} \subset L^2(\mathbb{R}^3,dx) \otimes \mathbb{C}^{2s+1}$$

with $l = 0,1,2\ldots,, m = -l,-l+1,\ldots,l-1,l, n = 0,1,2,\ldots, s_z = -s,-s+1,\ldots,s-1,s$, and the $|s,s_z\rangle \in \mathbb{C}^{2s+1}$ are unit eigenvectors of S_3 relative to s_z. As S_3 is Hermitian, the $2s+1$ vectors $|s,s_z\rangle$ are an orthonormal basis in $\mathbb{C}^{2s+1}$. $L^2(\mathbb{R}^3,dx)$ has a basis made by the $Y^l_m\psi_n$ of (10.48), Chapter 10. Proposition 10.25 ensures the $Y^l_m\psi_n \otimes |s,s_z\rangle$ form a basis for the product space. The $|l,m,s_z,n\rangle$ are not eigenvectors of $\mathscr{J}^2$. The purely algebraic *Clebsch-Gordan procedure*[8] [Mes99, CCP82] shows how to build, out of finite combinations of vectors $|l,m,s_z,n\rangle$, an eigenvector basis

$$|j,j_3,l,n\rangle$$

for $\mathscr{J}^2$, $\mathscr{J}_z$, $\mathscr{L}^2$, where $|l+s| \geq j \geq |l-s|$, $l = 0,1,2,\ldots$ $j_3 = -j,-j+1,\ldots,j+1,j$, $n = 0,1,2,\ldots$ (the j implicitly differ by integers). Then

$$\mathscr{J}^2|j,j_3,l,n\rangle = j(j+1)|j,j_3,n\rangle\,, \quad \mathscr{J}_3|j,j_3,n\rangle = j_z|j,j_3,n\rangle\,,$$

$$\mathscr{L}^2|j,j_3,n\rangle = l(l+1)|j,j_3,n\rangle\,.$$

The $|j,j_3,l,n\rangle$ belong in $\mathscr{D}$ being finite combinations of $|l,m,s,s_z,n\rangle$. As eigenvectors, they are analytic vectors for $\mathscr{J}^2$. Nelson's criterion tells $\mathscr{J}^2$ is essentially self-adjoint on $\mathscr{D}$. Then there exists a strongly continuous unitary $SU(2)$-representation on H, by Nelson's theorem, the generators of which are the self-adjoint operators $J_k := \overline{\mathscr{J}_k} = \overline{\mathscr{L}_k} \otimes I + I \otimes S_k$. (Notice $\overline{\mathscr{L}_k} \otimes I = \overline{\mathscr{L}_k \otimes I}$ since I is in that case defined on a finite-dimensional space.)

In the exercises we will show the strongly continuous unitary representation obtained when exponentiating the J_k, if $s = 0$, is an $SO(3)$-representation, and coincides with the known one from Example 12.17(1), where $\Gamma \in IO(3)$ specialises to $\Gamma = R \in SO(3)$. (The representation is strongly continuous owing to Example 12.37(1).) This fact easily implies (see exercises), when $s \neq 0$, that the $SU(2)$-representation arising by exponentiating the generators J_k as in Nelson's theorem has the form:

$$SU(2) \ni e^{-i\theta\frac{1}{2}\mathbf{n}\cdot\sigma} \mapsto e^{-i\theta\mathbf{n}\cdot\mathbf{J}} = e^{-i\theta\mathbf{n}\cdot\mathbf{L}} \otimes V^s\left(e^{-i\frac{\theta}{2}\mathbf{n}\cdot\sigma}\right) \tag{12.98}$$

where $L_k := \overline{\mathscr{L}_k}$ is the self-adjoint operator associated to the kth component of the orbital angular momentum, as in Chapter 9. Furthermore

$$\left(e^{-i\theta\mathbf{n}\cdot\mathbf{L}}\psi\right)(\mathbf{x}) = \psi\left(e^{\theta\mathbf{n}\cdot\mathbf{T}}\mathbf{x}\right)\,, \tag{12.99}$$

[8] Back when the author was an undergraduate, the procedure was impertinently known among students by the cheeky name of computation of "Flash Gordon coefficients".

where

$$SU(2) \ni e^{-i\theta\frac{1}{2}\mathbf{n}\cdot\sigma} \mapsto e^{-\theta\mathbf{n}\cdot\mathbf{T}} \in SO(3)$$

is the covering map $SU(2) \to SO(3)$ mentioned in Example 12.52(6).

Remarks 12.72. *Because of Proposition 12.54, the physical assumption is that the projective $SO(3)$-representation induced by the unitary $SU(2)$-representation (12.98) corresponds to the $SO(3)$ action on the spin-s particle, when viewing $SO(3)$ as symmetry group of the system.* ■

12.3.2 The superselection rule of the angular momentum

We consider a generic quantum system admitting a continuous projective representation of the rotation group $SO(3)$ illustrating the physical effect of rotating states. We may view the representation as a strongly continuous unitary $SU(2)$-representation by Bargmann's theorem and Proposition 12.54. Using Peter–Weyl we conclude the system's Hilbert space decomposes in a sum $\mathsf{H} = \bigoplus_{s\in A}\mathsf{H}_s$ of closed orthogonal spaces H_s, on which irreducible, hence finite-dimensional, unitary representations of $SU(2)$ act. Each such is unitarily equivalent to one V^s of the previous section, where now $s(s+1)$ will not correspond to the spin squared of a particle, but rather to the squared eigenvalue of the total angular momentum J^2 on V^s, including orbital and spin components. From the previous section the parameter s can only be integer or semi-integer, $s = 0, 1/2, 1, 3/2, 2, \ldots$, so the index set A cannot be larger than the set of those values. Suppose the A of our physical system contains either type of values. Let J_3 be the self-adjoint generator of rotations about the z-axis, however fixed, corresponding to the component of the total angular momentum along z by definition. Consider a pure state, given by $\Psi = \psi_s + \psi_{s'} \in \mathsf{H}_s + \mathsf{H}_{s'}$, with s integer and s' semi-integer. Irrespective of the axis x_3, remembering the expression for S_3 of the previous section:

$$e^{-i2\pi J_3}\Psi = e^{-i2\pi S_3^{(s)}}\psi_s + e^{-i2\pi S_3^{(s')}}\psi_{s'} = \psi_s - \psi_{s'} \neq \Psi\,.$$

This is physically nonsense, for it says that a 2π revolution about an axis alters the pure state $\Psi(\Psi|\ )$. Therefore, when A contains both integers and semi-integers, we need to assume a **superselection rule for the angular momentum** that forbids coherent superpositions of pure states with total angular momentum (the s giving the irreducible $SU(2)$-representations) partly integer and partly semi-integer. As remarked in Chapter 7.4.5, a pure state can have *undefined* angular momentum, when the state's vector is a combination of vectors corresponding to pure states with different angular momenta. The superselection rule, however, forces the values to be all either integer or semi-integer.

12.3.3 The Galilean group and its projective unitary representations

In classical physics the transformations of orthonormal Cartesian coordinates of two inertial frames $\mathscr{I}$, $\mathscr{I}'$ are elements of the **Galilean group** $\mathscr{G}$. In this sense Galilean transformations are *passive* transformations. With the obvious notation we can write them as:

$$\begin{cases} t' = t + c\,, \\ x'_i = c_i + tv_i + \sum_{j=1}^{3} R_{ij} x_j\,, \quad \text{i=1,2,3} \end{cases} \tag{12.100}$$

where $c \in \mathbb{R}$ (*not* the speed of light!), $c_i \in \mathbb{R}$ and $v_i \in \mathbb{R}$ are any constants, and the numbers R_{ij} define a matrix $R \in O(3)$. Every element of $\mathscr{G}$ is then given by four quantities $(c, \mathbf{c}, \mathbf{v}, R) \in \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times O(3)$. Composing Galilean transformations rephrases as

$$(c_2, \mathbf{c_2}, \mathbf{v_2}, R_2) \cdot (c_1, \mathbf{c_1}, \mathbf{v_1}, R_1) = (c_1 + c_2,\, R_2\mathbf{c_1} + c_1\mathbf{v_2} + \mathbf{c_2},\, R_2\mathbf{v_1} + \mathbf{v_2},\, R_2R_1)\;. \tag{12.101}$$

This composition law turns $\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times O(3)$ into a group, the **Galilean group**. In particular, the neutral element is $(0, \mathbf{0}, \mathbf{0}, I)$ and the inverse :

$$(c, \mathbf{c}, \mathbf{v}, R)^{-1} = (-c, R^{-1}(c\mathbf{v} - \mathbf{c}), -R^{-1}\mathbf{v}, R^{-1})\,. \tag{12.102}$$

We may interpret Galilean transformations as *active* transformations, that actively move spacetime events seen as column vectors $(\mathbf{x}, t)^t$ of Cartesian coordinates (orthonormal, right-handed) in an inertial frame of reference fixed once and for all.

$\mathscr{G}$ acts by matrix multiplication if we identify the generic $(c, \mathbf{c}, \mathbf{v}, R) \in \mathscr{G}$ with the real 5×5 matrix:

$$\begin{bmatrix} R & \mathbf{v} & \mathbf{c} \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \tag{12.103}$$

and the columns $(\mathbf{x}, t)^t \in \mathbb{R}^4$ with $(\mathbf{x}, t, 1)^t \in \mathbb{R}^5$. In this way $\mathscr{G}$ becomes a Lie subgroup of $GL(5, \mathbb{R})$ (the analytic structure coincides with the one inherited from $\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times O(3)$).

In the sequel we shall reduce to the so-called **restricted Galilean group** $S\mathscr{G}$, the connected Lie subgroup where R has positive determinant, i.e. $R \in SO(3)$. We will not consider the *inversion of parity*, which is known not to be always a symmetry and must be treated separately, at least at a quantum level.

The universal covering $\widetilde{S\mathscr{G}}$, arises by replacing $SO(3)$ with $SU(2)$ (real Lie group of dimension 3 inside $GL(4, \mathbb{R})$). As a matter of fact $\widetilde{S\mathscr{G}}$ is diffeomorphic to $\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \times SU(2)$ with product

$$(c_2, \mathbf{c_2}, \mathbf{v_2}, U_2) \cdot (c_1, \mathbf{c_1}, \mathbf{v_1}, U_1) = (c_1 + c_2,\, R(U_2)\mathbf{c_1} + c_1\mathbf{v_2} + \mathbf{c_2},\, R(U_2)\mathbf{v_1} + \mathbf{v_2},\, U_2U_1)\;, \tag{12.104}$$

where $SU(2) \ni U \mapsto R(U) \in SO(3)$ is the covering homomorphism of Example 12.52(6) (see also the exercises). This Lie group is the universal covering of $S\mathscr{G}$, being simply connected (as product of simply connected spaces) and having the Lie algebra of $S\mathscr{G}$.

An interesting basis, in physics, of the Lie algebra of $\widetilde{\mathscr{SG}}$ is given by the 10 vectors

$$-\mathbf{h}\,,\mathbf{p}_i\,,\mathbf{j}_i\,,\mathbf{k}_i \quad \text{i=1,2,3,} \tag{12.105}$$

(note the conventional $-$ sign in the first one), where:

(i) $-\mathbf{h}$ generates the one-parameter subgroup $\mathbb{R} \ni c \mapsto (c, \mathbf{0}, \mathbf{0}, I)$ of **time translations**;
(ii) the three $\mathbf{p}_i$ span the Abelian subgroup $\mathbb{R}^3 \ni \mathbf{c} \mapsto (0, \mathbf{c}, \mathbf{0}, I)$ of **space translations**;
(iii) the three $\mathbf{j}_i$ span the subgroup $SO(3) \ni R \mapsto (0, \mathbf{0}, \mathbf{0}, R)$ of **space rotations**;
(iv) the three $\mathbf{k}_i$ generate the Abelian subgroup $\mathbb{R}^3 \ni \mathbf{v} \mapsto (0, \mathbf{0}, \mathbf{v}, I)$ of **pure Galilean transformations**.

The generators obey commutation relations that detect the structure constants:

$$[\mathbf{p}_i, \mathbf{p}_j] = \mathbf{0}\,, \quad [\mathbf{p}_i, -\mathbf{h}] = \mathbf{0}\,, \quad [\mathbf{j}_i, -\mathbf{h}] = \mathbf{0}\,, \quad [\mathbf{k}_i, \mathbf{k}_j] = \mathbf{0}\,,$$

$$[\mathbf{j}_i, \mathbf{p}_j] = \sum_{k=1}^{3} \varepsilon_{ijk}\mathbf{p}_k\,, \quad [\mathbf{j}_i, \mathbf{j}_j] = \sum_{k=1}^{3} \varepsilon_{ijk}\mathbf{j}_k\,, \quad [\mathbf{j}_i, \mathbf{k}_j] = \sum_{k=1}^{3} \varepsilon_{ijk}\mathbf{k}_k\,,$$

$$[\mathbf{k}_i, -\mathbf{h}] = \mathbf{p}_i\,, \quad [\mathbf{k}_i, \mathbf{p}_j] = \mathbf{0}\,. \tag{12.106}$$

The Galilean group is in all likelihood the most important group of all classical physics, given that classical laws are invariant under the active action of this group. Galilean invariance is a way to express the equivalence of all inertial frame systems, interpreting passively the group transformations. We expect the restricted Galilean group, *seen as group of active transformations from now on*, to be a symmetry group for any quantum system, at least in low-speed regimes (compared to the speed of light, when relativistic effects are petty).

Projective unitary $\mathscr{SG}$-representations describing the action of the symmetry group $\mathscr{SG}$ on a physical system are well understood (see [Mes99, CCP82], for example). To start discussing them, take a physical system given by a particle of spin s (cf. previous section) and mass $m > 0$, not subject to any forces. Fix an inertial frame system $\mathscr{I}$ with right-handed orthonormal Cartesian coordinates that identify the rest space with $\mathbb{R}^3$. The system's Hilbert space H is $L^2(\mathbb{R}^3, dx) \otimes \mathbb{C}^{2s+1}$. Pure states are wavefunctions with spin:

$$\sum_{|s_3| \leq s} \psi_{s_3} \otimes |s, s_3\rangle\,.$$

The wavefunctions $\widetilde{\psi} \in L^2(\mathbb{R}^3, dk)$ are given in *momentum representation*, and are images under the unitary Fourier-Plancherel transform (cf. Chapter 3)

$$\widehat{\mathscr{F}} : L^2(\mathbb{R}^3, dx) \to L^2(\mathbb{R}^3, dk)$$

of wavefunctions ψ in *position representation*: $\widetilde{\psi} = \widehat{\mathscr{F}}\psi$. In particular (Proposition 5.31), the observable momentum P_j is given on $L^2(\mathbb{R}^3, dk)$ by the operator $\widetilde{P}_j = \widehat{\mathscr{F}} P_j \widehat{\mathscr{F}}^{-1}$, i.e. by the multiplication by $\hbar k_j$ on $L^2(\mathbb{R}^3, dk)$. From now on we set $\hbar = 1$ for simplicity. Assume $s = 0$ for a moment. In this representation of H, the action of each element of $S\mathscr{G}$ is induced by a unitary operator $\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)}$ that, up to a phase, possibly depending on $(c,\mathbf{c},\mathbf{v},U)$, is defined as:

$$\left(\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)}\widetilde{\psi}\right)(\mathbf{k}) := e^{i(c\mathbf{v}-\mathbf{c})\cdot(\mathbf{k}-m\mathbf{v})}e^{i\frac{c}{2m}(\mathbf{k}-m\mathbf{v})^2}\widetilde{\psi}\left(R(U)^{-1}(\mathbf{k}-m\mathbf{v})\right). \quad (12.107)$$

When $s \neq 0$, the unitary transformations $\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)}$ are replaced by

$$\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)} \otimes V^s(U), \quad (12.108)$$

where V^s was introduced in (12.95).

Back in position representation, i.e. viewing pure states of a spin-zero particle as unit vectors in $L^2(\mathbb{R}^3, dx)$, the unitary operators $\widetilde{Z^{(m)}}_g$ correspond to unitary operators $Z^{(m)}_g := \widehat{\mathscr{F}}^{-1}\widetilde{Z^{(m)}}_g\widehat{\mathscr{F}}$ under the Fourier-Plancherel transform. In the sequel we will use the two representations without distinction, even though the explicit action of $Z^{(m)}_g$ in position representation will have to wait until the next chapter.

Remark 12.73. (1) Let us evaluate the action on $(c,\mathbf{c},\mathbf{v},U)^{-1}$ rather than $(c,\mathbf{c},\mathbf{v},U)$, for this is more illuminating

$$\left(\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)^{-1}}\widetilde{\psi}\right)(\mathbf{k}) := e^{i\mathbf{c}\cdot(R(U)\mathbf{k}+m\mathbf{v})}e^{-i\frac{c}{2m}(R(U)\mathbf{k}+m\mathbf{v})^2}\widetilde{\psi}\left(R(U)\mathbf{k}+m\mathbf{v}\right). \quad (12.109)$$

To give a meaning to this, decompose $(c,\mathbf{c},\mathbf{v},U)^{-1}$ into

$$(c,\mathbf{c},\mathbf{v},U)^{-1} = (0,\mathbf{0},\mathbf{0},U)^{-1}\cdot(0,\mathbf{0},\mathbf{v},I)^{-1}\cdot(0,\mathbf{c},\mathbf{0},I)^{-1}\cdot(c,\mathbf{0},\mathbf{0},I)^{-1},$$

and let us examine the single actions one by one. From the right

$$\left(\widetilde{Z^{(m)}}_{(c,0,\mathbf{0},I)^{-1}}\widetilde{\psi}\right)(\mathbf{k}) = e^{-i\frac{c}{2m}\mathbf{k}^2}\widetilde{\psi}(\mathbf{k}).$$

In the next chapter we will see that multiplying by the phase $e^{-i\frac{c}{2m}\mathbf{k}^2}$ corresponds to rewinding by a time lapse c. Taking in also the second one,

$$\left(\widetilde{Z^{(m)}}_{(0,\mathbf{c},\mathbf{0},I)^{-1}\cdot(c,\mathbf{0},\mathbf{0},I)^{-1}}\widetilde{\psi}\right)(\mathbf{k}) = e^{i\mathbf{c}\cdot\mathbf{k}}e^{-i\frac{c}{2m}\mathbf{k}^2}\widetilde{\psi}(\mathbf{k}).$$

The multiplication by $e^{i\mathbf{c}\cdot\mathbf{k}}$ corresponds (under Fourier-Plancherel) to an active translation by $-\mathbf{c}$ of the wavefunction. Subsuming the third one, we obtain

$$\left(\widetilde{Z^{(m)}}_{(0,\mathbf{0},\mathbf{v},I)^{-1}\cdot(0,\mathbf{c},\mathbf{0},I)^{-1}\cdot(c,\mathbf{0},\mathbf{0},I)^{-1}}\widetilde{\psi}\right)(\mathbf{k}) = e^{i\mathbf{c}\cdot(\mathbf{k}+m\mathbf{v})}e^{-i\frac{c}{2m}(\mathbf{k}+m\mathbf{v})^2}\widetilde{\psi}(\mathbf{k}+m\mathbf{v}).$$

If $\mathbf{k}$ is understood as momentum vector, $\mathbf{k} \to \mathbf{k}+m\mathbf{v}$ is precisely the transformation of momentum under a Galilean transformation that changes the velocity of the frame of reference, but does not contain space or time translations, nor rotations. The transformation corresponds to an active transformation of the wavefunction under a pure Galilean transformation associated to $-\mathbf{v}$. Eventually, incorporating the rotation $R(U)$, i.e. actively transforming the wavefunction by $R(U)^{-1}$, gives:

$$\left(\widetilde{Z^{(m)}}_{(0,\mathbf{0},\mathbf{0},U)^{-1}\cdot(0,\mathbf{0},\mathbf{v},I)^{-1}\cdot(0,\mathbf{c},\mathbf{0},I)^{-1}\cdot(c,\mathbf{0},\mathbf{0},I)^{-1}}\widetilde{\psi}\right)(\mathbf{k}) =$$

$$e^{i\mathbf{c}\cdot(R(U)\mathbf{k}+m\mathbf{v})}e^{-i\frac{c}{2m}(R(U)\mathbf{k}+m\mathbf{v})^2}\widetilde{\psi}(R(U)\mathbf{k}+m\mathbf{v}) \; .$$

Overall the right-hand side of (12.109) corresponds to the *combined* action (in agreement with the Galilean product) of the subgroups of transformations. Bearing in mind (12.102) our discussion now justifies (12.107).

(2) The operators $\widetilde{Z^{(m)}}_g$ (i.e. the $Z^{(m)}{}_g$, in the position representation) are associated to the universal covering $\widetilde{S\mathscr{G}}$ rather than the group $S\mathscr{G}$ itself. We made this choice to apply the theory of previous sections. We know, in fact, projective representations of a group are obtained from the universal covering's projective representations, and this is particularly convenient because the Galilean group contains a subgroup isomorphic to $SO(3)$. We saw in the previous section that if the spin s is a semi-integer, the projective unitary $SO(3)$-representations of physical interest are unitary $SU(2)$-representations. ∎

Using definition (12.107), the representation $\widetilde{S\mathscr{G}} \ni g \mapsto Z^{(m)}_g$ (equivalently, $\widetilde{S\mathscr{G}} \ni g \mapsto \widetilde{Z^{(m)}}_g$ working in the momentum representation) is *projective* unitary, due to the presence of a multiplier function

$$\omega^{(m)}(g',g) = e^{im\left(-\frac{1}{2}c'\mathbf{v}^2 - c'(R(U')\mathbf{v})\cdot\mathbf{v}' + (R(U')\mathbf{v})\cdot\mathbf{c}'\right)}, \quad g=(c,\mathbf{c},\mathbf{v},U),\, g'=(c',\mathbf{c}',\mathbf{v}',U') \tag{12.110}$$

after a boring computation. The result (clearly) remains valid in case the spin s is non-zero, and the unitary operators $\widetilde{Z^{(m)}_g}$ generalise to the unitary (12.108), because the representation $U \mapsto V^s(U)$ on the spin space $\mathbb{C}^{2s+1}$ is unitary and does not affect the multiplier function.

It is easy to prove the projective unitary representation $\widetilde{S\mathscr{G}} \ni g \mapsto \widetilde{Z^{(m)}}_g$ (equivalently $\widetilde{S\mathscr{G}} \ni g \mapsto Z^{(m)}{}_g$ in the position representation) is strongly continuous. To that end, as operators are unitary, $\omega^{(m)}$ is continuous and $\omega^{(m)}(e,e)=1$, it suffices to prove $\widetilde{Z^{(m)}}_g\widetilde{\psi} \to \widetilde{\psi}$ as $g \to e$, for any $\psi \in \mathsf{H}$. This is an easy consequence of the explicit form of $\widetilde{Z^{(m)}}_g$.

We do not know whether the projective unitary representation $\widetilde{S\mathscr{G}} \ni g \mapsto Z^{(m)}{}_g$ is equivalent to a unitary representation, by multiplying $Z^{(m)}{}_g$ by suitable phases $\chi(g)$. The Galilean Lie algebra shows that Bargmann's Theorem 12.57 does *not* hold. But the aforementioned theorem gives sufficient conditions, not necessary ones, so we are not in a position to answer the question. What we will see now is that the representations found are intrinsically projective: they cannot be made unitary by a clever choice of phase.

In order to keep general, we consider *every* possible projective unitary representation $\widetilde{S\mathscr{G}} \ni g \mapsto Z^{(m)}_g$, on any Hilbert space, with multipliers as in (12.110), but irrespective of the fact the $Z^{(m)}_g$ are as in (12.107) or (12.108) on $L^2(\mathbb{R}^3, dk) \otimes \mathbb{C}^{2s+1}$.

Proposition 12.74. *Let* $\widetilde{\mathscr{SG}} \ni g \mapsto Z_g^{(m)}$ *be projective unitary representations with multipliers (12.110).*

(a) *Given* m*, it is not possible to define the phases of* $Z_g^{(m)}$ *to obtain a unitary* $\widetilde{\mathscr{SG}}$*-representation (neither strongly continuous).*
(b) *Representations with distinct numbers* m *cannot belong to the same unitary equivalence class.*

Proof. We prove (a) and (b) simultaneously. If two representations with $m_1 > m_2$ belonged to the same equivalence class, there would exist a map $\chi = \chi(g)$ such that

$$\omega^{(m_1)}(g', g)\left(\omega^{(m_2)}(g', g)\right)^{-1} = \frac{\chi(g' \cdot g)}{\chi(g')\chi(g)}, \quad g, g' \in \widetilde{\mathscr{SG}}. \tag{12.111}$$

Writing $m := m_1 - m_2 > 0$, this is the same as

$$\omega^{(m)}(g', g) = \frac{\chi(g' \cdot g)}{\chi(g')\chi(g)}, \quad g, g' \in \widetilde{\mathscr{SG}}. \tag{12.112}$$

We claim that for any given $m > 0$ there is no function χ satisfying (12.112), proving the theorem.

By contradiction if such a χ existed, letting $V_g := \chi(g) Z_g^{(m)}$ would force the multipliers of $\widetilde{\mathscr{SG}} \ni g \mapsto V_g$ to be 1, hence the representation would be unitary. Consider the elements in $\widetilde{\mathscr{SG}}$ of the form $f := (0, \mathbf{c}, \mathbf{0}, I)$ and $g := (0, \mathbf{0}, \mathbf{v}, I)$. By (12.101) they commute, so $f^{-1} \cdot g^{-1} \cdot f \cdot g = e$. The corresponding computation for $Z^{(m)}$, keeping (12.107) in account, gives $Z_{f^{-1}}^{(m)} Z_{g^{-1}}^{(m)} Z_f^{(m)} Z_g^{(m)} = e^{-i2m\mathbf{c}\cdot\mathbf{v}} Z_e^{(m)}$. This becomes, with our assumptions:

$$\left(\chi(f^{-1})\chi(g^{-1})\chi(f)\chi(g)\right)^{-1} V_{f^{-1}} V_{g^{-1}} V_f V_g = e^{-i2m\mathbf{c}\cdot\mathbf{v}} \chi(e)^{-1} I;$$

as the multipliers of V are trivial because V is unitary by assumption, we have $f \cdot g = g \cdot f$, and permuting the order of the coefficients χ_h:

$$\begin{aligned}&\left(\chi(f^{-1})\chi(f)\chi(g^{-1})\chi(g)\right)^{-1} V_{f^{-1}\cdot f\cdot g^{-1}\cdot g}\\ &\qquad = \left(\chi(f^{-1})\chi(f)\chi(g^{-1})\chi(g)\right)^{-1} V_e = e^{-i2m\mathbf{c}\cdot\mathbf{v}} \chi(e)^{-1} I.\end{aligned}$$

Therefore

$$\frac{\chi(f^{-1} \cdot f)}{\chi(f^{-1})\chi(f)} \frac{\chi(g^{-1} \cdot g)}{\chi(g^{-1})\chi(g)} = \chi(e) e^{-i2m\mathbf{c}\cdot\mathbf{v}}.$$

Using (12.112) this identity becomes $\omega(f, f^{-1})\omega(g, g^{-1}) = \chi(e) e^{-i2m\mathbf{c}\cdot\mathbf{v}}$. Computing the left-hand side explicitly, with the help of (12.110), yields

$$1 = \chi(e) e^{-i2m\mathbf{c}\cdot\mathbf{v}}.$$

This has to be true for any $\mathbf{c}, \mathbf{v} \in \mathbb{R}^3$, hence $m = 0$ and $\chi(e) = 1$. But $m = 0$ was excluded right from the start. The contradiction invalidates the initial assumption, so χ does not exist. □

By virtue of this proposition, and since the quantity m labelling equivalence classes of projective unitary representations has a very explicit physical meaning, we might think that the symmetry group of a non-relativistic quantum system of mass m, instead of being the Galilean group, is the central extension $\widehat{\widetilde{\mathscr{SG}}}_m$ given by the multiplier function of the value m of the mass. From the general theory the representation $\widetilde{\mathscr{SG}} \ni g \mapsto Z_g^{(m)}$ arises thus: (a) build the central $U(1)$-extension $\widehat{\widetilde{\mathscr{SG}}}_m$, with multiplier function (12.110) ($\widehat{\widetilde{\mathscr{SG}}}_m$ is a product manifold since $\omega^{(m)}$ is analytic on $\widetilde{\mathscr{SG}} \times \widetilde{\mathscr{SG}}$); (b) restrict to $\widetilde{\mathscr{SG}}$ the strongly continuous unitary representation

$$\widehat{\widetilde{\mathscr{SG}}}_m \ni (\chi, g) \mapsto \chi Z_g^{(m)} .$$

On that account (intrinsically) projective unitary $\widetilde{\mathscr{SG}}$-representations are substituted by unitary $\widehat{\widetilde{\mathscr{SG}}}_m$-representations. There is a price to pay: the symmetry group changes when the mass varies. Consider the strongly continuous unitary representation

$$\widehat{\widetilde{\mathscr{SG}}}_m \ni (\chi, g) \mapsto \chi Z_g^{(m)} .$$

Restrict to $\mathscr{D} \subset L^2(\mathbb{R}^3, dk)$, space of smooth complex functions $\widetilde{\psi} = \widetilde{\psi}(\mathbf{k})$ with compact support. By (12.107) every map

$$\widehat{\widetilde{\mathscr{SG}}}_m \ni (\chi, g) \mapsto \chi \widetilde{Z^{(m)}}_g \widetilde{\psi}$$

is smooth whenever $\widetilde{\psi} \in \mathscr{D}$. Hence $\mathscr{D}$ is contained in the Gårding space of $\widehat{\widetilde{\mathscr{SG}}}_m$. With a minor notational misuse we indicate by $\mathscr{D}$ the inverse Fourier-Plancherel image of $\mathscr{D}$ inside $L^2(\mathbb{R}^3, dx)$. Consider the 11 one-parameter Lie subgroups of $\widehat{\widetilde{\mathscr{SG}}}_m$ generated by the Lie algebra basis:

$$1 \oplus \mathbf{0},\ 0 \oplus \mathbf{h},\ 0 \oplus \mathbf{p}_i,\ 0 \oplus \mathbf{j}_i,\ 0 \oplus \mathbf{k}_i\,, \quad i = 1, 2, 3\,.$$

Composing each one with $\widehat{\widetilde{\mathscr{SG}}}_m \ni (\chi, g) \mapsto \chi Z_g^{(m)}$ produces eleven strongly continuous one-parameter unitary groups. Let us find their self-adjoint generators. If we *restrict to* $\mathscr{D}$ when differentiating in the strong topology, the generators are (note the $-$ sign of H):

$$I,\ -H\restriction_{\mathscr{D}},\ P_i\restriction_{\mathscr{D}},\ L_i\restriction_{\mathscr{D}},\ K_i\restriction_{\mathscr{D}}\,, \quad i = 1, 2, 3.$$

P_k and L_k are the self-adjoint operators representing momentum and orbital angular momentum about the kth axis, which we met already. The self-adjoint operators $H := \widehat{\mathscr{F}}^{-1}\widetilde{H}\widehat{\mathscr{F}}$, called **Hamiltonian operator**, and K_i, called **boost** along the ith axis, are defined as:

$$(\widetilde{H}\widetilde{\psi})(\mathbf{k}) := \frac{\mathbf{k}^2}{2m}\widetilde{\psi}(\mathbf{k})$$

$$\text{where} \quad D(\widetilde{H}) := \left\{ \widetilde{\psi} \in L^2(\mathbb{R}^3, dk) \,\middle|\, \int_{\mathbb{R}^3} |\mathbf{k}|^4 |\widetilde{\psi}(\mathbf{k})|^2 dk < +\infty \right\} \quad (12.113)$$

and

$$K_j := -mX_j\,. \tag{12.114}$$

Since $\mathscr{D}$ is a core for all the above, the self-adjoint generators of one-parameter group representations of $\widehat{\widetilde{S\mathscr{G}}}_m$ associated to:

$$1\oplus\mathbf{0},\ 0\oplus\mathbf{h},\ 0\oplus\mathbf{p}_i,\ 0\oplus\mathbf{j}_i,\ 0\oplus\mathbf{k}_i\,,\quad i=1,2,3 \tag{12.115}$$

must concide with the corresponding:

$$I,\ -H,\ P_i,\ L_i,\ K_i\,,\quad i=1,2,3.$$

Each one, as an observable, has a physical meaning. We will talk about the observable H in the next chapter. By considering Lie algebra relations, for instance on $\mathscr{D}$, we realise we are actually working with a central extension of the Galilean group, because one bracket (the last one below) is new: the fault is of a *central charge* that coincides with the mass:

$$[-iP_i,-iP_j]=0\,,\quad [-iP_i,iH]=0\,,\quad [-iL_i,iH]=0\,,\quad [-iK_i,-iK_j]=\mathbf{0}\,,$$

$$[-iL_i,-iP_j]=\sum_{k=1}^{3}\varepsilon_{ijk}(-iP_k)\,,\quad [-iL_i,-iL_j]=\sum_{k=1}^{3}\varepsilon_{ijk}(-iL_k)\,,$$

$$[-iL_i,-iK_j]=\sum_{k=1}^{3}\varepsilon_{ijk}(-iK_k)\,,\quad [-iK_i,iH]=-iP_i\,,\quad [-iK_i,-iP_j]=-im\delta_{ij}I\,.$$

Remark 12.75. (1) Since $K_j=-mX_j$, the unitary representation $(\chi,g)\mapsto\chi\widetilde{Z^{(m)}}_g$ incorporates operators that obey Weyl's relations on $L^2(\mathbb{R}^3,dk)$. By Proposition 11.18(b) $L^2(\mathbb{R}^3,dk)$ is irreducible for these operators, hence for the representation $\widehat{\widetilde{S\mathscr{G}}}_m\ni(\chi,g)\mapsto\chi\widetilde{Z^{(m)}}_g$. In this sense the non-relativistic spin-zero quantum particle is an *elementary object* for the Galilean group.
(2) If we take into account the portion of Hilbert space due to the spin, the difference from above is that to have $\widehat{\widetilde{S\mathscr{G}}}_m$ act on states we must replace L_k by $J_k=L_k+S_k$ in every formula. That is to say, the unitary representation reads

$$\widehat{\widetilde{S\mathscr{G}}}_m\ni(\chi,g)\mapsto\chi\widetilde{Z^{(m)}}_g\otimes V^s(U)\,,$$

where $g=(c,\mathbf{c},\mathbf{v},U)$. The irreducibility seen for the case $s=0$ extends, so for the particle with spin s the above representation is irreducible on $L^2(\mathbb{R}^3,dk)\otimes\mathbb{C}^{2s+1}$. ∎

12.3.4 Bargmann's rule of superselection of the mass

We now consider more complicated systems than a free particle. We refer to the next chapter for the general matter, and recall here that when we study an isolated system of N interacting particles of masses $m_1,\dots,m_N$, the theory's Hilbert space splits:

$$L^2(\mathbb{R}^3,dx)\otimes\mathsf{H}_{int}\otimes\mathbb{C}^{2s_1+1}\otimes\cdots\otimes\mathbb{C}^{2s_N+1}\,.$$

The Hilbert space H_{int} is relative to the system's internal orbital degrees of freedom (the particles mutual positions, for example in terms of *Jacobi coordinates*, e.g. see [AnMo12] for a more explicit construction). $L^2(\mathbb{R}^3, dx)$ is the Hilbert space of the *centre of mass*. The centre of mass of the system is the unique particle of mass $M := \sum_{n=1}^N m_n$ determined by the observables X_k (the position of the centre of mass) and P_k (total momenta of the system), $k = 1,2,3$, of the usual form on $L^2(\mathbb{R}^3, dx)$. Each factor $\mathbb{C}^{2s_n+1}$ is the spin space of one particle. Via Fourier transform $L^2(\mathbb{R}^3, dx)$ can be $L^2(\mathbb{R}^2, dk)$, which we will assume from now on.

In this context – exactly as in classical mechanics – the symmetry group $S\mathscr{G}$ acts by

$$\widetilde{S\mathscr{G}} \ni (c, \mathbf{c}, \mathbf{v}, U) \;\mapsto\; Z^{(M)}_{(c,\mathbf{c},\mathbf{v},U)} \otimes V^{(int)}_{R(U)} W^{(int)}_c \otimes V^{S_1}(U) \otimes \cdots \otimes V^{S_N}(U)\,.$$

Above,

$$SO(3) \ni R \mapsto V^{(int)}_R \quad \text{and} \quad \mathbb{R} \ni c \mapsto W^{(int)}_c$$

are representations – both *unitary* and strongly continuous – of the rotation subgroup of $S\mathscr{G}$ (of elements $(0, \mathbf{0}, \mathbf{0}, R)$), and of time translations (of type $(c, \mathbf{0}, \mathbf{0}, I)$) respectively. In addition, $V^{(int)}_R W^{(int)}_c = W^{(int)}_c V^{(int)}_R$ for every $R \in SO(3)$, $c \in \mathbb{R}$. These two representations depend on how we define orbital coordinates and on the kind of inner interactions among the particles. The transformation $Z^{(M)}_{(c,\mathbf{c},\mathbf{v},U)}$ acts only on the freedom degrees of the centre of mass. Since every representation involved is unitary except $Z^{(M)}$, the multiplier function $\omega^{(M)}$ of the overall representation on $L^2(\mathbb{R}^3, dk) \otimes \mathsf{H}_{int} \otimes \mathbb{C}^{2s_1+1} \otimes \cdots \otimes \mathbb{C}^{2s_N+1}$ is the same we had before, using the total mass M as parameter m. Therefore the previous proposition reaches to this much more general instance of quantum system.

Let us look at a physical system S obtained by putting together a finite number, though *not fixed*, of the previous systems. Or even more generally, we may assume that the value of the mass of S, for some reason, is *not fixed*. The mass of S may thus range over several values m_i, with $i \in I$ at most countable, we assume. It is only natural to associate to the mass a quantum observable, i.e. a self-adjoint operator M whose spectrum is the values of mass (even if all that follows is completely general, explicit models have been constructed in [Giu96, AnMo12]). Likewise, we define a Hilbert space for the system:

$$\mathsf{H}_S = \bigoplus_{m \in \sigma(M)} \mathsf{H}^{(m)}_S\,,$$

where the $\mathsf{H}^{(m)}_S$ are the eigenspaces of the mass operator with distinct eigenvalues $m > 0$. What happens if the Galilean group acts on S? On each $H^{(m)}_S$ a different projective unitary representation $Z^{(m)}$, depending on m, will act. The representation of the restricted Galilean group will thus look like:

$$S\mathscr{G} \ni g \mapsto Z_g := \bigoplus_{m \in \sigma(M)} \chi^{(m)}(g) Z^{(m)}_g\,. \tag{12.116}$$

We claim this structure leads to a superselection rule. Since the representation is projective unitary, the multiplier

$$\Omega(g,g') := Z(gg')^{-1}Z(g)Z(g'),$$

computed with (12.116), produces

$$\Omega(g,g')I = \bigoplus_{m\in\sigma(M)} \omega^{(m)}(g,g')I_m,$$

where the $\omega^{(m)}$ account for possible new phases $\chi^{(m)}$ and the I_m on the right are identities on each $\mathsf{H}_S^{(m)}$. Since

$$I = \bigoplus_{m\in\sigma(M)} I_m,$$

so

$$\Omega(g,g')I = \bigoplus_{m\in\sigma(M)} \Omega(g,g')I_m,$$

necessarily:

$$\omega^{(m_1)}(g,g') = \omega^{(m_2)}(g,g') = \Omega(g,g') \quad \text{for every } m_1, m_2 \in \sigma(M).$$

But this is not possible, because it implies, solving for $\chi^{(m)}$, the false equation (12.111).

The net result is this: if the Galilean group is to be a symmetry group for our physical system, we are compelled to ban pure states arising from combinations in different subspaces $\mathsf{H}_S^{(m)}$. So we have found a superselection rule related to the mass, known as **Bargmann's superselection rule for the mass**. Coherent sectors of this rule are the $\mathsf{H}_S^{(m)}$ with given mass. The result is deeply rooted in the fact that physically-interesting projective representations of the Galilean group do not come from unitary representations, and the mass appears in the multiplier function.

A physically more appropriate situation is that in which one replaces the restricted Galilean group with the (proper orthochronous) Poincaré group: then the superselection rule ceases to be valid, because irreducible projective representations of the Poincaré group always arise from irreducible unitary representations [BaRa86], and states with undefinite (relativistic) mass are allowed.

Remarks 12.76. Since m multiplies the exponent in (12.110), we may introduce a central extension $\mathscr{G}_1$ of the Galilean group (of the universal covering to be precise) that does not depend on m. It is enough to redefine the multiplier setting $m = 1$ in the right-hand side of (12.110). The value of mass is subsequently fixed by a particular unitary representation (raising to the mth power the multiplier and the variables $\chi \in U(1)$) when a physical system is chosen with that mass and that is invariant by the Galilean group. This $\mathscr{G}_1$ should be considered as the quantum version of the Galilean group. This approach, adopted in [Giu96], lets the superselection rule of the mass arise dynamically by enlarging the system with more degrees of freedom, already at the classical level. The mass becomes, *a priori*, a (classical) variabile and defines a self-adjoint operator (the mass operator of the physical system) after quantisation. The approach was improved in [AnMo12], in particular presenting a physical procedure giving rise to the superselection rule once assumed the discreteness of the mass spectrum. ■

Exercises

12.1. Referring to Example 12.17(1), with $IO(3) \ni \Gamma = (\mathbf{t}, R)$, prove

$$\gamma^*_\Gamma(\mathbf{P}) = U_\Gamma^{-1}\mathbf{P}U_\Gamma = R\mathbf{P}\,, \tag{12.117}$$

where $\mathbf{P}$ is the triple of operators corresponding to the components of momentum, and the displayed identity holds on the Schwartz space $\mathscr{S}(\mathbb{R})$ taken as domain of the momenta.

12.2. Referring to Examples 12.17(1) and (2) and retaining the convention of Exercise 12.1, prove:

$$\gamma^*_{\mathscr{P}}(\mathbf{X}) = \mathscr{P}^{-1}\mathbf{X}\mathscr{P} = -\mathbf{X}\,, \quad \gamma^*_{\mathscr{P}}(\mathbf{P}) = \mathscr{P}^{-1}\mathbf{P}\mathscr{P} = -\mathbf{P} \tag{12.118}$$

while

$$\gamma^*_{\mathscr{T}}(\mathbf{X}) = \mathscr{T}^{-1}\mathbf{X}\mathscr{T} = \mathbf{X}\,, \quad \gamma^*_{\mathscr{T}}(\mathbf{P}) = \mathscr{T}^{-1}\mathbf{P}\mathscr{T} = -\mathbf{P}\,. \tag{12.119}$$

$\mathbf{P}$ is the triple corresponding to the momentum's components, and the displayed identities hold on the Schwartz space $\mathscr{S}(\mathbb{R})$ taken as domain of position operators and momenta.

12.3. Consider the self-adjoint operators L_1, L_2, L_3 representing the components of the *orbital angular momentum* (Chapter 10). If $\mathbf{L}$ indicates the relative column vector, then

$$\mathbf{L}\restriction_{\mathscr{S}(\mathbb{R}^3)} = \mathbf{X}\restriction_{\mathscr{S}(\mathbb{R}^3)} \wedge \mathbf{P}\restriction_{\mathscr{S}(\mathbb{R}^3)}\,.$$

Restrict domains to $\mathscr{S}(\mathbb{R}^3)$ and prove the following facts. Referring to Example 12.17(1), with $SO(3) \ni \Gamma = (\mathbf{0}, R)$:

$$\gamma^*_\Gamma(\mathbf{L}) = U_\Gamma^{-1}\mathbf{L}U_\Gamma = R\mathbf{L}\,, \tag{12.120}$$

$$\gamma^*_{\mathscr{P}}(\mathbf{L}) = \mathscr{P}^{-1}\mathbf{L}\mathscr{P} = \mathbf{L}\,, \tag{12.121}$$

$$\gamma^*_{\mathscr{T}}(\mathbf{L}) = \mathscr{T}^{-1}\mathbf{L}\mathscr{T} = -\mathbf{L}\,. \tag{12.122}$$

$SO(3)$ is the subgroup in $O(3)$ with positive determinant $(+1)$, and the wedge product is defined by the above formal determinant in a right-handed basis.

12.4. Decompose the Hilbert space H_S of system S in coherent sectors, so that the space of admissible pure states reads:

$$\mathfrak{S}_p(\mathsf{H}_S)_{adm} = \bigsqcup_{k\in K} \mathfrak{S}_p(\mathsf{H}_{Sk}) .$$

Equip $\mathfrak{S}_p(\mathsf{H}_S)$ with distance $d(\rho,\rho') := ||\rho-\rho'||_1 := tr(|\rho-\rho'|)$, where $||\ ||_1$ is the natural trace-class norm. Prove the sets $\mathfrak{S}_p(\mathsf{H}_{Sk})$ are the connected components of $\mathfrak{S}_p(\mathsf{H}_S)_{adm}$. (It might be useful to recall $\rho = \psi(\psi|\ )$, $\rho' = \psi'(\psi'|\ )$ in $\mathfrak{S}_p(\mathsf{H}_{Sk})$ imply $||\rho-\rho'||_1 = 2\sqrt{1-|(\psi|\psi')|^2}$, as was proved in the chapter).

Sketch of the solution. Consider pure states $\rho,\rho' \in \mathfrak{S}_p(\mathsf{H}_{Sk})$ with $\rho = \psi(\psi|\ )$ and $\rho' = \psi'(\psi'|\ )$, and ψ not parallel to ψ' (otherwise they give the same state). Define $\psi_t = t\psi + (1-t)\psi'$ and the curve $[0,1] \ni t \mapsto \frac{\psi_t}{||\psi_t||^2}(\psi_t|\ )$. Prove the curve is continuous and entirely contained in $\mathfrak{S}_p(\mathsf{H}_{Sk})$, making $\mathfrak{S}_p(\mathsf{H}_{Sk})$ path-connected, hence connected. To prove the $\mathfrak{S}_p(\mathsf{H}_{Sk})$ are disjoint, it is sufficient to find $||\rho-\rho'||_1$ for $\rho \in \mathfrak{S}_p(\mathsf{H}_{Sk})$, $\rho' \in \mathfrak{S}_p(\mathsf{H}_{Sk'})$ with $k \neq k'$. In that case the vectors of ρ, ρ' are orthogonal, so $\rho-\rho'$ is actually the decomposition into positive and negative parts of $\rho-\rho'$ itself. Hence $|\rho-\rho'| = \rho+\rho'$, i.e. $||\rho-\rho'||_1 = 2$. Consider an open set $A_k \supset \mathfrak{S}_p(\mathsf{H}_{Sk})$ union of open balls of radius $1/2$ centred in $\mathfrak{S}_p(\mathsf{H}_{Sk})$, and an open set $A_{k'} \supset \mathfrak{S}_p(\mathsf{H}_{Sk'})$ union of similar balls centred in $\mathfrak{S}_p(\mathsf{H}_{Sk'})$. These two sets cannot intersect by the triangle inequality, so $\mathfrak{S}_p(\mathsf{H}_{Sk})$ and $\mathfrak{S}_p(\mathsf{H}_{Sk'})$ are disjoint.

12.5. Prove the distance $d(\rho,\rho')$ of pure states (Exercise 12.4) satisfies:

$$d\left(\psi(\psi|\ ), \psi'(\psi|\ )\right) = \left|\left|\psi(\psi|\ ) - \psi'(\psi'|\ )\right|\right|_{\mathfrak{B}(\mathsf{H})}$$

for any unit $\psi,\psi' \in \mathsf{H}$, where $||\ ||_{\mathfrak{B}(\mathsf{H})}$ is the standard norm of operators.

12.6. Let $U : \mathsf{H} \to \mathsf{H}$ be an antiunitary operator on the Hilbert space H and $A : D(A) \to \mathsf{H}$ a self-adjoint operator on H. Prove:

(a) $U^{-1}AU : U^{-1}(D(A)) \to \mathsf{H}$ is self-adjoint;
(b) $\sigma(U^{-1}AU) = \sigma(A)$;
(c) $\mathscr{B}(\mathbb{R}) \ni E \mapsto U^{-1}P^{(A)}_E U$ is the spectral measure associated to $U^{-1}AU$ by the spectral theorem:

$$U^{-1}\int_{\mathbb{R}} \lambda dP^{(A)}(\lambda)U = \int_{\mathbb{R}} \lambda d(U^{-1}P^{(A)}U)(\lambda) ;$$

(d) $U^{-1}e^{itA}U = e^{-itU^{-1}AU}$.

Hint. (a) and (b) descend from the definitions of self-adjointness and spectrum. (c) follows from proving that for bounded maps $f : \mathbb{R} \to \mathbb{C}$, we have $U^{-1}\int_{\mathbb{R}} f(x)dP^{(A)}(x)U = \int_{\mathbb{R}} f(x)d(U^{-1}P^{(A)}U)(x)$. This comes directly from the definition of *integral of a bounded map* in a PVM (Chapter 8). Observing that any self-adjoint operator satisfies $T = \text{s-lim}_{n\to+\infty}\int_{\mathbb{R}} \chi_{[-n,n]}(x)dP^{(T)}(x)$, Stone's theorem and (a) imply (d).

12.7. Prove formula (12.63).

Outline of the solution. The first in (12.63) descends from U_Γ unitary, $U_{\Gamma_0}^{-1} = U_{\Gamma_0^{-1}}$ and $U_{\Gamma'}U_\Gamma = U_{\Gamma' \circ \Gamma}$. Hence it is enough to show, for any $\psi \in L^2(\mathbb{R}^3, dx)$:

$$||U_\Gamma \psi - \psi|| \to 0 \quad \text{as } \Gamma \to (\mathbf{0}, I).$$

Let us prove this for compactly-supported continuous maps ϕ. If ϕ is one such, $ISO(3) \times \mathbb{R}^3 \ni (\Gamma, \mathbf{x}) \mapsto \phi(\Gamma^{-1}\mathbf{x})$ is continuous. Then if Γ restricts to a relatively compact neighbourhood J of the identity, there is $K \geq 0$ such that $|\phi(\Gamma^{-1}\mathbf{x})| \leq K$ if $(\Gamma, \mathbf{x}) \in J \times \mathbb{R}^3$. Because of Γ there is a compact $S \subset \mathbb{R}^3$ containing every support of $\phi(\Gamma^{-1}\cdot)$. So there is $\phi_0 \in L^2(\mathbb{R}^3, dx)$ such that $|(U_\Gamma \phi)(\mathbf{x}) - \phi(\mathbf{x})| \leq |\phi_0(\mathbf{x})|$ if $(\Gamma, \mathbf{x}) \in J \times \mathbb{R}^3$: it suffices to choose a continuous ϕ_0 with absolute value, pointwise on S, larger than $2K$, and vanishing fast outside S. Since $(U_\Gamma \phi)(\mathbf{x}) \to \phi(\mathbf{x})$ pointwise, by dominated convergence $||U_\Gamma \psi - \psi|| \to 0$ as $\Gamma \to (\mathbf{0}, I)$, in L^2 norm. Let us pass to ψ generic in $L^2(\mathbb{R}^3, dx)$. If $\varepsilon > 0$, take ϕ continuous with compact support and such that $||\psi - \phi|| < \varepsilon/3$. Then

$$||U_\Gamma \psi - \psi|| \leq ||U_\Gamma \psi - U_\Gamma \phi|| + ||U_\Gamma \phi - \phi|| + ||\phi - \psi|| = ||U_\Gamma \phi - \phi|| + 2||\phi - \psi||,$$

since U_Γ is isometric so $||U_\Gamma \psi - U_\Gamma \phi|| = ||\psi - \phi||$. Choose Γ close enough to $(\mathbf{0}, I)$. By the above argument, $||U_\Gamma \phi - \phi|| \leq \varepsilon/3$. Hence for any $\varepsilon > 0$, if Γ is close enough to $(\mathbf{0}, I)$ we have $||U_\Gamma \psi - \psi|| \leq \varepsilon$.

12.8. Using Exercise 12.1, prove $\mathbf{t} \cdot \mathbf{P}\restriction_{\mathscr{S}(\mathbb{R}^3)}$ is essentially self-adjoint.

Hint. If $\mathbf{t} = \mathbf{0}$ the claim is trivial. Otherwise we know $P_1 \restriction_{\mathscr{S}(\mathbb{R}^3)}$ is essentially self-adjoint. Consider the unitary operator U_R representing an active rotation moving the axis $\mathbf{t}/|\mathbf{t}|$ onto $\mathbf{e}_3$. Show $U_R \mathbf{t} \cdot \mathbf{P}\restriction_{\mathscr{S}(\mathbb{R}^3)} U_R^{-1} = |\mathbf{t}| P_3 \restriction_{\mathscr{S}(\mathbb{R}^3)}$ and conclude.

12.9. Using Exercise 12.1, prove formula (12.64).

Hint. Prove the statement for P_3, passing from wavefunctions in $\mathbf{x}$ to wavefunctions in $\mathbf{k}$ via the Fourier trnasform. Extend to the general case as in the previous exercise. Note that U unitary and $A : D(A) \to \mathsf{H}$ closable imply UAU^{-1} is closable and

$$\overline{UAU^{-1}} = U\overline{A}U^{-1},$$

defining UAU^{-1} on $U(D(A))$.

12.10. Start from (12.29) and prove identity (12.30).

Hint. First, substitute g, g', g'' with one or more neutral elements e. Then write g^{-1} in place of g' and/or g''.

12.11. Let G be a *connected* topological group and $\mathsf{G} \ni g \mapsto \gamma_g$ a strongly continuous projective representation (Proposition 12.34) on the Hilbert space H_S, associated to a physical system. Suppose H_S decomposes in coherent sectors H_{Sk}. Can some γ_g map a certain sector to a different sector?

Hint. Decompose $\mathfrak{S}_p(\mathsf{H})$ in a disjoint union of pure states of each sector, and equip each with norm $|| \; ||_1$. Remember that continuous maps preserve connected sets.

12.12. Prove the Lie algebra of the real Lie group $SU(2)$ is the real vector space of skew-Hermitian matrices. Then show $SU(2)$ is simply connected.

Hint. $SU(2)$ is closed in $GL(4,\mathbb{R})$, hence a Lie group. Therefore one-parameter groups are of type $\mathbb{R} \ni t \mapsto e^{tA}$, with A in the Lie algebra $su(2)$. Impose $e^{tA}(e^{tA})^* = I$ and $tr(e^{tA}) = 1$ for every t, and infer how A has to look like. *Vice versa* if A is skew-Hermitian, verify the above two conditions hold. As for simply connectedness, parametrise the group by 4 real parameters so that $SU(2)$ is in one-to-one correspondence with the unit sphere in $\mathbb{R}^4$. Show the parametrisation is a homeomorphism.

12.13. Prove that $U \in SU(2)$ iff there exist a unit vector $\mathbf{n} \in \mathbb{R}^3$ and a real number θ such that:

$$U = e^{-i\theta \mathbf{n}\cdot \frac{\sigma}{2}} .$$

Hint. Use the spectral theorem for the unitary operator $U \in SU(2)$, keeping in account that the Pauli matrices and I form a real basis of 2×2 Hermitian matrices. Conversely, if $U = e^{-i\theta \mathbf{n}\cdot \frac{\sigma}{2}}$, what are U^*U and $\det U$?

12.14. Prove the matrices T in (12.72) satisfy:

$$RT_kR^t = \sum_{i=1}^{3} (R^t)_{ki} T_i \quad \text{for any } R \in SO(3).$$

Hint. Use $(T_i)_{jk} = -\varepsilon_{ijk}$ and write the above equations componentwise. Recall ε_{ijk} are the coefficients of a pseudotensor invariant under proper rotations.

12.15. Show $R \in SO(3)$ iff there exist a unit $\mathbf{n} \in \mathbb{R}^3$ and a real θ such that:

$$R = e^{\theta \mathbf{n}\cdot \mathbf{T}} .$$

Hint. Prove the claim for $\mathbf{n} = \mathbf{e}_3$ by taking, for $R \in SO(3)$, a rotation about $\mathbf{e}_3$. Show every $R \in SO(3)$ always admits an eigenvector $\mathbf{n}$. Rotate the axes so to move $\mathbf{n}$ onto $\mathbf{e}_3$, and recall the previous exercise. If, conversely, $R = e^{-i\theta \mathbf{n}\cdot \mathbf{T}}$, what are R^tR and $\det R$?

12.16. Demostrate that for every $U \in SU(2)$ there exists a unique $R_U \in SO(3)$ such that:

$$U\mathbf{t}\cdot \sigma U^* = (R_U \mathbf{t})\cdot \sigma \quad \text{for any } \mathbf{t} \in \mathbb{R}^3.$$

Then verify

$$SU(2) \ni U \mapsto R_U \in SO(3)$$

is a surjective homomorphism that coincides with:

$$R : SU(2) \ni e^{-i\theta \mathbf{n}\cdot\frac{\sigma}{2}} \mapsto e^{\theta \mathbf{n}\cdot\mathbf{T}} \in SO(3) .$$

Eventually prove the kernel is $\{\pm I\} \subset SU(2)$.

Sketch of the solution. Note $|\mathbf{t}|^2 = \det(\mathbf{t}\cdot\sigma)$, and conclude every $U \in SU(2)$ determines a unique transformation of $\mathbb{R}^3$ mapping $\mathbf{t}$ to some $\mathbf{t}'$, with $|\mathbf{t}| = |\mathbf{t}'|$, defined by $U\mathbf{t}\cdot\sigma U^* = U\mathbf{t}'\cdot\sigma U^*$. The transformation $\mathbf{t} \to \mathbf{t}'$ is then an orthogonal matrix $R(U) \in O(3)$. That $R : SU(2) \ni U \mapsto R(U) \in O(3)$ is a homomorphism is immediate by construction. In the case $U_\theta = e^{-i\theta\frac{\sigma_3}{2}}$ one checks in various ways (e.g. directly, expanding the exponentials) that $R(U_\theta) = e^{\theta T_3}$. The general case relies on exercise (12.14), rotating $\mathbf{e}_3$ onto an arbitrary unit vector $\mathbf{n}$. $R(U_\theta) = e^{\theta\mathbf{n}\cdot\mathbf{T}}$ clearly implies $R(U) \in SO(3)$. Surjectivity is a consequence of the fact every $SO(3)$ matrix can be written as $e^{\theta\mathbf{n}\cdot\mathbf{T}}$. The kernel is computed reducing to the one-parameter subgroup generated by σ_3, by rotation of $\mathbf{n}$. The result becomes thus obvious by direct computation.

12.17. Referring to Chapter 12.3.1, prove the strongly continuous unitary $SU(2)$-representation obtained from exponentiating the $\overline{\mathscr{L}_k}$ is the representation $SO(3) \ni R \mapsto U_R$ of Example 12.17 (where $\Gamma \in IO(3)$ is now restricted to $\Gamma = R \in SO(3)$), that is strongly continuous (cf. Example 12.37(1)).

Hint. By Nelson's Theorem 12.62 it suffices to check the one-parameter groups $\theta \mapsto U_{e^{\theta\mathbf{n}\cdot\mathbf{T}_\mathbf{X}}}$, with $\mathbf{n} = \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$, are generated by the self-adjoint L_1, L_2, L_3. It is convenient to work with polar coordinates, using the core of L_1, L_2, L_3 given by spherical harmonics multiplied by a basis of $L^2(\mathbb{R}_+, r^2dr)$.

12.18. Show that the $SU(2)$-representation obtained by exponentiating the generators J_k, by Nelson's theorem, has the form:

$$SU(2) \ni e^{-i\theta\frac{1}{2}\mathbf{n}\cdot\sigma} \mapsto e^{-i\theta\mathbf{n}\cdot\mathbf{J}} = e^{-i\theta\mathbf{n}\cdot\mathbf{L}} \otimes V^s\left(e^{-i\frac{\theta}{2}\mathbf{n}\cdot\sigma}\right) .$$

Hint. Employ the properties of the tensor product of operators to prove

$$e^{-i\theta\mathbf{n}\cdot\mathbf{J}} = e^{-i\theta\mathbf{n}\cdot\mathbf{L}} \otimes V^s\left(e^{-i\frac{\theta}{2}\mathbf{n}\cdot\sigma}\right) .$$

So we have to prove the representation $SO(3) \ni R \mapsto U_R$ of the previous exercise can be written as $U_{e^{\theta\mathbf{n}\cdot\mathbf{T}}} = e^{-i\theta\mathbf{n}\cdot\mathbf{L}}$. This is certainly true, for instance with $\mathbf{n} = \mathbf{e}_3$. As for the general case: on one hand we have

$$U_R^* e^{-i\theta\mathbf{n}\cdot\mathbf{L}} U_R = e^{-i\theta\mathbf{n}\cdot\mathbf{U_R^*LU_R}} ,$$

and Exercise 12.3; on the other the result of Exercise 12.14 holds.

13

Selected advanced topics in Quantum Mechanics

Give up telling God what to do with his dice.
Niels Bohr, to Einstein

With this chapter we complete the list of axioms for non-relativistic Quantum Mechanics, introducing *time evolution* and *compound systems*. Certain notions, here defined formally, have already been introduced in the final part of the previous chapter when talking about symmetry groups. More advanced reference texts, which we have followed here and there, are [Pru81] and [DA10].

In the first section we will state the *axiom of time evolution*, described by a strongly continuous one-parameter unitary group that is generated by the *Hamiltonian operator* of the system. Within this framework we will define *dynamical symmetries* as a special kind of the symmetries seen earlier. Then we shall discuss the nature of *Schrödinger's equation* and introduce the important concept of *stationary state*. As classical example of this formalism we will make explicit the action of the Galilean group in the position representation (we saw it in the momentum representation in the previous chapter). There, we will also explain how wavefunctions transform under changes of inertial frame systems. Then we will pass to the basic theory of non-relativistic scattering. We will make a few remarks on the existence of the unitary time-evolution operator in absence of time homogeneity (we will examine the convergence in $\mathfrak{B}(\mathsf{H})$ of the *Dyson series* for a Hamiltonian), and discuss the antiunitarity of the time-reversal symmetry.

In the following section we will present a version of the so-called *Pauli theorem*, whose concern is the possibility of defining the "operator time" as self-adjoint conjugate to the Hamiltonian. In this respect we will briefly discuss the notion of POVM, that may be employed to give a weaker meaning to the time observable.

Heisenberg's picture of observables is introduced in section three, where we will discuss the relationship between constants of motion and dynamical symmetries, present the quantum version of *Nöther's theorem* and study the case of constants of motion associated to generators of a Lie group, including the one-parameter subgroup of time evolution. A digression will give us the chance to present the mathematical problems raised by the so-called *Ehrenfest theorem*. The section will close with the study of the constants of motion associated to the Galilean group.

Section four will be devoted to the theory of *compound quantum systems*: systems with an inner structure and multi-particle systems. We will consider, in particular, *en-*

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_13, © Springer-Verlag Italia 2013

tangled states and discuss ome problems related to the so-called *EPR paradox* and the notion of *decoherence*. Eventually, we will pass to the general theory of *systems of identical particles*, and conclude with the *spin-statistic* correlation.

13.1 Quantum dynamics and its symmetries

As we quickly recalled in Chapter 7.2.1, physical systems evolve in time acording to their dynamics. In the classical Hamiltonian formulation of mechanics the evolution in time of a system's state is described in phase spacetime by the solutions to *Hamilton's equations*. Let us consider the situation in which the *Hamiltonian function H does not depend explicitly on time* in the coordinates of a given inertial frame $\mathscr{I}$. We will talk in this case of time being *homogeneous* with respect to the considered physical system. Hamilton's equations are *autonomous* PDEs, i.e. the time variable does not show up explicitly if the equations are written in those canonical coordinates and the phase spacetime splits naturally in a product $\mathbb{R} \times \mathscr{F}$, where $\mathscr{F}$ is the *phase space*. Solutions to Hamilton's equations determine a one-parameter group of diffeomorphisms $\phi_\tau : \mathscr{F} \to \mathscr{F}$ mapping the initial state (taken to be sharp for simplicity) $r \in \mathscr{F}$ at time 0 into the state $\phi_\tau(r) \in \mathscr{F}$ at time τ. The basic mathematical tool to construct the *time-evolution operator* – the one-parameter group $\{\phi_\tau\}_{\tau\in\mathbb{R}}$ – is the Hamiltonian H of the system, which identifies with the total *mechanical energy* of the frame system $\mathscr{I}$ [GPS01, FaMa06]. In the sequel we will present the quantum analogues of the Hamiltonian function and the evolution operator.

13.1.1 Axiom A6*: time evolution*

The quantum case is not dissimilar to the classical setting. The following axiom comprises time evolution in a quantum system S, described on the Hilbert space H_S for given inertial frame $\mathscr{I}$, with homogeneous time. The axiom defines the *Hamiltonian* (operator) of the quantum system as the generator of the one-parameter unitary group capturing the evolution, hence the dynamics, of the quantum state. (We will return to this in Chapter 13.1.6 when looking at a more general situation.) Using the notion of time evolution makes possible to treat *dynamical symmetries* and, as we will see later, state the quantum *Nöther theorem.*

A6. *Let S be a quantum system described on the Hilbert space H_S associated to the inertial frame $\mathscr{I}$. There exists a self-adjoint operator H, called* **Hamiltonian of the system** *S in the frame $\mathscr{I}$, corresponding to the observable of total mechanical energy of S in the frame $\mathscr{I}$, such that*

(i) *$\sigma(H)$ is lower bounded;*

(ii) *setting $U_\tau := e^{-\frac{i\tau}{\hbar}H}$, if the system's state at time t is $\rho_t \in \mathfrak{S}(\mathsf{H}_S)$, the state at time $t+\tau$ is:*

$$\rho_{t+\tau} = \gamma_\tau^{(H)}(\rho_t) := U_\tau \rho_t U_\tau^{-1}\,. \tag{13.1}$$

The strongly continuous one-parameter unitary group $\mathbb{R} \ni \tau \mapsto U_\tau$ *is called* **time-evolution operator** *of* S *in the frame* $\mathscr{I}$*, and the continuous projective representation of* $\mathbb{R}$ *induced by* U*,* $\mathbb{R} \ni \tau \mapsto \gamma_\tau^{(H)}$*, is said* **dynamical flow** *of* S *in the frame* $\mathscr{I}$*.*

Remark 13.1. (1) From now on, unless strictly necessary for better physical clarity, we will omit to write $\hbar$ explicitly in formulas, and set $\hbar = 1$.

(2) The evolution of states is thus given by a continuous projective representation of the Abelian group $\mathbb{R}$. This fact enables us to phrase differently axiom **A6**, using the results of the preceding chapter.

With the intent to weaken the axiom's assumptions as much as possible, and think the evolution as a function $\rho \mapsto \gamma_\tau(\rho)$ mapping states to states for any $\tau \in \mathbb{R}$, we may require γ_τ to satisfy the following conditions, rather reasonable from the physical viewpoint:

(i) γ_τ preserves the convexity of the space of states (Kadison symmetry), or equivalently, it preserves transition probabilities (Wigner symmetry);
(ii) γ_τ is additive: $\gamma_\tau \circ \gamma_{\tau'} = \gamma_{\tau+\tau'}$, for $\tau, \tau' \in \mathbb{R}$;
(iii) (viewing symmetries γ_τ *à la* Wigner) γ_τ is continuous as in Definition 12.31 or equivalently, continuous in the topology of $\mathfrak{S}_p(\mathsf{H}_S)$ induced by $|| \ ||_1$, as in (12.42).

Then Theorem 12.36 proves the projective representation $\mathbb{R} \ni \tau \mapsto \gamma_\tau$ has the form predicted by axiom **A6** above. One of the possible self-adjoint generators – which exist and differ by an additive constant, by Theorem 12.36 – is the system's Hamiltonian, by definition. But we still need to impose the spectrum be bounded from below. In defining the Hamiltonian, the ambiguity coming from the additive constant is actually present in physics, because the energy of a classical system (non-relativistic) is given up to constant. (However, the picture is not so obvious dealing with advanced topics as the mass superselection rule, when noticing that classical physics arises as an approximation from relativistic physics [AnMo12].)

(3) That the Hamiltonian spectrum of a real physical systems is bounded stems from thermodynamical stability. Unless we consider an ideal system, perfectly isolated from the environment, and that in reality does not exist (also by deep theoretical motivations that demand quantum field theory to be explained properly), the lower limit constraining the spectrum of H (the mechanical energy) is unavoidable. In absence of a threshold there could be transitions in S to states with decreasingly lower energy. This (infinite!) energy loss towards the outside, in some form or other (particles, electromagnetic waves), would in practice make the system collapse. The lower limit of $\sigma(H)$ has other important theoretical repercussions we will see later.

(4) The *inverse* symmetry to time evolution is called **time displacement**. We met this symmetry when studying the Galilean group. Physically it is an *active* transformation of S. In other words, for given τ, it is a Kadison symmetry $\gamma_\tau^{(-H)} : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$ that transforms the state ρ at a generic given time t_0 into the state $\tau_\tau(\rho)$ *at the same time* t_0, so that $\gamma_\tau^{(H)}\left(\gamma_\tau^{(-H)}(\rho)\right)$ coincides with ρ. By construction $\gamma_\tau^{(-H)} = \left(\gamma_\tau^{(H)}\right)^{-1}$.

Evidently the unitary generator of $\gamma_\tau^{(-H)}$ is $-H$, as the name already suggests. This explains the "$-$" sign used for the self-adjoint generators of the Galilean group. ■

Now let us suppose H_S is decomposed in coherent sectors H_{Sk}, $k \in K$. Then the space of admissible pure states $\mathfrak{S}_p(\mathsf{H}_S)_{adm}$ splits in the disjoint union of sets $\mathfrak{S}_p(\mathsf{H}_{Sk})$, and mixed states are convex combinations of elements in the various $\mathfrak{S}(\mathsf{H}_{Sk})$. The next result shows that the dynamical flow preserves this splitting, as expected.

Proposition 13.2. *Let S be a quantum system described on the Hilbert space H_S associated to the inertial frame $\mathscr{I}$, with dynamical flow $\gamma^{(H)}$. Suppose H_S splits in coherent sectors H_{Sk}, $k \in K$. Then the dynamical flow preserves both pure and mixed states. More precisely:*

(a) *if $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$ for some $k \in K$, then $\gamma_t^{(H)}(\rho) \in \mathfrak{S}(\mathsf{H}_{Sk})$ for every $t \in \mathbb{R}$.*

(b) *If $\rho \in \mathfrak{S}_p(\mathsf{H}_{Sk})$ for some $k \in K$, then $\gamma_t^{(H)}(\rho) \in \mathfrak{S}_p(\mathsf{H}_{Sk})$ for every $t \in \mathbb{R}$.*

Proof. Since

$$\gamma_t^{(H)}(\psi(\psi|\ )) = e^{-itH}\psi\left(e^{-itH}\psi|\ \right),$$

clearly the representation $\gamma^{(H)}$ maps pure states to pure states, so mixed to mixed ones. Restrict $\gamma^{(H)}$ to pure states. Fix $\rho \in \mathfrak{S}_p(\mathsf{H}_{Sk})$ and consider the path $\mathbb{R} \ni t \mapsto \gamma_t^{(H)}(\rho)$. By Proposition 12.34 it is continuous for $||\cdot||_1$. We know the $\mathfrak{S}_p(\mathsf{H}_{Sk})$ are the connected components of $\mathfrak{S}_p(\mathsf{H}_S)_{adm}$ for the topology induced by the aforementioned norm (Exercise 12.4), so the curve is confined to live in one component only. The latter is $\mathfrak{S}_p(\mathsf{H}_{Sk})$, since the path passes through there at $t = 0$. If $U_t = e^{-itH}$, then, for any unit $\psi \in \mathsf{H}_{Sk}$ we have $U_t\psi \in \mathsf{H}_{Sk}$ for all t. Consider now $\rho \in \mathfrak{S}(\mathsf{H}_{Sk})$ and its spectral decomposition $\rho = \sum_{j\in J} p_j\psi_j(\psi_j|\ )$. The series converges strongly and by construction $\psi_j \in \mathsf{H}_{Sk}$, $j \in J$, is a unit vector. Therefore for any $t \in \mathbb{R}$:

$$\gamma_t^{(H)}(\rho) = U_t \sum_{j\in J} p_j\psi_j(\psi_j|\ )U_t^{-1} = \sum_{j\in J} p_j U_t\psi_j(\psi_j|U_t^*\ )$$

$$= \sum_{j\in J} p_j U_t\psi_j(U_t\psi_j|\ ) \in \mathfrak{S}(\mathsf{H}_{Sk}),$$

ending the proof. □

Remarks 13.3. From now the system's Hilbert space H_S will not contain coherent sectors, apart from a few cases we will comment upon. We leave it to the reader to generalise the ensuing definitions and results to the multi-sector case. ■

13.1.2 Dynamical symmetries

Time evolution allows to refine the notion of symmetry seen in the previous chapter and define *dynamical symmetries*.

Consider a quantum system S with dynamical flow $\gamma^{(H)}$. Let us assume, as we said, the Hilbert space consists of a single coherent sector. Take a symmetry σ (Kadison or Wigner) acting on states, paying attention that now states evolve in time following the dynamics of the flow $\gamma^{(H)}$. If we apply σ to the evolved state $\gamma_t^{(H)}(\rho)$ and obtain $\rho'_t := \sigma(\gamma_t^{(H)}(\rho))$, nothing guarantees the function $\mathbb{R} \ni t \mapsto \rho'_t$ describes the possible evolution under $\gamma^{(H)}$ of a certain state (necessarily $\rho'_0 = \sigma(\gamma_0^{(H)}(\rho)) = \sigma(\rho)$). Conversely, if this happens (for any choice of initial state ρ), σ is called a *dynamical symmetry*, because its action *is compatible* with the system's dynamics.

A variant to having $\mathbb{R} \ni t \mapsto \sigma(\gamma_t^{(H)}(\rho))$ describing the evolution of a state in S, is to take a whole family of symmetries $\sigma^{(t)}$ parametrised by time $t \in \mathbb{R}$. To have a *time-dependent* dynamical symmetry we require $\mathbb{R} \ni t \mapsto \sigma^{(t)}(\gamma_t^{(H)}(\rho))$ still be an evolution under $\gamma^{(H)}$ for some state of S.

More formally:

Definition 13.4. *Let S be a quantum system described on the Hilbert space H_S made of one sector and associated to the frame system $\mathscr{I}$, with Hamiltonian H and dynamical flow $\gamma^{(H)}$. A symmetry $\sigma : \mathfrak{S}(\mathsf{H}_S) \to \mathfrak{S}(\mathsf{H}_S)$ is called a* **dynamical symmetry** *for S if*

$$\gamma_t^{(H)} \circ \sigma = \sigma \circ \gamma_t^{(H)} \quad \textit{for every } t \in \mathbb{R}. \tag{13.2}$$

A family of symmetries parametrised by time $\{\sigma^{(t)}\}_{t\in\mathbb{R}}$ is a **time-dependent dynamical symmetry** *when:*

$$\gamma_t^{(H)} \circ \sigma^{(0)} = \sigma^{(t)} \circ \gamma_t^{(H)} \quad \textit{for every } t \in \mathbb{R}. \tag{13.3}$$

The first result we prove characterises dynamical symmetries. Part (c) is a consequence of the spectral lower bound of H and characterises dynamical symmetries when $\sigma(H)$ is unbounded, as for the majority of concrete physical systems.

Theorem 13.5. *Let S be a quantum system described on the Hilbert space H_S associated to the inertial reference frame $\mathscr{I}$ with Hamiltonian H (hence, with lower-bounded spectrum) and dynamical flow $\gamma^{(H)}$.*

(a) *Consider a family of symmetries labelled by time $\{\sigma^{(t)}\}_{t\in\mathbb{R}}$ and induced by unitary (or antiunitary) operators $V^{(\sigma^{(t)})} : \mathsf{H}_S \to \mathsf{H}_S$. Then $\{\sigma^{(t)}\}_{t\in\mathbb{R}}$ is a time-dependent dynamical symmetry for S if and only if*

$$\chi_t V^{(\sigma^{(t)})} e^{-itH} = e^{-itH} V^{(\sigma^{(0)})} \quad \textit{for every } t \in \mathbb{R} \textit{ and some unit } \chi_t \in \mathbb{C}\,.$$

(b) *Consider a symmetry σ induced by a unitary (or antiunitary) $V^{(\sigma)} : \mathsf{H}_S \to \mathsf{H}_S$. Then σ is a dynamical symmetry for S if and only if*

$$e^{-iat} V^{(\sigma)} e^{-itH} = e^{-itH} V^{(\sigma)} \quad \textit{for every } t \in \mathbb{R} \textit{ and some } a \in \mathbb{R}\,.$$

(c) *Consider a symmetry σ induced by a unitary (or antiunitary) $V^{(\sigma)} : \mathsf{H}_S \to \mathsf{H}_S$ and suppose $\sigma(H)$ is not bounded above. Then σ is a dynamical symmetry for S if and only if*

$$V^{(\sigma)} e^{-itH} = e^{-itH} V^{(\sigma)} \quad \textit{for every } t \in \mathbb{R}\,,$$

or equivalently, if and only if the following hold:

(i) *$V^{(\sigma)}$ is unitary and*
(ii) *$V^{(\sigma)} H = H V^{(\sigma)}$.*

Proof. (a) and (b) Remember that for $S : \mathsf{H}_S \to \mathsf{H}_S$ unitary (or antiunitary), $S\psi(\psi|S^{-1}\cdot) = S\psi(\psi|S^*\cdot) = S\psi(S\psi|\cdot)$. Set $U_t := e^{-itH}$, $V^{(t)} := V^{(\sigma^{(t)})}$ and use the identity with the unitary $S := (V^{(t)}U_t)^{-1}U_tV^{(0)}$. Then (13.3) implies, for any pure $\rho = \psi(\psi|\ )$:

$$(V^{(t)}U_t)^{-1}U_tV^{(0)}\psi\left((V^{(t)}U_t)^{-1}U_tV^{(0)}\psi\right|\ \Big) = \psi(\psi|\ )\,,$$

hence for some unit $\chi_t \in \mathbb{C}$:

$$(V^{(t)}U_t)^{-1}U_tV^{(0)}\psi = \chi_t\psi \quad \text{for all } \psi \in \mathsf{H}.$$

The same proof used for the corresponding fact in Theorem 12.10 tells χ_t does not depend on ψ. Therefore if $\sigma^{(t)}$ is a time-dependent dynamical symmetry:

$$\chi_t V^{(\sigma^{(t)})}U_t = U_tV^{(\sigma^{(0)})} \quad \text{for all } t \in \mathbb{R} \text{ and some } \chi_t \in \mathbb{C},\ |\chi_t| = 1\,.$$

Conversely, if the condition holds, trivially $\sigma^{(t)}$ is a time-dependent dynamical symmetry. Statement (b) is a subcase, except for the proof that $\chi_t = e^{ict}$ for some $c \in \mathbb{R}$, which we will see at the end.

(c) We claim that if σ is a dynamical symmetry then (i), (ii) hold. By (a), if σ is a dynamical symmetry:

$$\chi_t V^{(\sigma)}U_t = U_tV^{(\sigma)} \quad \text{for some unit } \chi_t \in \mathbb{C}\,. \tag{13.4}$$

Hence $\chi_t I = (V^{(\sigma)}U_t)^{-1}U_tV^{(\sigma)}$ and $\chi_t(\psi|\phi) = \left(V^{(\sigma)}U_t\phi\middle|U_tV^{(\sigma)}\psi\right)$. Choose $\phi \in D(H)$ not orthogonal to $\psi \in V^{(\sigma)-1}(D(H))$ (since $D(H)$ is dense), apply Stone's theorem and conclude $t \mapsto \chi_t$ is smooth everywhere. We rewrite (13.4) as:

$$\chi_t U_t = e^{\pm itV^{(\sigma)-1}HV^{\sigma}}\,, \tag{13.5}$$

with $-$ if $V^{(\sigma)}$ is unitary, $+$ if antiunitary. Using Stone's theorem in (13.5) we obtain $D(V^{(\sigma)-1}HV^{(\sigma)}) \subset D(H) = D(cI+H)$ and

$$\mp V^{(\sigma)-1}HV^{(\sigma)}\restriction_{D(H)} = cI + H \quad \text{where } c := i\frac{d\chi_t}{dt}|_{t=0}\,. \tag{13.6}$$

Note c must be real since $\mp V^{(\sigma)-1}HV^{(\sigma)} - H$ is symmetric on $D(H)$. Actually (13.6) holds on the entire domain of $V^{(\sigma)-1}HV^{(\sigma)}$ because the latter is self-adjoint and does not have self-adjoint extensions ($cI+H$) other than $\mp V^{(\sigma)-1}HV^{(\sigma)}$ itself. Therefore

$$V^{(\sigma)-1}HV^{(\sigma)} = \mp cI \mp H\,. \tag{13.7}$$

In particular (cf. Exercise 12.6 in the antiunitary case):

$$\sigma(H) = \sigma(V^{(\sigma)-1}HV^{(\sigma)}) = \sigma(\mp cI \mp H) = \mp c \mp \sigma(H)\,.$$

If $\sigma(H)$ is bounded below but not above, the identity cannot be valid if on the right side we have the minus sign, irrespective of the constant c. Hence $V^{(\sigma)}$ must be unitary. Therefore $\inf\sigma(H) = \inf(c+\sigma(H)) = c + \inf\sigma(H)$ and $c = 0$, for $\inf\sigma(H)$ is finite by hypothesis ($\sigma(H) \neq \emptyset$ is bounded). We obtained that a dynamical symmetry σ fulfills (i) and (ii): $V^{(\sigma)}$ is unitary and $V^{(\sigma)}H = HV^{(\sigma)}$. If so, $H = V^{(\sigma)-1}HV^{(\sigma)}$. Exponentiating,

$$V^{(\sigma)}e^{-itH} = e^{-itH}V^{(\sigma)} \quad \text{for every } t \in \mathbb{R}\,.$$

Eventually, σ is a dynamical symmetry, ending part (c).

We still have to finish part (b). If σ is a simmetry, but $\sigma(H)$ is upper bounded, using the proof of (c) we arrive at (13.7) but cannot say $c = 0$. Exponentiating (13.7) gives:

$$e^{-ict}U_t = V^{(\sigma)-1}U_tV^{(\sigma)}\,,$$

whence

$$e^{-iat}V^{(\sigma)}e^{-itH} = e^{-itH}V^{(\sigma)}$$

($a = \pm c$ according to whether $V^{(\sigma)}$ is unitary or antiunitary). This ends part (b) and the theorem. □

13.1.3 Schrödinger's equation and stationary states

Consider a pure initial state $\rho \in \mathfrak{S}_p(\mathsf{H}_S)$. As already noticed, the evolution is such that any evolved state ρ_t is pure. Theoretical physicists refer to this property[1] by saying the evolution of quantum states is *unitary*. If $t \mapsto \rho_t \in \mathfrak{S}_p(\mathsf{H}_S)$ denotes the evolution of a pure state, we can determine any ρ_t, up to phase, by a vector ψ_t normalised to 1. Choosing the simplest phases for the pure states involved, the equation governing the evolution of pure states becomes (reintroducing the constant $\hbar$):

$$\psi_{t'} = e^{-\frac{i(t'-t)}{\hbar}H}\psi_t\,.$$

We can manipulate this relation to obtain an equation of great historical relevance. For this we impose that $\psi_t \in D(H)$ equal $\psi_{t'} \in D(H)$ for any other time $t' \in \mathbb{R}$. In fact, $\psi_t \in D(H)$ means $\int_{\mathbb{R}} \lambda^2 d\mu^{(H)}_{\psi_t} < +\infty$, where $\mu^{(H)}_{\psi_t}(E) = (\psi_t|P^{(H)}(E)\psi_t) = (\psi_{t'}|e^{+\frac{i\tau}{\hbar}H}P^{(H)}(E)e^{-\frac{i\tau}{\hbar}H}\psi_{t'})$, for $t - t' = \tau$. On the other hand, trivially,

$$e^{+\frac{i\tau}{\hbar}H}P^{(H)}(E)e^{-\frac{i\tau}{\hbar}H} = P^{(H)}(E)\,,$$

since $P^{(H)}(E)$ is a projector of the PVM of H. Hence $\int_{\mathbb{R}} \lambda^2 d\mu^{(H)}_{\psi_t} < +\infty$ is equivalent to $\int_{\mathbb{R}} \lambda^2 d\mu^{(H)}_{\psi_{t'}} < +\infty$, i.e. $\psi_{t'} \in D(H)$. Let us suppose $\psi_t \in D(H)$ for some t, from which $\psi_{t'} \in D(H)$ for every t'. Applying Stone's theorem to

$$\psi_{t'} = e^{-\frac{i(t'-t)}{\hbar}H}\psi_t\,,$$

and interpreting the outcoming derivative in strong sense, we obtain

$$i\hbar\frac{d}{dt}\psi_t = H\psi_t\,. \tag{13.8}$$

[1] Especially in relationship to the evolution of states of quantum fields in spacetimes comprising dynamical black holes, where the unitary evolution is rather problementic.

This is the fundamental **time-dependent Schrödinger equation**. We have to notice (13.8) only holds if $\psi_t \in D(H)$, whereas the evolution equation (13.1) has a general reach.

Let us make a few comments on the Schrödinger equation and then pass to more general matters.

Consider a system formed by one particle of mass m (without spin for simplicity) subjected to a force with sufficiently regular potential energy $V = V(\mathbf{x})$, in the inertial frame $\mathscr{I}$ with right-handed orthonormal coordinates. Following the discussion about *Dirac's correspondence principle*, at the end of Chapter 11, one expects the Hamiltonian of this system to correspond, in quantum sense, to a certain self-adjoint extension H of the symmetric operator

$$H_0 := \frac{1}{2m}\sum_{i=1}^{3} P_i^2 + V(\mathbf{X}),$$

initially defined on some invariant dense subspace where P_i and X_i are well defined. This choice formally fulfills Dirac's correspondence, at least with respect to the commutation relations of the operator above and X_k, P_k, when working on domains where everything is well defined. This expectation turns out to be correct and the Hamiltonian observables do have the mentioned form in the physical world, like for systems formed by atoms and molecules [Mes99, CCP82].

We shall identify the particle's Hilbert space with $L^2(\mathbb{R}^3, dx)$ so that position operators are multiplicative. If we work with functions that are regular enough, the starting expression for H is

$$H_0 = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}), \tag{13.9}$$

where Δ is the known Laplacian on $\mathbb{R}^3$ and $V(\mathbf{X})$ is the multiplication by the original function $V = V(\mathbf{x})$. Schrödinger's equation then reads:

$$\left[-\frac{\hbar^2}{2m}\Delta + V(X)\right]\psi_t(x) = i\hbar\frac{\partial}{\partial t}\psi_t(x),$$

which is precisely how Schrödinger wrote it in his astounding 1926 papers. Beware, however, the equation should not be taken literally as a usual PDE, because: (1) the t-derivative is not meant pointwise, but in Hilbert sense[2]; (2) the equation is valid up to zero-measure sets for x, since wavefunctions belong to $L^2(\mathbb{R}^3, dx)$. If we were to find "naïve" solutions (functions $f(t,x)$ in t and x), we would then have to prove they solve (13.8) in the unknown $\psi_t = f(t,\cdot) \in L^2(\mathbb{R}^3, dx)$.

Let us return to how to define the Hamiltonian operator from the symmetric differential operator (13.9) defined on a dense domain. We have to verify, case by case, if the operator admits self-adjoint extensions or if it is essentially self-adjoint. In

[2] Observe, nevertheless, that if the derivative exists both in the ordinary and in L^2 sense, the two concide by Proposition 2.29 for $p = 2$.

this respect the symmetric operator H_0 commutes with the operator $C : L^2(\mathbb{R}^3, dx) \to L^2(\mathbb{R}^3, dx)$ representing the complex conjugation of L^2 functions. By von Neumann's Theorem 5.43, then, there are self-adjoint extensions. The general theory of self-adjoint extensions of operators like H_0 was developed and harvested by T. Kato [Kat66]. For several potentials of interest, like the attractive Coulomb potential and the harmonic oscillator, one can prove H_0 is essentially self-adjoint. We saw these results in Examples 10.51, Chapter 10.4, as consequences of general theorems. There is a whole branch of functional analysis in Hilbert spaces devoted to this sort of problems. We mention just one easy corollary of Theorem 10.49.

Theorem 13.6 (Kato). *Consider the differential operator on* $\mathbb{R}^3$*:*

$$H_0 := -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}), \tag{13.10}$$

defined on some dense domain $D(H_0) \supset \mathscr{S}(\mathbb{R}^3)$*. Suppose*

$$V(\mathbf{x}) = \sum_{j=1}^{N} \frac{g_j}{|\mathbf{x} - \mathbf{x}_j|} + U(\mathbf{x}), \tag{13.11}$$

where g_j *are constants,* $\mathbf{x}_j \in \mathbb{R}^3$ *are given points and* $U : \mathbb{R}^3 \to \mathbb{R}$ *is measurable and (essentially) bounded. Then:*

(a) H_0 *is essentially self-adjoint on* $D(H_0)$*,* $\mathscr{D}(\mathbb{R}^3)$ *and* $\mathscr{S}(\mathbb{R}^3)$*.*
(b) *The common self-adjoint extension* $\overline{H_0}$ *of the operators in (a) coincides the self-adjoint* $\overline{-\Delta} + V$ *defined on* $D(\overline{-\Delta})$*.*
(c) $\sigma(\overline{H_0})$ *is bounded from below.*

In general, if the Hamiltonian H of a certain system has point spectrum $\sigma_p(H)$, every eigenvector ψ_E of H, for $E \in \sigma_p(H)$, has a trivial evolution:

$$U_t\psi_E = e^{-i\frac{tE}{\hbar}}\psi_E\,.$$

This says the pure state $\rho_E := \psi_E(\psi_E|\;)$ associated to ψ_E ($||\psi_E|| = 1$) *does not evolve in time*. These very special states are said **stationary states** of the system. When one studies the macroscopic system of an atom or a molecule, to begin with the heavier parts – nuclei – are described as classical systems, that act by electric Coulomb forces on peripheral electrons viewed as quantum particles. The electrons' quantum states are stationary for their Hamiltonian. More on this in Example 13.8(3).

Remark 13.7. (1) Referring to Theorem 13.6 it can be proved [CCP82] that if some g_j vanish and the remaining are strictly negative then $\sigma_p(\overline{H_0}) \neq \varnothing$.
(2) By virtue of Theorem 10.50, H_0 continues to be essentially self-adjoint on $\mathscr{D}(\mathbb{R}^3)$ and the only self-adjoint extension is bounded from below provided U is non-negative and lower bounded. In that case [CCP82] if $g_j = 0$ for every j and U is regular enough and tends to infinity as $|\mathbf{x}| \to +\infty$, then $\sigma(\overline{H_0}) = \sigma_p(\overline{H_0}) \neq \varnothing$.

(3) One of the highest mountain tops the inexperienced student has to conquer when taking on QM is to understand the motivations behind the regularity constraints imposed on the eigenvalues of the theory's Hamiltonian. The characteristic equation

$$H_0\psi_E = E\psi_E\,, \quad \text{with } E \in \mathbb{R}\,, \quad \psi_E \in L^2(\mathbb{R}^3, dx)\,,$$

should give, roughly speaking, the stationary states of the system whose Hamiltonian is determined by H_0. Consider, as often in physics, an operator of the form (13.10) where the $U : \mathbb{R}^3 \to \mathbb{R}$ of (13.11) has finite discontinuities on some regular surfaces σ_k, $k = 1, 2, \ldots, N$ (disjoint from one another and from other isolated singularities of V) and is continuous everywhere else. We also want U to be bounded (by remark (2) we could just require lower boundedness). QM manuals typically require the functions ψ_E further satisfy the following conditions:

(1) away from the singularities of V the ψ_E are C^2 (actually C^∞);
(2) the ψ_E solve $H_0\psi_E = E\psi_E$ for some $E \in \mathbb{R}$, i.e. interpreting the operator as differential, away from the singularities of V;
(3) on singular surfaces σ_k the maps ψ_E and the normal derivatives are continuous;
(4) at isolated singularities ψ_E admits finite limits.

The constraints are sometimes justified in a sort-of-whimsical way in textbooks (this happens in particular for the analogous statements for $\mathbb{R}^1$).

What we can say is, first, that H_0 is *not* the operator representing the Hamiltonian observable, because H_0 is not self-adjoint! The operator in question is a self-adjoint extension of H_0. Theorem 13.6 warrants, under the assumptions made, H_0 is essentially self-adjoint on $\mathscr{D}(\mathbb{R}^3)$, so there is one self-adjoint extension that coincides with the closure of H_0 and with its adjoint as well: $\overline{H_0} = H_0^*$. Stationary states are given by the spectrum of H_0^*, i.e. by solutions to

$$H_0^*\psi_E = E\psi_E\,, \quad E \in \mathbb{R}\,, \quad \psi_E \in D(H_0^*)\,.$$

This equation, since $\mathscr{D}(\mathbb{R}^3)$ is dense in $L^2(\mathbb{R}^3, dx)$, may be written:

$$(\varphi|H_0^*\psi_E) = E(\varphi|\psi_E)\,, \quad E \in \mathbb{R}\,, \quad \text{for every } \varphi \in \mathscr{D}(\mathbb{R}^3) \text{ and a given } \psi_E \in D(H_0^*).$$

Using the definition of adjoint, the equation reads

$$(H_0\varphi|\psi_E) = E(\varphi|\psi_E)\,, \quad E \in \mathbb{R}\,, \quad \text{for any } \varphi \in \mathscr{D}(\mathbb{R}^3) \text{ and a given } \psi_E \in D(H_0^*).$$

Put differently, we seek functions $\psi_E \in L^2(\mathbb{R}^3, dx)$ such that, for any $\varphi \in \mathscr{D}(\mathbb{R}^3)$:

$$\int_{\mathbb{R}^3} \left(-\frac{\hbar^2}{2m}\Delta\overline{\varphi(\mathbf{x})} + V(\mathbf{x})\overline{\varphi(\mathbf{x})} - E\overline{\varphi(\mathbf{x})} \right) \psi_E(\mathbf{x})\, dx = 0\,. \tag{13.12}$$

Hence the ψ_E do not necessarily solve $H_0^*\psi_E = E\psi_E$, for it is enough they solve it *weakly*: they must satisfy (13.12) for any $\varphi \in \mathscr{D}(\mathbb{R}^3)$. Issues of this kind [ReSi80] are dealt with by the general theory of *elliptic regularity*, which proves [CCP82] $\psi_E \in L^2(\mathbb{R}^3, dx)$ satisfies (13.12), with the aforementioned assumptions on the potential V, if and only if ψ_E satisfies conditions (1)–(4). ■

Examples 13.8. **(1)** The simplest example is the free spin-zero particle of mass $m > 0$, described on the Hilbert space $L^2(\mathbb{R}^3, dx)$ associated to the axes of an inertial system $\mathscr{I}$. Pure states are represented by *wavefunctions*, i.e. unit elements $\psi \in L^2(\mathbb{R}^3, dx)$. The Hamiltonian is simply:

$$H := \frac{1}{2m}\overline{\sum_{k=1}^{3} P_k\restriction^2_{\mathscr{S}(\mathbb{R}^3)}} = -\frac{\hbar^2}{2m}\overline{\Delta\restriction_{\mathscr{S}(\mathbb{R}^3)}} . \tag{13.13}$$

Let us briefly discuss its self-adjointness. Although everything should be clear from Proposition 10.44, we think it might be interesting to go over a few facts. The left-hand side of (13.13) is self-adjoint since

$$H_0 := \frac{1}{2m}\sum_{k=1}^{3} P_k\restriction^2_{\mathscr{S}(\mathbb{R}^3)} = -\frac{\hbar^2}{2m}\Delta\restriction_{\mathscr{S}(\mathbb{R}^3)}$$

is essentially self-adjoint. The proof is direct via the unitary Fourier-Plancherel operator $\widehat{\mathscr{F}}$, noting that in the space $L^2(\mathbb{R}^3, dk)$ of transformed maps $\widetilde{\psi} := \widehat{\mathscr{F}}(\psi)$, the above operator multiplies by:

$$\mathbf{k} \mapsto \frac{\hbar^2}{2m}\mathbf{k}^2 ,$$

and has dense domain $D(\widehat{\mathscr{F}}^{-1}H_0\widehat{\mathscr{F}}) := \mathscr{S}(\mathbb{R}^3)$. By construction $\widehat{\mathscr{F}}^{-1}H_0\widehat{\mathscr{F}}$ is symmetric, and it is easy to prove its essential self-adjointness by showing $Ker((\widehat{\mathscr{F}}^{-1}H_0\widehat{\mathscr{F}})^* \pm I) = \{0\}$, or proving each vector of $\mathscr{D}(\mathbb{R}^3) \subset \mathscr{S}(\mathbb{R}^3)$ is analytic for $\widehat{\mathscr{F}}^{-1}H_0\widehat{\mathscr{F}}$. The same holds for H_0, since $\widehat{\mathscr{F}}$ is unitary.

By construction if $H := \overline{H_0}$, then $\widetilde{H} := \widehat{\mathscr{F}}^{-1}H\widehat{\mathscr{F}}$ acts as multiplicative operator:

$$\left(\widetilde{H}\widetilde{\psi}\right)(\mathbf{k}) = \frac{\hbar^2}{2m}\mathbf{k}^2\widetilde{\psi}(\mathbf{k}) ,$$

where

$$D(\widetilde{H}) = \left\{ \widetilde{\psi} \in L^2(\mathbb{R}^3, dk) \,\middle|\, \int_{\mathbb{R}^3} |\mathbf{k}|^4 |\widetilde{\psi}(\mathbf{k})|^2 dk < +\infty \right\} .$$

An alternative definition for H comes from taking the unique self-adjoint extension of H_0 defined on $\mathscr{D}(\mathbb{R}^3)$ instead of $\mathscr{S}(\mathbb{R}^3)$:

$$H_0 := \frac{1}{2m}\sum_{k=1}^{3} P_k\restriction^2_{\mathscr{D}(\mathbb{R}^3)} = -\frac{\hbar^2}{2m}\Delta\restriction_{\mathscr{D}(\mathbb{R}^3)} .$$

However, H_0 is still essentially self-adjoint and its self-adjoint extension is the previous H. Or, we could define H_0 on $\widehat{\mathscr{F}}(\mathscr{D}(\mathbb{R}^3))$, and find the same result. All this descends from Proposition 10.44.

(2) An interesting case in $\mathbb{R}^3$ is where the free Hamiltonian is modified by the potential energy of the *attractive Coulomb potential*:

$$V(\mathbf{x}) = \frac{eQ}{|\mathbf{x}|} ,$$

where $e < 0$, $Q > 0$ are constants expressing the electric charges of the particle and the centre of attraction respectively. The assumptions of Kato's Theorem 10.49 (or 10.47) hold ($m, \hbar > 0$ are constants that play no role, since we can multiply the operator by $2m/\hbar^2$ without loss of generality). Therefore:

$$H_0 := -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})$$

is essentially self-adjoint, whether defined on $\mathscr{D}(\mathbb{R}^3)$ or $\mathscr{S}(\mathbb{R}^3)$. If $-Q = e$ is the charge of the electron ($-1.60 \cdot 10^{-19}$ C), and $m = m_e$ its mass ($9.11 \cdot 10^{-31}$ Kg), the only self-adjoint extension $\overline{H_0}$ corresponds to the Hamiltonian of an electron inside the electric field of a proton (neglecting spin effects and envisaging the proton as a classical object of infinite mass). This gives the simplest quantum description of the Hamiltonian operator of the hydrogen atom. Although V is not bounded from below, it is important to note the spectrum of the operator is always bounded, so also the admissible values of energy are constrained. This implies the hydrogen atom is an energetically stable system: it cannot collapse to decreasingly lower energy levels by emitting a quantity of energy, eventually infinite, when interacting with the electromagnetic field (i.e. in the energy of photons emitted by the atom: this way will not be treated in this physically-very-elementary book). Observe that the analogous classical model, for which the electron and the attractive centre are dimensionless points, would not have total energy bounded from below[3]. Studying the spectrum of $\overline{H_0}$ [Mes99, CCP82] shows $\sigma_c(\overline{H_0}) = [0, +\infty)$, while $\sigma_p(\overline{H_0}) = \{E_n\}_{n=1,2,\ldots}$, where

$$E_n = -\frac{2\pi R\hbar c}{n^2} \quad n = 1,2,3,\ldots \tag{13.14}$$

$R = me^4/(4\pi c\hbar^3)$ is the *Rydberg constant* and c the speed of light. Eigenvectors have a complicated expression [Mes99, CCP82]. For each of the values E_n, $n = 1,2,3$, the corresponding eigenspace has a finite basis in spherical coordinates:

$$\psi_{nlm}(r,\theta,\phi) = -\sqrt{\left(\frac{2}{na_0}\right)^2 \frac{(n-l-1)!}{2n[(n+l)!]^3}}\, e^{-\frac{r}{na_0}} \left(\frac{2r}{na_0}\right)^l L^{2l+1}_{n+l}\left(\frac{2r}{na_0}\right) Y_l^m(\theta,\phi)\,, \tag{13.15}$$

where $l = 0,1,\ldots,n-1$ and $m = -l, -l+1, \ldots, l-1, l$. The maps Y_l^m are the spherical harmonics (10.42), $a_0 = \hbar^2/e^2 m_e = 0{,}529$ Å is the radius of Bohr's first orbit and $L_n^\alpha(x)$, for $x \geq 0$, is the **Laguerre polynomial**:

$$L_n^\alpha(x) := \frac{d^\alpha}{dx^\alpha}\left[e^x \frac{d^n}{dx^n}(x^n e^{-x})\right], \quad n \in \mathbb{N},\ \alpha = 0,1,\ldots,n.$$

By examining the interaction between photons and the hydrogen atom [Mes99, CCP82] we know the electron, initially in a stationary state determined by an eigenvector of $\overline{H_0}$ with eigenvalue E_n, can change state and pass to a new stationary state of

[3] Such a classical model would not, anyway, be consistent because of the *Bremsstrahlung* of the accelerated electron; as is well known, this fact produces mathematical inconsistencies when the electronic radius tends to zero.

energy $E_m < E_n$ transferring the energy excess to a photon. The reverse process may occur, whereby the electron acquires energy from a photon and passes from states of energy E_m to E_n. Due to the interactions with photons, it can be proved that only the state of minimum energy $E_1 = 2\pi R\hbar c$, the so-called **ground state**, is stable, while the others are all unstable, so the electron decays to the ground state after a certain mean lifetime to be determined. (Therefore the name *stationary state* is not completely appropriate for the system formed by an atom and the electromagnetic field described by photons, and one should rather just speak of eigenvalues of the Hamiltonian for the hydrogen atom.) The collection of energy differences $E_n - E_m$ determine all possible photonic, i.e. light, frequencies that a gas of hydrogen atoms can emit or absorb, by Einstein's formula $E_n - E_m = h\nu_{n,m}$. The latter relates the frequency $\nu_{n,m}$ of photons emitted by the atom to the energy needed by photons that switch from energy E_n to E_m (see Chapter 6). XIX century spectroscopists, though puzzled by them, knew the frequency values $\nu_{n,m}$, long before QM was formulated [Mes99, CCP82]. Finding the same values and being able to *explain them* in a completey theoretical manner is certainly one of the pinnacles of the past century's physics.

(3) A second interesting situation, in $\mathbb{R}^3$, is that in which to the Hamiltonian of the free particle of example (1) we add the *Yukawa potential*:

$$V(\mathbf{x}) = \frac{-e^{-\mu|\mathbf{x}|}}{|\mathbf{x}|},$$

where $\mu > 0$ is a positive constant. Here, too, $H_0 = -\frac{\hbar^2}{2m}\Delta + V(x)$ is essentially self-adjoint if either defined on $\mathscr{D}(\mathbb{R}^3)$ or on $\mathscr{S}(\mathbb{R}^3)$, because of Kato's Theorem 10.49 (or 10.47). The Yukawa potential describes, roughly speaking, the interaction processes between a *pion* and a *strong force* originating from a macroscopic source.

(4) Referring to example (1), the action of the evolution operator is evident using Fourier's representation:

$$(\widetilde{U}_t\widetilde{\psi})(\mathbf{k}) = \left(e^{-\frac{it}{\hbar}\widetilde{H}}\widetilde{\psi}\right)(\mathbf{k}) = e^{-\frac{it\hbar}{2m}\mathbf{k}^2}\,\widetilde{\psi}(\mathbf{k})\,. \tag{13.16}$$

The proof is immediate from the spectral decompositions of $\widetilde{H}$ and the commutation of the spectral measures of P_1, P_2, P_3:

$$e^{-\frac{it}{\hbar}\widetilde{H}} = e^{-\frac{it}{2\hbar m}\widetilde{P}_1^2}\,e^{-\frac{it}{2\hbar m}\widetilde{P}_2^2}\,e^{-\frac{it}{2\hbar m}\widetilde{P}_3^2}\,,$$

where each $\widetilde{P}_j = \widehat{\mathscr{F}}^{-1}P_j\widehat{\mathscr{F}}$ multiplies by

$$\left(\widetilde{P}_j\widetilde{\psi}\right)(\mathbf{k}) = \hbar k_j\,\widetilde{\psi}(\mathbf{k})\,.$$

Back in position representation we look at the evolution of a wavefunction determining the state $U_t\rho U_t^*$ when $\rho = \psi(\psi|\ )$. This is

$$\psi(t,\mathbf{x}) := \left(e^{-i\frac{t}{\hbar}H}\psi\right)(\mathbf{x}) = \int_{\mathbb{R}^3}\frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}}\,\widetilde{\psi}(\mathbf{k})e^{-i\frac{\hbar t}{2m}\mathbf{k}^2}\,dk \tag{13.17}$$

where

$$\psi(\mathbf{x}) = \psi(0,\mathbf{x}) := \int_{\mathbb{R}^3} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}} \widetilde{\psi}(\mathbf{k})\, dk\,, \tag{13.18}$$

for $\psi \in \mathscr{S}(\mathbb{R}^3)$. In general the integrals should be understood in the sense of the Fourier-Plancherel transform.

(5) In the previous example equation (13.17) can be written without Fourier transforming the initial datum ψ, as this proposition establishes.

Proposition 13.9. *Take* $\psi \in \mathscr{S}(\mathbb{R}^3)$ *and* $H = \overline{-\frac{\hbar}{2m}\Delta}$, $\hbar, m > 0$ *(the Laplacian* Δ *is initially defined on* $\mathscr{S}(\mathbb{R}^3)$ *or equivalently* $\mathscr{D}(\mathbb{R}^3)$*).*

(a) *For any given* $t \in \mathbb{R}$*, the map* $\psi(t,\mathbf{x}) := \left(e^{-i\frac{t}{\hbar}H}\psi\right)(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$*, belongs to* $\mathscr{S}(\mathbb{R}^3)$.

(b) *If* $t \neq 0$ *and* $\mathbf{x} \in \mathbb{R}^3$*:*

$$\psi(t,\mathbf{x}) = \left(\frac{m\hbar}{2\pi i t}\right)^{3/2} \int_{\mathbb{R}^3} e^{im\hbar|\mathbf{x}-\mathbf{y}|^2/(2t)}\psi(\mathbf{y})dy \tag{13.19}$$

where the multi-valued square root is computed by branching the complex plane along the negative real axis.

(c) *Let* $C_\psi := \left(\frac{m\hbar}{2\pi}\right)^{3/2} \int_{\mathbb{R}^3} |\psi(\mathbf{x})|dx$*. Then*

$$||\psi(t,\cdot)||_\infty \le C_\psi |t|^{-3/2} \quad \text{for every } t \neq 0\,. \tag{13.20}$$

Proof. (a) The Fourier transform $\widetilde{\psi}$ of $\psi \in \mathscr{S}(\mathbb{R}^3)$ is in $\mathscr{S}(\mathbb{R}^3)$. Multipliying by the exponential $e^{-i\hbar\mathbf{k}^2/(2m)}$ produces a map of $\mathscr{S}(\mathbb{R}^3)$. Since $\mathscr{S}(\mathbb{R}^3)$ is Fourier-invariant, equation (13.17) implies $\psi(t,\cdot) \in \mathscr{S}(\mathbb{R}^3)$.

(b) Equation (13.17) can be rewritten using the Fourier transform and Lebesgue's dominated convergence:

$$\psi(t,\mathbf{x}) = \int_{\mathbb{R}^3} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}} e^{-i\frac{\hbar t}{2m}\mathbf{k}^2} \left(\int_{\mathbb{R}^3} \frac{e^{-i\mathbf{k}\cdot\mathbf{y}}}{(2\pi)^3}\psi(\mathbf{y})\,dy\right) dk$$

$$= \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^3} \frac{e^{i\mathbf{k}\cdot\mathbf{x}}}{(2\pi)^{3/2}} e^{-i\frac{\hbar(t-i\varepsilon)}{2m}\mathbf{k}^2} \left(\int_{\mathbb{R}^3} \frac{e^{-i\mathbf{k}\cdot\mathbf{y}}}{(2\pi)^3}\psi(\mathbf{y})\,dy\right) dk\,.$$

If $\varepsilon > 0$, Fubini–Tonelli allows to write

$$\psi(t,\mathbf{x}) = \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^3} \left(\int_{\mathbb{R}^3} e^{i(\mathbf{k}\cdot(\mathbf{x}-\mathbf{y}) - \hbar(t-i\varepsilon)\mathbf{k}^2/(2m))}dk\right)\psi(\mathbf{y})\,dy\,.$$

The inner Gaussian integral can be computed explicitly (e.g., with residue techniques):

$$\psi(t,\mathbf{x}) = \lim_{\varepsilon\to 0^+} \left(\frac{m\hbar}{2\pi i(t-i\varepsilon)}\right)^{3/2} \int_{\mathbb{R}^3} e^{im\hbar|\mathbf{x}-\mathbf{y}|^2/(2(t-i\varepsilon))}\psi(\mathbf{y})dy\,.$$

For $t \neq 0$ and $\mathbf{x} \in \mathbb{R}^3$ fixed, we can take the limit inside the integral due to dominated convergence: the integrand, in fact, is in absolute value smaller than $|\psi| \in L^1(\mathbb{R}^3, dx)$, uniformly in $\varepsilon > 0$. This gives (13.19).

(c) Follows from (b) directly. □

Equation (13.19) holds on $\mathbb{R}^d$ if we replace the exponent $3/2$ with $d/2$. (The wavefunctions $\psi(t,\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$, evolve under the evolution operator generated by the self-adjoint closure of $-\frac{1}{2m}\Delta$, where $\Delta : \mathscr{S}(\mathbb{R}^d) \to L^2(\mathbb{R}^d, dx)$ is the Laplacian in d dimensions.)

Since the integral of $|\psi(t,\mathbf{x})|^2$ over $\mathbb{R}^3$ is constant in time, and $|\psi(t,\mathbf{x})|^2$ at any point $\mathbf{x} \in \mathbb{R}^3$ is infinitesimal by (13.20), a wavefunction that is initially non-zero on a small region in space must increase its support as time goes by, and "spread out" to increasingly larger regions. ■

13.1.4 The action of the Galilean group in position representation

Example 13.8(4) allows to make explicit, in position representation, the Galilean group's action, which we saw at the end of Chapter 12 in momentum representation for the free particle of spin s. If $(\tau, \mathbf{c}, \mathbf{v}, U)$ is the generic element of the universal covering $\widetilde{S\mathscr{G}}$ of the restricted Galilean group, the mentioned representation is induced by the unitary operators $Z^{(m)}_{(\tau,\mathbf{c},\mathbf{v},U)}$ that act, in momentum representation, as (12.107):

$$\left(\widetilde{Z^{(m)}}_{(\tau,\mathbf{c},\mathbf{v},U)}\widetilde{\psi}\right)(\mathbf{k}) := e^{i(\tau\mathbf{v}-\mathbf{c})\cdot(\mathbf{k}-m\mathbf{v})} e^{i\frac{\tau}{2m}(\mathbf{k}-m\mathbf{v})^2}\,\widetilde{\psi}\left(R(U)^{-1}(\mathbf{k}-m\mathbf{v})\right) .$$

In position representation, anti-transforming with Fourier-Plancherel $\psi = \widehat{\mathscr{F}}^{-1}\widetilde{\psi}$, easily gives

$$\left(U_t Z^{(m)}_{(\tau,\mathbf{c},\mathbf{v},U)}\psi\right)(\mathbf{x}) = e^{im(\mathbf{v}\cdot\mathbf{x}-\mathbf{v}^2 t/2)}\psi\left(t-\tau, R(U)^{-1}(\mathbf{x}-\mathbf{c}) - (t-\tau)R(U)^{-1}\mathbf{v}\right)$$

for $\psi \in L^2(\mathbb{R}^2, dx)$. Put otherwise, if $\psi'(t,\mathbf{x}) := \left(U_t Z^{(m)}_{(\tau,\mathbf{c},\mathbf{v},U)}\psi\right)(\mathbf{x})$ is the wavefunction acted upon by the element $(\tau, \mathbf{c}, \mathbf{v}, U)$ of the (universal covering of the) Galilean group at $t = 0$, which evolves to time t, we have:

$$\psi'(t,\mathbf{x}) = e^{im(\mathbf{v}\cdot\mathbf{x}-\mathbf{v}^2 t/2)}\psi\left((\tau,\mathbf{c},\mathbf{v},U)^{-1}(t,\mathbf{x})\right) \tag{13.21}$$

by (12.102). For particels with spin s, as we saw in the previous chapter, for fixed inertial frame the Hilbert space is $L^2(\mathbb{R}^3, dx) \otimes \mathbb{C}^{2s+1}$ and wavefunctions are unit vectors

$$\Psi = \sum_{s_z=-s}^{s} \psi_{s_z} \otimes |s, s_z\rangle ,$$

where $|s, s_z\rangle$ form the canonical basis of $\mathbb{C}^{2s+1}$ in which the spin operator S_z is diagonal with eigenvalues s_z.

By this decomposition $L^2(\mathbb{R}^3, dx) \otimes \mathbb{C}^{2s+1}$ becomes naturally isomorphic to the orthogonal sum of $2s+1$ copies of $L^2(\mathbb{R}^3, dx)$; consequently, the vectors Ψ identify *spinors* of order s, that is, column vectors of wavefunctions for particles without spin:

$$\Psi \equiv (\psi_s, \psi_{s-1}, \cdots, \psi_{-s+1}, \psi_{-s})^t .$$

Similarly, let $\Psi'_t := \left(U_t Z^{(m)}_{(\tau,\mathbf{c},\mathbf{v},U)} \otimes V^{(s)}(U)\Psi\right)$, where $V^{(s)}(U)$ is the action of $U \in SU(2)$ on spinors for particles of spin s (cf. Chapter 12.3.1). Then the active Galilean action, in terms of spinors, reads:

$$\psi'_{s'_z}(t,\mathbf{x}) = e^{im(\mathbf{v}\cdot\mathbf{x}-\mathbf{v}^2t/2)} \sum_{s_z=-s}^{s} V^{(s)}(U)_{s'_z s_z} \psi_{s_z}\left((\tau,\mathbf{c},\mathbf{v},U)^{-1}(t,\mathbf{x})\right) , \tag{13.22}$$

where $V^{(s)}(U)_{ij}$ is the matrix entry of $V^{(s)}(U)$ in the canonical basis of $\mathbb{C}^{2s+1}$.

Now think Galilean transformations *passively*, hence view the $Z^{(m)}_{(\tau,\mathbf{c},\mathbf{v},U)}$ as unitary operators between *distinct Hilbert spaces associated to different frame systems* that describe the *same* physical system. We can thus describe the transformations of quantum states between different frame systems. The basic idea is that when acting on a state by an active Galilean transformation, and then changing to the transformed reference system by the *same* active map, in the *new* frame the transformed state must look like the original, pre-transformation, one. Therefore the law of passive transformations of states (coordinate change) corresponds to the inverse active transformation seen above, meaning that we replace $(\tau,\mathbf{c},\mathbf{v},U)$ with $(\tau,\mathbf{c},\mathbf{v},U)^{-1}$ in (13.22). Let us see this recipe implemented. Take two inertial frames $\mathscr{I}$, $\mathscr{I}'$ with right-handed Cartesian coordinates x_1,x_2,x_3 and x'_1,x'_2,x'_3 and time coordinate t, t' respectively. Suppose the coordinate change is the Galilean transformation:

$$\begin{cases} t' = t+\tau\,, \\ x'_i = c_i + tv_i + \sum_{j=1}^{3} R_{ij}x_j\,, \quad \text{i=1,2,3} \end{cases} \tag{13.23}$$

where $\tau \in \mathbb{R}$, $c_i \in \mathbb{R}$, $v_i \in \mathbb{R}$, $R \in SO(3)$. Consider a particle of spin s, so the theory's Hilbert space is $\mathsf{H} := L^2(\mathbb{R}^3,dx)\otimes\mathbb{C}^{2s+1}$ for $\mathscr{I}$, and $\mathsf{H}' := L^2(\mathbb{R}'^3,dx')\otimes\mathbb{C}^{2s+1'}$ for $\mathscr{I}'$. The spaces $\mathbb{R}^3$ and $\mathbb{R}'^3$ are identified with the rest spaces of their frame systems by the coordinates. The canonical bases of $\mathbb{C}^{2s+1}$, $\mathbb{C}^{2s+1'}$ are the eigenvector bases of the spin operators along the third axes, S_3 and $S_{3'}$. Choose a matrix $U \in SU(2)$ whose image under the covering of $SO(3)$ is R. (Note the parameters $\mathbf{v},U$ also show up in the phase factor, and are given up to sign, as seen in the previous chapter: this sign may change the vectors representing a pure state, but does not alter the state itself.) Consider a pure state described in $\mathscr{I}$ by the unit vector Ψ and its evolution in $\mathscr{I}$. The state Ψ corresponds to a state Ψ' in $\mathscr{I}'$, together with its evolution. The relationship between the spinors Ψ and Ψ' evolves according to

$$\psi'_{s'_z}(t',\mathbf{x}') = e^{-im(\mathbf{v}\cdot R(U)\mathbf{x}+\mathbf{v}^2t/2)} \sum_{s_z=-s}^{s} V^{(s)}(U)_{s'_z s_z} \psi_{s_z}(t+\tau, R(U)\mathbf{x}+\tau\mathbf{v}+\mathbf{c}) , \tag{13.24}$$

obtained replacing $(\tau,\mathbf{c},\mathbf{v},U)$ with $(\tau,\mathbf{c},\mathbf{v},U)^{-1}$ in (13.22) (the parameters $\mathbf{v},U$ also appear in the phase, and the ones of the inverse Galilean transformation must be used). For spin $s=0$, in particular:

$$\psi'(t',\mathbf{x}') = e^{-im(\mathbf{v}\cdot R(U)\mathbf{x}+\mathbf{v}^2t'/2)} \psi(t+\tau, R(U)\mathbf{x}+\tau\mathbf{v}+\mathbf{c}) , \tag{13.25}$$

where the coordinates $(t,\mathbf{x})$ and $(t',\mathbf{x}')$ are related by (13.23).

Remarks 13.10. Notice how the term $e^{-im(\mathbf{v}\cdot R(U)\mathbf{x}+\mathbf{v}^2t/2)}$ *cannot* be removed by taking another representative for the projective ray, since the phase depends on the variable $\mathbf{x}$). The found equation, therefore, is not the transformation we would expect, intuitively, if imagining that the zero-spin wavefunction and each component of the wavefunction with spin $s \neq 0$ are *scalar fields* on the spacetime of classical physics. The scalar field interpretation of wavefunctions in position representation is *a priori* not automatic, and totally false (not just for one choice of phase) in relativistic theories, where wavefunctions in position representation (within the so-called *Newton-Wigner formalism* [BaRa86]) are highly nonlocal objects[4]. ■

13.1.5 Basic notions of scattering processes

Consider a quantum system, for instance a single quantum particle, described on the Hilbert space H (after an inertial system has been fixed) and whose evolution is given by a Hamiltonian operator $H = H_0 + V$; H_0 is the Hamiltonian of the "non-iteracting theory" that we may think, to fix ideas, as described by the purely kinetic Hamiltonian of the particle, even if we could consider more involved multi-particle quantum systems. V represents therefore the interaction with an external field or the self-interaction, often unknown or partially known. In certain circumstances, in the distant past or future a state described by ψ behaves "as if it evolved" under the non-interacting Hamiltonian H_0. This happens typically in *scattering* processes.

Consider for example one particle: initially free, it interacts briefly with a scattering centre – a system we can treat as semi-classical – and then returns free. Experimentally speaking, we can say the system is prepared at $t \to -\infty$ in an approximatively free state, and after the interaction, as $t \to +\infty$, it manifests itself in a state that can still be seen as free. Examining the difference between prepared state and observed one gives informations on the structure of the scattering centre, and more generally on the type of interaction described by V. In more complicated situations there is no scattering centre, and one has to deal with two or more particles, even a system with an unknown number of particles, that (self-)interact very briefly and return swiftly to a non-interacting setup.

We will introduce the basic mathematical ideas to formalise all that, referring the reader to advanced texts for details and generalisations to several particles (or relativistic processes with unknown number of particles) [ReSi80, Pru81, Mes99, CCP82]. The fact that *for certain state vectors in the system*, generically indicated by ϕ, the evolution in time is approximated by the non-interacting evolution in the far future, is expressed by

$$\lim_{t\to+\infty} ||e^{-itH}\phi - e^{-itH_0}\psi|| = 0 \tag{13.26}$$

for some state ψ distinct from, but dertemined by, ϕ. Equivalently, since e^{itH} is unitary:

$$\lim_{t\to+\infty} ||\phi - e^{itH}e^{-itH_0}\psi|| = 0\,. \tag{13.27}$$

---

[4] One should not confuse a wavefunction in position representation with the field of second quantisation, which is a local object instead.

The argument can be clearly replicated for $t \to -\infty$, i.e. describing what happens long before the interaction takes place, when the evolution is taken to be free. For several reasons, both theoretical and experimental, it is convenient to describe scattering using vectors like ψ, *that evolve by the Hamiltonian of the non-interacting theory*, rather than ϕ, *which evolves under the interacting Hamiltonian.* This motivates the introduction of **wave operators** $\Omega_\pm$, also known as **Møller operators**:

$$\Omega_\pm \psi := \lim_{t\to\pm\infty} e^{itH} e^{-itH_0} \psi\,, \quad \psi \in \mathsf{H}, \tag{13.28}$$

assuming the limit exists.

If the operators $\Omega_\pm : \mathsf{H} \to \mathsf{H}$ exist they must be *isometries*, since strong limits of unitary operators. More precisely $\Omega_\pm$ are *partial isometries* (Definition 3.61) with initial space the whole H. Consequently the final spaces

$$\mathsf{H}_\pm := Ran(\Omega_\pm) \tag{13.29}$$

are closed in H by Proposition 3.62. By construction if $\phi_\pm \in \mathsf{H}_\pm$, so $\phi_\pm = \Omega_\pm \psi$ for some $\psi \in \mathsf{H}$, it follows [5]

$$||e^{-itH}\phi_\pm - e^{-itH_0}\psi|| \to 0 \quad \text{as } t \to \pm\infty. \tag{13.30}$$

Hence $\mathsf{H}_\pm$ determine the class of states whose long-time future or past evolution can be approximated by the free evolution of the states obtained by swapping the $\Omega_\pm$. Equation (13.30), exactly as we wanted, tells that the state *of the interacting system* $\phi_\pm$, evolving under the full Hamiltonian $H = H_0 + V$, has the asymptotic behaviour (for $t \to \pm\infty$ respectively) of the state ψ *in the non-interacting system*, which evolves under the free Hamiltonian H_0.

In details:

Proposition 13.11. *If the surjective isometry* $\Omega_\pm : \mathsf{H} \to \mathsf{H}_\pm$ *of (13.28) is defined, then*

$$e^{-itH}\Omega_\pm = \Omega_\pm e^{-itH_0}\,. \tag{13.31}$$

Consequently

$$e^{-itH}\mathsf{H}_\pm \subset \mathsf{H}_\pm\,, \quad e^{-itH}\restriction_{\mathsf{H}_\pm} = \Omega_\pm e^{-itH_0}\Omega_\pm^{-1}\,, \quad H\restriction_{\mathsf{H}_\pm\cap D(H)} = \Omega_\pm H_0 \Omega_\pm^{-1}\,, \tag{13.32}$$

and in particular:

$$\sigma(H\restriction_{\mathsf{H}_\pm\cap D(H)}) = \sigma(H_0)\,. \tag{13.33}$$

Proof. As for the first statement

$$e^{-itH}\Omega_\pm\psi = \lim_{s\to\pm\infty} e^{i(s-t)H}e^{-isH_0}\psi = \lim_{z\to\pm\infty} e^{izH}e^{-izH_0}e^{-itH_0}\psi = \Omega_\pm e^{-itH_0}\psi\,,$$

whence $e^{-itH}\mathsf{H}_\pm \subset \mathsf{H}_\pm$ and $e^{-itH}\restriction_{\mathsf{H}_\pm} = \Omega_\pm e^{-itH_0}\Omega_\pm^{-1}$. Stone's theorem easily implies the other relation. Eventually, $\Omega_\pm : \mathsf{H} \to \mathsf{H}_\pm$ isometric proves (13.33) by the last identity in (13.32) (Exercise 8.9). □

[5] In fact $||e^{-itH}\Omega_\pm\psi - e^{-itH_0}\psi|| = ||\Omega_\pm\psi - e^{itH}e^{-itH_0}\psi|| \to ||\Omega_\pm\psi - \Omega_\pm\psi|| = 0$ as $t \to \pm\infty$.

For the usual non-relativistic particle the non-interacting Hamiltonian H_0, that accounts for the kinetic energy only, has spectrum $\sigma(H_0) = \sigma_c(H_0) = [0, +\infty)$. Under the above proposition's assumptions, then, $\sigma(H\restriction_{\mathsf{H}_\pm \cap D(H)}) = \sigma_c(H\restriction_{\mathsf{H}_\pm \cap D(H)}) = [0 + \infty)$.

So let us assume the wave operators $\Omega_\pm : \mathsf{H} \to \mathsf{H}_\pm$ exist on some physical system, and suppose the system has been prepared in a state that, as $t \to -\infty$, tends to be described by the non-interacting evolution of $e^{-itH_0}\psi_{in}$. Hence the system's real state will be described, at $t = 0$, by the state $\phi_- := \Omega_- \psi_{in}$. After the interaction, as $t \to +\infty$, the state will be described approximatively by a non-interacting vector $e^{-itH_0}\psi_{out}$. The real state, at $t = 0$, is described by $\phi_+ := \Omega_+ \psi_{out}$. The probability of the process is thus:

$$|(\phi_+|\phi_-)|^2 = |(\Omega_+ \psi_{out}|\Omega_- \psi_{in})|^2 = |(\psi_{out}|\Omega_+^* \Omega_- \psi_{in})|^2 .$$

Define the **scattering operator**, also called S **matrix** :

$$S := \Omega_+^* \Omega_- : \mathsf{H} \to \mathsf{H} . \tag{13.34}$$

The transition amplitude from a state that behaves as a non-interacting state $e^{-itH_0}\psi_{in}$, as $t \to -\infty$, to the state that behaves like a non-interacting state $e^{-itH_0}\psi_{out}$ as $t \to +\infty$, equals:

$$(\psi_{out}|S\psi_{in}) . \tag{13.35}$$

In this picture, the interaction V is completely "withheld" by S, while we can consider the states $\psi_{in/out}$ as being indeed free. Overall we have, as we were saying at the beginning, a recipe to describe the scattering in terms of states in a non-interacting system. To conclude, we have a proposition.

Proposition 13.12. *If the surjective isometries* $\Omega_\pm : \mathsf{H} \to \mathsf{H}_\pm$ *of (13.28) exist, and* $\mathsf{H}_+ = \mathsf{H}_-$ *– in particular under asymptotic completeness (see Remark 13.13(1)) – the scattering operator (13.34) is unitary.*

Proof. It is enough to prove $S^*S = SS^* = I$. Since $\Omega_\pm$ is a partial isometry with initial space H and final space $\mathsf{H}_\pm$, by Proposition 3.63

$$\Omega_\pm^* \Omega_\pm = I , \quad \Omega_\pm \Omega_\pm^* = P_{\mathsf{H}_\pm}$$

where $P_{\mathsf{H}_\pm} : \mathsf{H} \to \mathsf{H}$ is the orthogonal projector on $\mathsf{H}_\pm$. Therefore

$$S^*S = \Omega_-^* \Omega_+ \Omega_+^* \Omega_- = \Omega_-^* P_{\mathsf{H}_+} \Omega_- = \Omega_-^* P_{\mathsf{H}_-} \Omega_- = \Omega_-^* \Omega_- = I .$$

Similarly

$$SS^* = \Omega_+^* \Omega_- \Omega_-^* \Omega_+ = \Omega_+^* P_{\mathsf{H}_-} \Omega_+ = \Omega_+^* P_{\mathsf{H}_+} \Omega_+ = \Omega_+^* \Omega_+ = I ,$$

ending the proof. □

Remark 13.13. **(1)** Next to $\mathsf{H}_\pm$ it is useful to introduce the **space of stationary states** H_p given by the closure of the span of eigenvectors of H, which describe stationary states (see Remark 9.12(2)). Physically, one expects elements $\phi \in \mathsf{H}_p$ to represent precisely states whose evolution *cannot* be approximated, at large times, by non-interacting states. That is because, in particular, the evolution of such a state $\phi(\phi|\ )$ (with $||\phi|| = 1$) is trivial, for ϕ is an eigenvector of H. What we expect, said more accurately, is the orthogonal sum

$$\mathsf{H} = \mathsf{H}_\pm \oplus \mathsf{H}_p \,. \tag{13.36}$$

In this case one speaks about **asymptotic completeness**. Note that (13.36) implies:

$$\mathsf{H}_+ = \mathsf{H}_- = \mathsf{H}_p^\perp \tag{13.37}$$

and by (13.33), also

$$\sigma_{pc}(H) = \sigma(H_0) \tag{13.38}$$

(Remark 9.12(2)). At last, asymptotic completeness and (13.37) make the operator S unitary by Proposition 13.12.
(2) The next easy result relates the orthogonality of $\mathsf{H}_\pm$ and H_p with the properties of the evolution operator generated by H_0.

Proposition 13.14. *If the surjective isometry $\Omega_\pm : \mathsf{H} \to \mathsf{H}_\pm$ of (13.28) exists, and $(\psi|e^{-itH_0}\psi') \to 0$ as $t \to \pm\infty$ for any $\psi, \psi' \in \mathsf{H}$, then $\mathsf{H}_\pm \perp \mathsf{H}_p$.*

Proof. Define $\phi_\pm := \Omega_\pm \psi$ and suppose $H\phi_E = E\phi_E$. Then

$$(\phi_\pm|\phi_E) = \lim_{t\to\pm\infty}(e^{itH}e^{-itH_0}\psi|\phi_E) = \lim_{t\to\pm\infty}(e^{-itH_0}\psi|e^{itH}\phi_E)$$

$= \lim_{t\to\pm\infty} e^{-iEt}(e^{-itH_0}\psi|\phi_E) = 0.$ □

(3) The short and compressed description of scattering theory we have presented does not work in quantum field theory, because of the *impossibility* of defining the unitary operators $\Omega_\pm$ under simple, physically plausible hypotheses on the theory's invariance under the group of space translations (spatial homogeneity). The obstruction is exquisitely theoretical and goes under the name of *Haag theorem* [Haa96]. In order to overcome the problem we can turn to the *LSZ formalism* [Haa96], in which scattering descriptions employ the weak topology. However, these issues assume an ivory-tower flavour, so to speak, when compared to the much more serious problem of *renormalisation*. ■

Example 13.15. Take a free spin-zero particle (in the sequel $\hbar = 1$) of mass $m > 0$, subject to a square-integrable potential V on $\mathbb{R}^3$ in a given inertial system. Then $\mathsf{H} = L^2(\mathbb{R}^3, dx)$, $H_0 = \overline{-\frac{1}{2m}\Delta}$ (the Laplacian Δ is as usual initially defined on $\mathscr{D}(\mathbb{R}^3)$ or $\mathscr{S}(\mathbb{R}^3)$), and $V \in L^2(\mathbb{R}^3)$. By Theorem 10.47 (redefining the coordinates of $\mathbb{R}^3$ so to comprise the factor $(2m)^{-1}$) $H = H_0 + V$ is self-adjoint on $D(H_0)$, so $D(H) = D(H_0)$. We wish to show the wave operators $\Omega_\pm$ are well defined and that $\mathsf{H}_\pm \perp \mathsf{H}_p$.

First, a technical lemma.

Lemma 13.16. *If* $\mathsf{H} = L^2(\mathbb{R}^3, dx)$*,* $H_0 = \overline{-\frac{1}{2m}\Delta}$ *and* $V \in L^2(\mathbb{R}^3)$*, let* $H = H_0 + V$ *and* $U_0(t) := e^{-itH_0}$*,* $U(t) := e^{-itH}$*. Then*

$$\frac{d}{dt}U(-t)U_0(t)\psi = U(-t)iVU_0(t)\psi\,, \quad \psi \in D(H_0) = D(H). \tag{13.39}$$

Hence, for $T > t$*:*

$$||(U(-T)U_0(T) - U(-t)U_0(t))\psi)|| \leq \int_t^T ||VU_0(s)\psi||ds\,. \tag{13.40}$$

Proof. Set $\Omega_t := U(-t)U_0(t)$. Then

$$\frac{d}{dt}\Omega_t\psi = \lim_{h\to 0}\frac{U(-(t+h))U_0(t+h) - U(-t)U_0(t)}{h}\psi\,.$$

Decompose the limit

$$\frac{d}{dt}\Omega_t\psi = \lim_{h\to 0}\frac{U(-(t+h))(U_0(t+h) - U_0(t))}{h}\psi + \lim_{h\to 0}\frac{(U(-(t+h)) - U(-t))U_0(t)}{h}\psi\,.$$

Since $U_0(t)D(H_0) \subset D(H_0) = D(H)$, Stone's theorem shows the second limit equals

$$U(-t)iHU_0(t)\psi\,.$$

As for the first limit, we compute the norm squared of the difference between $-iU(-t)H_0U_0(t)\psi$ and $\frac{U(-(t+h))(U_0(t+h)-U_0(t))}{h}\psi$, by Stone's theorem and the unitarity of $U(-(t+h))$:

$$\lim_{h\to 0}\frac{U(-(t+h))(U_0(t+h) - U_0(t))}{h}\psi = -iU(-t)H_0U_0(t)\psi\,.$$

As $H - H_0 = V$, we have:

$$\frac{d}{dt}\Omega_t\psi = U(-t)iVU_0(t)\psi\,,$$

for $\psi \in D(H_0)$, so for any $\phi \in \mathsf{H}$

$$\frac{d}{dt}(\phi|\Omega_t\psi) = (\phi|U(-t)iVU_0(t)\psi)\,.$$

The right-hand side is continuous, so the fundamental theorem of calculus gives

$$(\phi|\Omega_T\psi) - (\phi|\Omega_t\psi) = \int_t^T (\phi|U(-s)iVU_0(s)\psi)ds\,.$$

But $U(s)$ is unitary,

$$|(\phi|(\Omega_T - \Omega_t)\psi)| \leq \int_t^T ||\phi||||VU_0(s)\psi||ds\,,$$

and choosing $\phi = (\Omega_T - \Omega_t)\psi$ we recover (13.40):

$$||(\Omega_T - \Omega_t)\psi)|| \leq \int_t^T ||VU_0(s)\psi||ds\,,$$

ending the proof. □

This lemma allows to prove the existence of wave operators on the system considered.

Proposition 13.17. *Take* $\mathsf{H} = L^2(\mathbb{R}^3, dx)$, $H_0 = \overline{-\frac{1}{2m}\Delta}$, $V \in L^2(\mathbb{R}^3)$, *and consider the self-adjoint operator* $H = H_0 + V$.

(a) *The wave operators* $\Omega_\pm : \mathsf{H} \to \mathsf{H}_\pm$ *in (13.28) are well defined.*
(b) $\mathsf{H}_\pm \perp \mathsf{H}_p$.

Remarks 13.18. The theorem applies to the special case where V is a Yukawa potential (Example 13.8(3)). A stronger conclusion one can reach is asymptotic completeness, hence unitarity, of the scattering operator by assuming $V \in L^1(\mathbb{R}^3, dx) \cap L^2(\mathbb{R}^3, dx)$, as for the Yukawa potential [ReSi80]. ■

Proof. (a) Let us begin with the existence of Ω_+, for Ω_- is similar. If $\psi \in \mathscr{S}(\mathbb{R}^3) \subset D(H_0) = D(H)$, estimate (13.40) implies immediately:

$$||(U(-T)U_0(T) - U(-t)U_0(t))\psi)|| \leq \int_t^T ||V||_2 ||U_0(s)\psi||_\infty ds$$

because if $\psi \in \mathscr{S}(\mathbb{R}^3)$ then $U_0(t)\psi \in \mathscr{S}(\mathbb{R}^3)$ (cf. Example 13.8(5)). Using (13.20) we find:

$$||(U(-T)U_0(T)\psi - U(-t)U_0(t)\psi)|| \leq 2C_{V,\psi}\left(\frac{1}{\sqrt{|t|}} - \frac{1}{\sqrt{|T|}}\right). \tag{13.41}$$

This shows every sequence of vectors $\psi_n := U(-t_n)U_0(t_n)\psi$ is a Cauchy sequence when $t_n \to +\infty$ for $n \to +\infty$, so it converges to $\phi \in \mathsf{H}$. On the other hand equation (13.41) proves such limit does not depend on the sequence chosen. Hence if $\psi \in \mathscr{S}(\mathbb{R}^3)$ there exist a (unique) $\phi \in \mathsf{H}$ so that

$$\lim_{t\to+\infty} U(-t)U_0(t)\psi = \phi\,. \tag{13.42}$$

This extends easily to $\psi \in \mathsf{H}$, because $\mathscr{S}(\mathbb{R}^3)$ is dense in H. Let us prove the latter assertion. Set $\Omega_t := U(-t)U_0(t)$. By the above considerations $\Omega'\psi := \lim_{t\to+\infty}\Omega_t\psi$ is well defined provided $\psi \in \mathscr{S}(\mathbb{R}^3)$. Since this space is dense in H and every Ω_t is isometric, the operator Ω' extends to a linear isometry on H. To conclude it suffices to prove $\Omega_t\psi \to \Omega'\psi$, $t \to +\infty$, for any $\psi \in \mathsf{H}$. If $\psi \in \mathsf{H}$ consider a sequence $\{\psi_n\}_{n\in\mathbb{N}} \subset \mathscr{S}(\mathbb{R}^3)$ with $\psi_n \to \psi$, $n \to +\infty$ in H. Then

$$||\Omega_t\psi - \Omega'\psi|| \leq ||\Omega_t\psi - \Omega_t\psi_n|| + ||\Omega_t\psi_n - \Omega'\psi_n|| + ||\Omega'\psi_n - \Omega'\psi||\,.$$

Since Ω_t and Ω' are isometric, we can rewrite it as

$$||\Omega_t\psi - \Omega'\psi|| \leq ||\psi - \psi_n|| + ||\Omega_t\psi_n - \Omega'\psi_n|| + ||\psi_n - \psi||\,.$$

Given $\varepsilon > 0$, choose $n \in \mathbb{N}$ large enough so that $||\psi - \psi_n|| < 2\varepsilon/3$. For that n, by the first part of the proof, we can pick $T \in \mathbb{R}$ so that $||\Omega_t\psi_n - \Omega'\psi_n|| < \varepsilon/3$ for $t > T$. Hence we can determine, for every $\varepsilon > 0$, $T \in \mathbb{R}$ such that $t > T$ gives $||\Omega_t\psi - \Omega'\psi|| \leq \varepsilon$. And this holds for any $\psi \in \mathsf{H}$, ending (a).

(b) It is enough to prove, in our hypotheses, that Proposition 13.14 holds. We show that this descends from (13.20). Fix $\psi, \phi \in \mathsf{H}$ and consider corresponding sequences $\psi_n, \phi_n \in \mathscr{S}(\mathbb{R}^3)$ with $\psi_n \to \psi$, $\phi_n \to \phi$ in H as $n \to +\infty$. If $\delta_n := \psi - \psi_n$ and $\delta'_n := \psi' - \psi'_n$, we have

$$|(\psi|U_0(t)\psi')| \leq |(\delta_n|U_0(t)\delta'_n)| + |(\delta_n|U_0(t)\psi'_n)| + |(\psi_n|U_0(t)\delta'_n)| + |(\psi_n|U_0(t)\psi'_n)|\,.$$

Using Schwarz's inequality and the fact $U_0(t)$ is isometric we find

$$|(\psi|U_0(t)\psi')| \leq ||\delta_n||||\delta'_n|| + ||\delta_n||||\psi'_n|| + ||\psi_n||||\delta'_n|| + |(\psi_n|U_0(t)\psi'_n)|\,.$$

But the norm is obviously continuous in H, and $\delta_n \to 0$ and $\delta'_n \to 0$ as $n \to +\infty$. Hence for any given $\varepsilon > 0$, there is a large enough $n \in \mathbb{N}$ for which $||\delta_n||||\delta'_n||, ||\delta_n||||\psi'_n||, ||\psi_n||||\delta'_n|| < \varepsilon/4$. Therefore

$$|(\psi|U_0(t)\psi')| \leq 3\varepsilon/4 + |(\psi_n|U_0(t)\psi'_n)|\,.$$

Computing the inner product on $L^2(\mathbb{R}^3, dx)$ explicitly, and since $\psi_n, \psi'_n, U_0(t)\psi'_n \in \mathscr{S}(\mathbb{R}^3)$, we obtain

$$|(\psi_n|U_0(t)\psi'_n)| \leq ||U_0(t)\psi'_n||_\infty \int_{\mathbb{R}^3} |\psi_n(\mathbf{x})|dx\,.$$

By (13.20) there exists $T > 0$ for which the right-hand side above is bounded by $\varepsilon/4$ when $t > T$. Altogether, for any pair $\psi, \psi' \in \mathsf{H}$, if $\varepsilon > 0$ there is $T > 0$ such that $|(\psi|U_0(t)\psi')| \leq \varepsilon$ whenever $t > T$. □

■

13.1.6 The evolution operator in absence of time homogeneity and Dyson's series

We return to the notion of evolution operator to discuss a generalisation that has to do with Schrödinger's equation. An important remark, made in axiom **A6**, is that the evolution operator U_τ is actually *independent of the initial instant*. If we fix the state ρ at the initial time t, $U_\tau \rho U^*_\tau$ will be the state at time $t + \tau$. Had we fixed the same state ρ at initial time $t' \neq t$, the state at time $t' + \tau$ would have been $U_\tau \rho U^*_\tau$ again. So the system's laws of dynamics are unaffected in the time interval $[t, t']$. In other terms axiom **A6** presumes, for the system S in the frame $\mathscr{I}$, *homogeneity of time*. Classically, this situation corresponds to having the Hamiltonian not explicitly dependent on time in the coordinates of a certain frame. This is not the case in more general dynamical situations, like when S interacts with an evolving external world. If, on the contrary, S is isolated (though this is not the only possibility) and we describe it in an inertial system, then time is homogeneous, as in classical mechanics.

But if time is not homogeneous, time evolution is axiomatised as follows.

A6'. *Let the quantum system S be described in an inertial frame $\mathscr{I}$, with space of states H_S. There exists a familiy $\{U(t_2,t_1)\}_{t_2,t_1\in\mathbb{R}}$ of unitary operators on H_S, called* **evolution operators** *from t_1 to t_2, satisfying, for $t,t',t''\in\mathbb{R}$:*

(i) $U(t,t) = I$;
(ii) $U(t'',t')U(t',t) = U(t'',t)$;
(iii) $U(t',t) = U(t,t')^* = U(t',t)^{-1}$

and such that the function $\mathbb{R}^2 \ni (t,t') \mapsto U(t,t')$ is strongly continuous.
Furthermore, if ρ is the state at time t_0, the evolved state at time t_1 (which may precede t_0) is $U(t_1,t_0)\rho U(t_1,t_0)^*$.

The main difference with axiom **A6** is that now we cannot associate a self-adjoint generator to the family $\{U(t_2,t_1)\}_{t_2,t_1\in\mathbb{R}}$, and speaking of Hamiltonian of the system makes no longer sense, in general. We may still retain such a notion nonetheless (in the sense of a *time-dependent* Hamiltonian) by generalising Schrödinger's equation and defining the $U(t',t)$ as its solutions. Formally, the operator U_τ of **A6** satisfies Schrödinger's equation (with $\hbar = 1$):

$$\text{s-}\frac{d}{d\tau}U_\tau = -iHU_\tau\,.$$

For the generalised evolution operator $U(t',t)$, we can assume an analogous equation

$$\text{s-}\frac{d}{d\tau}U(\tau,t) = -iH(\tau)U(\tau,t)\,, \tag{13.43}$$

whenever to each instant τ is assigned an observable, called **Hamiltonian at time** τ, expressing the system's energy (in the given frame) at time τ. This energy is, in general, not a preserved quantity. In order to treat equation (13.43) rigorously a few delicate technical problems must be addressed concerning the distinct domains of the various $H(\tau)$; however, the equation remains extremely useful in a number of practical applications. The so-called *Dyson series*, pivotal in quantum electrodynamics and quantum field theory, is a formal solution to (13.43). To this end let us prove a result that illustrates the simplified situation where each Hamiltonian $H(\tau)$ is bounded and defined on the entire Hilbert space. In that case the collection of the $H(\tau)$ does determine via (13.43) a *continuous* family of evolution operators $U(t',t)$ given by Dyson's series.

Proposition 13.19. *Let H be a Hilbert space and $\mathbb{R}\ni t\mapsto H(t) = H(t)^* \in \mathfrak{B}(\mathsf{H})$ strongly continuous. Consider the* **Dyson series** *of the operators $U(t,s)$:*

$$U(t,s) := I + \sum_{n=1}^{\infty}(-i)^n \int_s^t dt_1 H(t_1)\int_s^{t_1} dt_2 H(t_2)\cdots\int_s^{t_{n-1}} dt_{n-1}H(t_n) \tag{13.44}$$

where iterated integrals are defined as in Proposition 9.27. Then the series converges uniformly. Moreover:

(a) *The $U(t,s)$ are unitary and satisfy (i), (ii), (iii) in* **A6'**.
(b) *The map $\mathbb{R} \ni (t,s) \mapsto U(t,s)$ is continuous in the uniform topology.*
(c) *The* **generalised Schrödinger equation** *holds:*

$$s\text{-}\frac{d}{dt}U(t,s) = -iH(t)U(t,s) \quad \textit{for every } t,s \in \mathbb{R}. \tag{13.45}$$

(d) *Expression (13.44) may be written:*

$$U(t,s) = \sum_{n=0}^{+\infty} \frac{(-i)^n}{n!} \int_s^t \int_s^t \cdots \int_s^t T\left[H(t_1)H(t_2)\cdots H(t_n)\right] dt_1 dt_2 \cdots dt_n\,. \tag{13.46}$$

Above,

$$T[H(t_1)H(t_2)\cdots H(t_n)] := H(\tau_n)H(\tau_{n-1})\cdots H(\tau_1)$$

is the **chronological reordering operator** *of the product: τ_n is the largest among $t_1,\ldots,t_n$, then comes $\tau_{n-1} \le \tau_n$ as second-largest and so on for every $t_1,\ldots,t_n$, up to the smallest value τ_1.*

Proof. First of all every term in Dyson's expansion

$$U_n(t,s) = (-i)^n \int_s^t dt_1 H(t_1) \int_s^{t_1} dt_2 H(t_2) \cdots \int_s^{t_{n-1}} dt_{n-1} H(t_n)$$

makes sense, since by Proposition 9.27(c) each integral on the right, starting from the right-most $(t_{n-1},s) \mapsto \int_s^{t_{n-1}} dt_{n-1}H(t_n)$, is an operator-valued map ranging in $\mathfrak{B}(\mathsf{H})$ and jointly strongly continuous in the integration limits (hence in the upper limit alone, too). The product (as pointwise operation) of two such maps is still strongly continuous and valued in $\mathfrak{B}(\mathsf{H})$, so it can be integrated. Using Proposition 9.27, where now the L^1 map is the characteristic function of the interval $[s,t_k]$, the nth term $U_n(t,s)$ in Dyson's series, $t,s \in [T,S]$, satisfies

$$||U_n(t,s)|| \le A_{a,b} := \frac{|b-a|^n}{n!}\left(\sup_{\tau\in[a,b]} ||H(\tau)||\right)^n, \quad (t,s) \in [a,b]^2\,. \tag{13.47}$$

As we observed in the proof of Proposition 9.27, since $\tau \mapsto H(\tau)$ is strongly continuous, $\sup_{\tau\in[a,b]} ||H(\tau)|| < +\infty$ by Banach–Steinhaus. Hence $0 \le A_{a,b} < +\infty$. Since the series of positive terms $A_{a,b}$ converges, the Dyson series converges in the uniform topology, uniformly in (s,t) on every compact set. Therefore if every Dyson term is uniformly continuous (proved next) then $(t,s) \mapsto U(t,s)$ is uniformly continuous. To show that the Dyson terms are uniformly continuous, we must resort to their recurrence relationship:

$$U_n(t,s) = -i\int_s^t H(\tau)U_{n-1}(\tau,s)d\tau\,. \tag{13.48}$$

It implies, working on the compact set $[a,b] \times [a,b]$,

$$||U_n(t,s) - U_n(t',s')|| \leq \left|\left|\int_{t'}^{t} H(\tau)U_{n-1}(\tau,s)d\tau\right|\right|$$
$$+\left|\left|\int_{s}^{t'} H(\tau)(U_{n-1}(\tau,s) - U_{n-1}(\tau,s'))d\tau\right|\right| + \left|\left|\int_{s}^{s'} H(\tau)U_{n-1}(\tau,s')d\tau\right|\right|,$$

so by Proposition 9.27(a):

$$||U_n(t,s) - U_n(t',s')|| \leq |t-t'| \sup_{(\tau,\sigma)\in[a,b]^2} ||H(\tau)|| ||U_{n-1}(\tau,\sigma)||$$
$$+(b-a) \sup_{\tau\in[a,b]} ||H(\tau)|| ||U_{n-1}(\tau,s) - U_{n-1}(\tau,s')||$$
$$+|s-s'| \sup_{(\tau,\sigma)\in[a,b]^2} ||H(\tau)|| ||U_{n-1}(\tau,\sigma)||\,.$$

Thus if $(t,s) \mapsto U_{n-1}(t,s)$ is uniformly continuous, so is $(t,s) \mapsto U_n(t,s)$; in particular

$$\sup_{\tau\in[a,b]} ||H(\tau)|| ||U_{n-1}(\tau,s) - U_{n-1}(\tau,s')|| \to 0 \quad \text{as } s \to s'$$

because the continuity of $(t,s) \mapsto U_{n-1}(t,s)$ on $[a,b]^2$ implies uniform continuity (besides, $\sup_{\tau\in[a,b]} ||H(\tau)|| < +\infty$ exists). The induction principle tells we can just prove that $U_1(t,s) = -i\int_s^t dt_1 H(t_1)$ is continuous. But this is true by (i) in Proposition 9.27(c). With this we proved (b) and part of (a). To finish (a) we will use (c), so let us prove that first. Applying Proposition 9.27(b, c) to the terms of the Dyson series computed on ψ, differentiating term by term and using (13.48), we arrive at

$$\frac{d}{dt}U(t,s)\psi = -iH(t)U(t,s)\psi\,, \quad \text{i.e.} \quad s\text{-}\frac{d}{dt}U(t,s) = -iH(t)U(t,s)\,, \qquad (13.49)$$

provided we can exchange sum and derivative. Using (13.47) together with $\sup_{t\in[a,b]} ||H(t)|| < +\infty$ tells the derivatives' series converges uniformly on compact sets in uniform topology, hence uniformly in the strong one. Hence the Dyson series can be differentiatied in t (strongly) term by term, which proves (13.49) and thus (c). Now we can finish claim (a). With a similar procedure, in particular employing Proposition 9.27(ii), we obtain $\frac{d}{ds}(\phi|U(t,s)\psi) = i(\phi|U(t,s)H(s)\psi)$. From this and (13.49) follows $\frac{d}{ds}(\phi|U(t,s)U(s,t)\psi) = i(\phi|U(t,s)(H(s) - H(s))U(s,t)\psi) = 0$, so in particular $(\phi|U(t,s)U(s,t)\psi) = (\phi|U(t,t)U(t,t)\psi)$. But $U(t,t) = I$, so $U(s,t) = U(t,s)^{-1}$. From (13.49) we have $\frac{d}{dt}||U(t,s)\psi||^2 = \frac{d}{dt}(U(t,s)\psi|U(t,s)\psi)$. The right-hand side is easy, and equals $(-iH(t)U(t,s)\psi|U(t,s)\psi) + (U(t,s)\psi| - iH(t)U(t,s)\psi) = 0$ by linearity on the right entry, antilinearity on the left, and because $H(t) = H(t)^*$. In other words $||U(t,s)\psi|| = ||U(s,s)\psi|| = ||\psi||$. Consequently every $U(t,s)$ is unitary, being isometric and onto. So we proved $U(t,s)^* = U(s,t) = U(t,s)^{-1}$. There remains to see (iii) of **A6'**. The operator $V(t,s) :=$

$U(t,s) - U(t,u)U(u,s)$ clearly satisfies $\frac{d}{dt}V(t,s)\psi = -iH(t)V(t,s)\psi$. Exactly as before $\frac{d}{dt}||V(t,s)\psi||^2 = \frac{d}{dt}(V(t,s)\psi|V(t,s)\psi) = 0$, by the inner product's linearity and by $H(t) = H(t)^*$. Hence $||V(t,s)\psi|| = ||V(s,s)\psi||$. But this is null, for $U(s,s) = I$ and $U(s,u)U(u,s) = I$. Eventually, then, $U(t,s)\psi = U(t,u)U(u,s)\psi$ for every $\psi \in \mathsf{H}$.

To show (13.46) it suffices, starting from the last relation, to express the iterated integrals of each series using suitable maps θ, and change names to variables, to get (13.44). For instance $T[H(t_1)H(t_2)] = \theta(t_1 - t_2)H(t_1)H(t_2) + \theta(t_2 - t_1)H(t_2)H(t_1)$, where $\theta(t) = 1$ for $t \geq 0$ and $\theta(t) = 0$ otherwise. Integrating the sum in $dt_1 dt_2$ over $[t,s]^2$, and swapping t_1, t_2 in the term with $\theta(t_2 - t_1)$, produces the second summand on the right in (13.44), apart from the constant $(-i)^2/2!$. □

Remark 13.20. **(1)** Dyson's series, written as in (13.46), resembles the series expansion of the time-ordered exponential. For that reason the series is often encountered, with $\hbar$ back in, in the integral form:

$$U(t,s) = T\left[e^{-\frac{i}{\hbar}\int_s^t H(\tau)d\tau}\right] . \tag{13.50}$$

If H does not depend on time, the right-hand side reduces precisely to $e^{-i\frac{(t-s)}{\hbar}H}$ as expected.

(2) We have already noted that the Dyson series is central in quantum field theory, and certainly in perturbation theory where the Hamiltonian decomposes as $H = H_0 + V$ and V is a correcting term to H_0 and the dynamics it generates. In such cases one proceeds by the so-called *Dirac's interaction picture* [Mes99, CCP82], in which the Dyson series plays a key part. In general concrete applications the Dyson series is used also when H is not bounded. For that reason the above theorem does not apply and the series should be understood in a weak sense of sorts [ReSi80]. ■

13.1.7 Antiunitary time reversal

Let us return to general matters in relation to the time-evolution axiom **A6**, i.e. *under time homogeneity*, and show two more important corollaries to the existence of a lower bound for the spectrum of the Hamiltonian H.

In the previous chapter we saw that if a system admits a symmetry (whether Kadison or Wigner is irrelevant to Theorem 12.12), the latter is a unitary or antiunitary transformation. If a system S with Hamiltonian H possesses the *time reversal* symmetry $\gamma_{\mathscr{T}}$ (cf. Example 12.17(2)), the unitary or antiunitary $\mathscr{T} : \mathsf{H}_S \to \mathsf{H}_S$ it determines (suppose the Hilbert space has one coherent sector) must satisfy

$$\gamma_{\mathscr{T}}\left(\gamma_t^{(H)}(\rho)\right) = \gamma_{-t}^{(H)}\left(\gamma_{\mathscr{T}}(\rho)\right)$$

(we set $\hbar = 1$ henceforth). Equivalently,

$$e^{-itH}\mathscr{T}\rho\mathscr{T}^{-1}e^{+itH} = \mathscr{T}e^{+itH}\rho e^{-itH}\mathscr{T}^{-1} \quad \text{for every } \rho \in \mathfrak{S}(\mathsf{H}_S). \tag{13.51}$$

Therefore time reversal, when present, is *not* a dynamical symmetry in the sense of Definition 13.4, owing to the sign flip of time in the dynamical flow. The following important result rephrases, partially, Proposition 13.2.

Theorem 13.21. *Consider a system S with Hamiltonian H (of lower-bounded spectrum) on the Hilbert space* H_S*. If the spectrum of H is unbounded above, every operator* $\mathscr{T} : \mathsf{H}_S \to \mathsf{H}_S$ *satisfying (13.51) is antiunitary and such that*

$$\mathscr{T}^{-1} H \mathscr{T} = H .$$

This applies in particular to the time-reversal symmetry, if existent.

Proof. If $V : \mathsf{H}_S \to \mathsf{H}_S$ is unitary (or antiunitary), then $V\psi(\psi|V^{-1}\cdot) = V\psi(\psi|V^*\cdot) = V\psi(V\psi|\cdot)$. Setting $U_t := e^{-itH}$ and taking the unitary $V := (\mathscr{T}U_{-t})^{-1}U_t\mathscr{T}$, for any pure state $\rho = \psi(\psi|\ )$ we have

$$(\mathscr{T}U_{-t})^{-1}U_t\mathscr{T}\psi\left((\mathscr{T}U_{-t})^{-1}U_t\mathscr{T}\psi\right|\ ) = \psi(\psi|\ ) .$$

Hence for some $\chi_t \in \mathbb{C}$ with $|\chi_t| = 1$:

$$(\mathscr{T}U_{-t})^{-1}U_t\mathscr{T}\psi = \chi_t\psi , \quad \psi \in H.$$

Replicating the argument of Theorem 12.10 shows χ_t does not depend on ψ. What is more, the map $\mathbb{R} \ni t \mapsto \chi_t$ is differentiable: take $\phi \in D(H)$, $\psi \in \mathscr{T}^{-1}D(H)$ with $(\phi|\psi) \neq 0$ ($D(H)$ is dense) and differentiate the identity $(\mathscr{T}U_{-t}\phi|\mathscr{T}^{-1}U_t\mathscr{T}\psi) = \chi_t(\phi|\psi)$, if $\mathscr{T}$ is unitary, or $\overline{(\mathscr{T}U_{-t}\phi|\mathscr{T}^{-1}U_t\mathscr{T}\psi)} = \chi_t(\phi|\psi)$ if $\mathscr{T}$ is antiunitary. Stone's theorem guarantees derivatives exist. Hence there is a differentiable map $\mathbb{R} \ni t \mapsto \chi_t$ such that $e^{-itH}\mathscr{T} = \mathscr{T}\chi_t e^{itH}$, so $\mathscr{T}^{-1}e^{-itH}\mathscr{T} = \chi_t e^{itH}$. Therefore

$$e^{\mp it\mathscr{T}^{-1}H\mathscr{T}} = \chi_t e^{itH}$$

with $-$ if $\mathscr{T}$ is unitary and $+$ if antiunitary (cf. Exercise 12.6 for the latter). Note $\mathscr{T}^{-1}H\mathscr{T}$ is self-adjoint, so the left-hand side is a strongly continuous unitary group parametrised by $t \in \mathbb{R}$. Applying Stone's theorem tells $D(\mathscr{T}^{-1}H\mathscr{T}) \subset D(H) = D(cI + H)$ and

$$\mp\mathscr{T}^{-1}H\mathscr{T}\restriction_{D(H)} = cI + H \quad \text{where } c := -i\frac{d\chi_t}{dt}|_{t=0} . \tag{13.52}$$

The constant c is real, for $\mp\mathscr{T}^{-1}H\mathscr{T} - H$ is symmetric on $D(H)$. As a matter of fact (13.52) is valid everywhere on the domain of the self-adjoint $\mp\mathscr{T}^{-1}H\mathscr{T}$, which has no self-adjoint extensions ($cI + H$) other than itself. Therefore

$$\mathscr{T}^{-1}H\mathscr{T} = \mp cI \mp H .$$

In particular (cf. Exercise 12.6 for the antiunitary case):

$$\sigma(H) = \sigma(\mathscr{T}^{-1}H\mathscr{T}) = \sigma(\mp cI \mp H) = \mp c \mp \sigma(H) .$$

Suppose $\sigma(H)$ is bounded below but not above: the last identity cannot hold if on the right there is a $-$ sign, whichever the constant c. Then $\mathscr{T}$ must be antiunitary, and $\inf\sigma(H) = \inf(c + \sigma(H)) = c + \inf\sigma(H)$. Hence $c = 0$, as $\inf\sigma(H)$ is finite by assumption ($\sigma(H) \neq \varnothing$ is bounded below). □

13.2 The time observable and Pauli's theorem. POVMs in brief

There is yet another consequence of the spectral lower bound of H, that addresses the problem of the existence of a quanum observable corresponding to the classical quantity of time, which satisfies canonical commutation relations with the Hamiltonian. The existence of such an operator could be suggested by Heisenberg's 'time-energy' uncertainty relationship mentioned in Chapter 6. In Chapter 11 we deduced Heisenberg's uncertainty principle for position and momentum as a theorem, following the CCR

$$[X,P] = i\hbar I\,.$$

We could expect a self-adjoint operator T corresponding to the *observable time* (*when* a phenomenon occurs, or its *duration* in a given quantum system); it should satisfy a similar commutation relation with the Hamiltonian, on some domain:

$$[T,H] = i\hbar I\,,$$

and therefore there should be an analogue time-energy uncertainty

$$(\Delta H)_\psi (\Delta T)_\psi \geq \hbar/2$$

exactly as for position-momentum. We saw in Chapter 11 that by interpreting in strong sense the position-momentum CCR, i.e. passing from operators to the exponential algebra, the exponential commutation relation determined the operators up to unitary transformations, by virtue of the Stone–von Neumann theorem. These alleged relations would read $e^{-i\frac{h}{\hbar}T}e^{-i\frac{t}{\hbar}H} = e^{i\frac{ht}{\hbar}}\,e^{-i\frac{t}{\hbar}H}e^{-i\frac{h}{\hbar}T}$. But in the case at stake that is not possible. There no way to define properly the operator time, and thus make sense to the time-energy relations: a no-go result that bears the name of *Pauli's theorem*. It is however possible to try to define the observable time, case by case, invoking the notion of *generalised observable*, which is useful in other contexts like the theory of quantum information.

13.2.1 Pauli's theorem

Putting together a series of results collected from previous chapters, we will prove our version of a result known as Pauli's theorem.

Theorem 13.22. *Consider a system S with Hamiltonian H (with lower-bounded spectrum) on the Hilbert space* H_S. *Suppose there exist a self-adjoint operator* $T : D(T) \to \mathsf{H}_S$ *and a subspace* $\mathscr{D} \subset D(H) \cap D(T)$ *in* H_S *on which* TH *and* HT *are well defined and the CCR (*$\hbar = 1$*)*

$$[T,H] = iI$$

holds. Then none of the following facts can occur.

(a) $\mathscr{D}$ *is dense and invariant under* T, H, *and the symmetric operator* $T^2 + H^2$ *is essentially self-adjoint on* $\mathscr{D}$.
(b) $\mathscr{D}$ *is dense, invariant under* T, H *and made of analytic vectors for both* T *and* H.

(c) *The exponential operators satisfy CCRs:*

$$e^{ihT}e^{itH} = e^{iht}\, e^{itH}e^{ihT}\ , \quad t,h \in \mathbb{R}.$$

Proof. If (a) were true, by Nelson's Theorem 12.62 $H\restriction_{\mathscr{D}}$ and $T\restriction_{\mathscr{D}}$ would be essentially self-sdjoint (making $\mathscr{D}$ a core for both self-adjoint H, T) and there would be a strongly continuous unitary representation of the unique simply connected Lie group whose Lie algebra is generated by I, H, T under the CCRs and the trivial $[T,I] = [H,I] = 0$. But that defines the Heisenberg group $\mathscr{H}(2)$, as seen in the previous chapter, and we would have proven (c). The same conclusion follows from assuming (b) because of Theorem 12.63. So let us suppose (c) holds. Going through the argument after Theorem 11.24, we could prove that the $W(t,h) := e^{iht/2}e^{itH}e^{ihT}$ satisfy Weyl's relations and the hypotheses of Mackey's Theorem 11.23. Then the Hilbert space H_S would split in an orthogonal sum $\mathsf{H}_S = \oplus_k \mathsf{H}_k$ of closed invariant spaces under e^{itH} and e^{ihT} for any t, h; and for any k there would be a unitary map $S_k : \mathsf{H}_k \to L^2(\mathbb{R}, dx)$, so $S_k e^{itH}\restriction_{\mathsf{H}_k} S_k^{-1} = e^{itX}$ in particular, with X denoting the standard position operator on $\mathbb{R}$. Applying Stone's theorem to $e^{itH}\mathsf{H}_s \subset \mathsf{H}_s$ we would obtain these consequences: first $H(\mathsf{H}_k \cap D(H)) \subset \mathsf{H}_k$, second $H\restriction_{\mathsf{H}_k \cap D(H)}$ is self-adjoint on H_k, and then $e^{itH}\restriction_{\mathsf{H}_k} = e^{itH\restriction_{\mathsf{H}_k \cap D(H)}}$. At this point the condition satisfied by S_k would read $e^{itH\restriction_{\mathsf{H}_k \cap D(H)}} = S_k^{-1}e^{itX}S_k$. Reapplying Stone's theorem would produce $H\restriction_{\mathsf{H}_k \cap D(H)} = S_k^{-1}XS_k$, hence

$$\sigma(H) \supset \sigma(H\restriction_{\mathsf{H}_k \cap D(H)}) = \sigma(S_k^{-1}XS_k) = \sigma(X) = \mathbb{R}\,.$$

(For the first inclusion it suffices to use the definition of spectrum.) But that is impossible because $\sigma(H)$ is bounded from below. □

13.2.2 Generalised observables as POVMs

The problem raised by Pauli's theorem about the definition of time is hard, and not yet completely construed. One attempt, that weakens the notions of observable and PVM, has found several other uses in QM, especially in *Quantum Information* [NiCh07].

Let us look into the proof of Proposition 7.52, which associates probability measures to observables seen as PVMs on $\mathbb{R}$: given a state $\rho \in \mathfrak{S}(\mathsf{H})$, we did *not* employ the characterisation of observables of Proposition 7.44 (property (a)); and concerning propriety (d), we only made use of *weak convergence* (implied by strong convergence). So we may rephrase Proposition 7.52 like this.

Proposition 13.23. *Let* H *be a Hilbert space and* $\{P(E)\}_{E\in\mathscr{B}(\mathbb{R})}$ *a collection of operators in* $\mathfrak{B}(\mathsf{H})$ *satisfying:*

(a)' $P(E) \geq 0$ *for every* $E \in \mathscr{B}(\mathbb{R})$.
(b)' $P(\mathbb{R}) = I$.
(c)' *For any countable set* $\{E_n\}_{n\in\mathbb{N}}$ *of pairwise-disjoint Borel sets in* $\mathbb{R}$,

$$w\text{-}\sum_{n=0}^{+\infty} P(E_n) = P\left(\cup_{n\in\mathbb{N}} E_n\right)\ .$$

If $\rho \in \mathfrak{S}(\mathsf{H})$, *the mapping* $\mu_\rho : E \mapsto tr(P(E)\rho)$ *is a probability measure on* $\mathbb{R}$.

The numbers $\mu_\rho(E)$ are the probabilities the experimental readings of the observable $\{P(E)\}_{E\in\mathscr{B}(\mathbb{R})}$ fall in the Borel set E. Sometimes it is convenient to adopt *generalised observables*, assuming they are given by maps $E\mapsto P(E)$ satisfying conditions (a)', (b)', (c)': these are weaker than the ones for PVMs, but still guarantee μ_ρ is a probability measure. In particular, the $P(E)$ are no longer orthogonal projectors, but mere bounded positive operators. All this leads to the following definition. We refer to [Ber66] for a broad mathematical treatise and [BGL95] for an extensive discussion on the applications to QM.

Definition 13.24. *Let* H *be a Hilbert space and* X *a topological space. A mapping* $A:\mathscr{B}(\mathsf{X})\to\mathfrak{B}(\mathsf{H})$ *is called* **positive operator-valued measure** *(POVM) on* X *if:*

(a)' $A(E)\geq 0$ *for any* $E\in\mathscr{B}(\mathsf{X})$.
(b)' $A(\mathsf{X})=I$.
(c)' *For any countable set* $\{E_n\}_{n\in\mathbb{N}}$ *of disjoint Borel subsets in* X*:*

$$w\text{-}\sum_{n=0}^{+\infty}A(E_n)=A(\cup_{n\in\mathbb{N}}E_n)\,.$$

A **generalised observable** *on* H *is a collection of operators* $\{A(E)\}_{E\in\mathscr{B}(\mathbb{R})}$ *such that* $\mathscr{B}(\mathbb{R})\ni E\mapsto A(E)$ *is a positive operator-valued measure.*

If A is a POVM on H, since $\mathfrak{B}(\mathsf{H})\ni A(E)\geq 0$ for every $E\in\mathscr{B}(\mathsf{X})$, we have $A(E)=A(E)^*$ by Proposition 3.54(f). Moreover, by Definition 13.24(c, d) $0\leq A(E)\leq I$, so $||A(E)||\leq 1$ from Proposition 3.54(a).

On a Hilbert space H the set $\mathfrak{E}(\mathsf{H})$ of $A\in\mathfrak{B}(\mathsf{H})$ with $0\leq A\leq I$ is called **space of effects**, and the **effects** are the operators A. The effects on H are the operators used to build every POVM on H, and their space is the analogue of $\mathfrak{P}(\mathsf{H})$ in defining observables via PVMs. $\mathfrak{E}(\mathsf{H})$ contains $\mathfrak{P}(\mathsf{H})$ and is partially ordered by the usual relation $\geq$: in contrast to $\mathfrak{P}(\mathsf{H})$, though, it is not a lattice. This prevents a generalised interpretation of orthogonal projectors as propositions on the system.

Extending axiom **A3** from post-measurement states to generalised observables is problematic. It is not possible to establish, in practice, in which state the system is after a measurement whose reading is $E'\in\mathscr{B}(\mathbb{R})$ if the observable is represented by a POVM $\{A(E)\}_{E\in\mathscr{B}(\mathbb{R})}$, and without further information. The extra data is assigned by decomposing each $A(E)=B(E)^*B(E)$ in the POVM, where the operators $B(E)\in\mathfrak{B}(\mathsf{H})$ are called **measuring operators**. If so, the post-measurement state is assumed to look like $(tr(A(E')\rho))^{-1}B(E')\rho B(E')^*$. For PVMs, clearly, $A(E)=B(E)$ are orthogonal projectors. See [BGL95] for details.

Here is an interesting application of generalised observables to the definition of time. Suppose time is defined as the observable associated to the lapse it takes a particle to hit a detector. By Pauli's Theorem 13.22 such observable is unlikely going to be defined via projectors if we impose that the observable is somehow "conjugated" to the Hamiltonian.

The attempts to define the *time observable* in terms of POVMs are very promising. Candidates for a *generalised time observable* T, e.g. the arrival time of a free particle, arise from a suitable POVM $T := \{A(E)\}_{E\in\mathscr{B}(\mathbb{R})}$ dependent on the system [Gia97, BF02]. Introducing measures $\mu^{(T)}_{\psi,\phi}(E) := (\psi|A(E)\phi)$, $E \subset \mathscr{B}(\mathbb{R})$, and setting $\mu^{(T)}_{\psi} := \mu^{(T)}_{\psi,\psi}$, we can define $\langle T\rangle_\psi$ and $(\Delta T)_\psi$ using the same definitions for PVMs. If T is built appropriately, on suitable domains, then $(\Delta T)_\psi(\Delta H)_\psi \geq \hbar/2$ and the analogues hold [Gia97, BF02], where H is the system's Hamiltonian. In analogy to PVMs, we may associate to the POVM T an operator, denoted by T, characterised by being the unique operator such that:

$$(\psi|T\phi) := \int_{\mathbb{R}} \lambda d\mu^{(T)}_{\psi,\phi}(\lambda)\,,$$

where $\psi \in \mathsf{H}$, $\phi \in D(T)$ dense and suitable. T turns out to be symmetric, but non self-adjoint. For a particle of mass $m > 0$ free to move along the real axis, the operator T of [Gia97] has the obvious form $T = \frac{m}{2}(XP^{-1} + P^{-1}X)$, on a suitable dense subspace of $L^2(\mathbb{R}, dx)$.

Remark 13.25. **(1)** Gleason's Theorem 7.24 has an important extension to generalised observables due to Busch [Bus03].

Theorem 13.26 (Busch). *Let* H *be a complex Hilbert space of finite dimension* ≥ 2 *or separable. For any map* $\mu : \mathfrak{E}(\mathsf{H}) \to [0,1]$ *such that* $\mu(I) = 1$ *and* $\mu\left(w\text{-}\sum_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mu(A_n)$ *for every sequence* $\{A_n\}_{n\in\mathbb{N}} \subset \mathfrak{E}(\mathsf{H})$ *satisfying* $w\text{-}\sum_{n=0}^{+\infty} A_n \leq I$*, there exists* $\rho \in \mathfrak{S}(\mathsf{H})$ *such that* $\mu(A) = tr(A\rho)$*,* $A \in \mathfrak{E}(\mathsf{H})$*.*

(2) An important theorem shows the tight relationship between PVMs and POVMs.

Theorem 13.27 (Neumark). *Let* X *be a topological space and* H *a Hilbert space. If* $A : \mathscr{B}(\mathsf{X}) \to \mathfrak{B}(\mathsf{H})$ *is a POVM, there exist a Hilbert space* H'*, an operator* $U : \mathsf{H} \to \mathsf{H}'$ *and a PVM* $P : \mathscr{B}(\mathsf{X}) \to \mathfrak{P}(\mathsf{H}')$ *such that* $A(E) = U^*P(E)U$ *for every* $E \in \mathscr{B}(\mathsf{X})$*.*

Condition (b)' in Definition 13.24 and the analogue for PVMs imply $U^*U = I_{\mathsf{H}}$, so U is an isometry (not surjective, otherwise A would be a PVM). Thus H is isomorphic to a (proper) closed subspace of H'. Yet $P(E)$ does not, in general, have a direct physical meaning, because H' is *not* the system's Hilbert space. ■

13.3 Dynamical symmetries and constants of motion

This section is devoted to extending to QM the outcome of the various versions of *Nöther's theorem*: in classical theories that theorem relates dynamical symmetries to constants of motion. In QM this relationship is as straightforward as it can get. To state the relative theorem we need to introduce the so-called *Heisenberg picture* of observables.

13.3.1 Heisenberg's picture and constants of motion

Take a quantum system S described in the inertial frame $\mathscr{I}$ with evolution operator $\mathbb{R} \ni \tau \mapsto e^{-i\tau H}$. Fix once for all the instant $t = 0$ for the initial conditions. Then consider the associated continuous projective representation of $\mathbb{R}$, $\mathbb{R} \ni t \mapsto \gamma_t^{(H)} := e^{-itH} \cdot e^{itH}$, and the *dual action* (cf. Chapter 12.1.6) on observables. If A is an observable (possibly an orthogonal projector representing an elementary property of S) we call

$$A_H(t) := \gamma_t^{(H)*}(A) = e^{itH} A e^{-itH}$$

the **Heisenberg picture** of A at time τ. By construction $\sigma(A_H(\tau)) = \sigma(A)$ and the observables's spectral measures satisfy $P^{(A_H(t))}(E) = \gamma_t^{(H)*}(P^{(A)}(E))$ for any $E \in \mathscr{B}(\mathbb{R})$.

In Heisenberg's picture, coherently with the symmetries' dual action of Chapter 12, quantum states do *not* evolve in time and the dynamics acts on observables. In particular, the expectation value of A on the state ρ_t, evolution till time t of the initial state ρ, can be computed either as $\langle A \rangle_{\rho_t}$ or equivalently as $\langle A_H(t) \rangle_\rho$, because

$$\langle A \rangle_{\rho_t} = tr\left(A U_t \rho U_t^{-1}\right) = tr\left(U_\tau^{-1} A U_t \rho\right) = \langle A_H(\tau) \rangle_\rho$$

if we put ourselves in the hypotheses of Proposition 11.8 (using the measure $\mu_\rho^{(A)}$ direcly shows that the result holds generally). And the same happens for the probability that the reading of A at time τ falls within the Borel set E, if ρ was the state at time 0:

$$tr(P_E^{(A)} \rho_t) = tr(P_E^{(A_H(t))} \rho) \, .$$

Remark 13.28. (1) To distinguish Heisenberg's picture from the ordinary picture in which states – not observables – evolve, the latter is often called **Schrödinger picture**, a convention we will adopt.
(2) It must be noted that an observable may depend on time in Schrödinger's picture as well. Better said, it is convenient to use a self-adjoint family $\{A_t\}_{t \in \mathbb{R}}$ parametrised by time t, and view it as a single observable denoted A_t. In such a case we say the observable **depends on time explicitly**. In Heisenberg's picture time dependency takes care of both (implicit and explicit) dependencies:

$$A_{Ht}(t) := \gamma_t^{(H)*}(A_t) = e^{itH} A_t e^{-itH} \, . \tag{13.53}$$

Now that we have seen the evolution of observables in Heisenberg's picture, we can introduce *constants of motion* by mimicking the classical definition. ■

Definition 13.29. *Let S be a quantum system described on the Hilbert space H_S associated to the inertial frame $\mathscr{I}$ with Hamiltonian H. An observable A is a* **constant of motion** *or a* **first integral** *if its Heisenberg picture is does not depend on time:*

$$A_H(t) = A_H(0) \quad \textit{for any } t \in \mathbb{R}. \tag{13.54}$$

An explicitly time-dependent observable A_t is called in the same way provided

$$A_{Ht}(t) = A_{H0}(0) \quad \textit{for any } t \in \mathbb{R}. \tag{13.55}$$

Remark 13.30. (1) An observable that does not depend on time explicitly is a constant of motion if and only if its Heisenberg and Schrödinger pictures coincide.

(2) The notions of Heisenberg's picture and constants of motion extend to situations where time is not homogeneous and with evolution operators different from $U(t_2,t_1)$. We will not worry about this.
(3) Identity (13.55) is oftentimes found in books written as

$$\frac{\partial A_{Ht}}{\partial t} + i[H, A_{Ht}(t)] = 0\,, \tag{13.56}$$

where the partial derivative refers to the explicit time variable only, i.e. the subscript $_{Ht}$. In practice if we do not care about domain issues, that equation is a trivial consequence of (13.55), and implies (13.55) if we also assume (13.53). The equivalence, in general false, is however troublesome to prove. At any rate, the concept of *constant of motion* is perfectly formalised, physically, by (13.55), with no need to differentiate in time and incur in spurious technical problems. ■

Notation 13.31. Lest we overburden notations for (explicitly) time-dependent observables, we will simply write $A_H(t)$ instead of $A_{Ht}(t)$ from now on, if no confusion arises. ■

We are ready to exhibit the relationship between constants of motion and dynamical symmetries. In classical physics one-parameter symmetry groups are known to correspond, in the various formulations of Nöther's theorem, to constants of motion. We wish to extend that to QM. Let us start with an easy case.

Proposition 13.32. *Let $\sigma(\cdot) := V^{(\sigma)} \cdot V^{(\sigma)-1}$ be a dynamical symmetry with $V^{(\sigma)}$ simultaneously unitary and self-adjoint. Then the observable $V^{(\sigma)}$ is a constant of motion.*

Proof. If U_t is the evolution operator, by Theorem 13.5(c) $U_t V^{(\sigma)} U_t^{-1} = V^{(\sigma)}$. □

It is not that infrequent that an interesting operator is together unitary and self-adjoint (and thus represents a symmetry and an observable). An example is the *parity inversion*, which we discussed in Examples 12.17. The situation is completely different from that of classical mechanics, where a system invariant under parity inversion (or any discrete symmetry) does not gain an associated constant of motion.

Let us deal with one-parameter groups of continuous symmetries, for which the link dynamical symmetries–constants of motion is forthright.

To begin with we consider a time-dependent observable $\{A_t\}_{t\in\mathbb{R}}$, in a certain system S with Hamiltonian H. If A_t is a constant of motion, then, by the previous definitions

$$e^{itH} A_t e^{-itH} = A_0\,.$$

If we exponentiate the self-adjoint operators in the equation gives

$$e^{-iae^{itH}A_t e^{-itH}} = e^{-iaA_0}\,,$$

an equation that known exponential properties transform into

$$e^{itH} e^{-iaA_t} e^{-itH} = e^{-iaA_0}\,,$$

i.e.

$$e^{-iaA_t}e^{-itH} = e^{-itH}e^{-iaA_0}\,, \quad a \in \mathbb{R}, t \in \mathbb{R}.$$

This equation's interpretation in terms of dynamical symmetries is quite relevant. It says that for any given $a \in \mathbb{R}$ the symmetries $\{\sigma_a^{(A_t)}\}_{t\in\mathbb{R}}$, with $\sigma_a^{(A_t)}(\cdot) := e^{-iaA_t} \cdot e^{iaA_t}$, form by Theorem 13.5 a time-dependent dynamical symmetry for S. If we restrict to $A_t = A$ time-independent, the same argument proves $\sigma_a^{(A)}(\cdot) := e^{-iaA} \cdot e^{iaA}$ is a dynamical symmetry for every $a \in \mathbb{R}$.

All this shows *constants of motion determine dynamical symmetries* but also *continuous* projective representations $\mathbb{R} \ni a \mapsto \sigma_a^{(A_t)}(\cdot)$ of $\mathbb{R}$, since $\mathbb{R} \ni a \mapsto e^{-iaA_t}$ is strongly continuous by Definition 12.31 (cf. Chapter 12.2.5).

Now we ask about *the converse*: given a family of time-dependent dynamical simmetries $\{\sigma_a^{(t)}\}_{t\in\mathbb{R}}$ where $\mathbb{R} \ni a \mapsto \sigma_a^{(t)}$ is a *continuous* projective representation of the group $\mathbb{R}$ for every $t \in \mathbb{R}$, is it possible to write each one of them as $\sigma_a^{(A_t)}(\cdot) := e^{-iaA_t} \cdot e^{iaA_t}$, so that the self-adjoint operators A_t give an (explicitly time-dependent) observable that is a constant of motion? According to Theorem 12.36 we can always find self-adjoint operators A_t such that $\sigma_a^{(A_t)}(\cdot) := e^{-iaA_t} \cdot e^{iaA_t}$ for every $a \in \mathbb{R}$. But these are determined up to a real constant $A_t \to A'_t := A_t - c(t)I$, so the point is whether one can fix the maps $c(t)$ so that

$$e^{itH}A'_t e^{-itH} = A'_0\,.$$

The answer of the next theorem, quantum version of Nöther's theorem, is yes.

Theorem 13.33 ("Quantum Nöther theorem"). *Let S be a quantum system, described on the Hilbert space H_S associated to the inertial frame $\mathscr{I}$, with Hamiltonian H and dynamical flow $\gamma^{(H)}$. If constants of motion and dynamical symmetries refer to $\gamma^{(H)}$, the following facts holds.*

(a) *If A is a constant of motion:*

$$\sigma_a^{(A)}(\cdot) := e^{-iaA} \cdot e^{iaA}$$

is a dynamical symmetry for every $a \in \mathbb{R}$, and $\mathbb{R} \ni a \mapsto \sigma_a^{(A)}(\cdot)$ is continuous.
(b) *Let $\{A_t\}_{t\in\mathbb{R}}$ be a time-dependent observable and a constant of motion. As $t \in \mathbb{R}$ varies,*

$$\sigma_a^{(A_t)}(\cdot) := e^{-iaA_t} \cdot e^{iaA_t}$$

is a time-dependent dynamical symmetry for every $a \in \mathbb{R}$, and $\mathbb{R} \ni a \mapsto \sigma_a^{(A_t)}(\cdot)$ is continuous $\forall t \in \mathbb{R}$.
(c) *Let σ_a be a dynamical symmetry and $\mathbb{R} \ni a \mapsto \sigma_a$ a continuous projective representation, $\forall a \in \mathbb{R}$. Then there exists a constant of motion A such that*

$$\sigma_a(\cdot) := e^{-iaA} \cdot e^{iaA}\,, \quad a \in \mathbb{R}.$$

(d) *Let $\{\sigma_a^{(t)}\}_{t\in\mathbb{R}}$ be a time-dependent dynamical symmetry $\forall a \in \mathbb{R}$, and $\mathbb{R} \ni a \mapsto \sigma_a^{(t)}$ a continuous projective representation $\forall t \in \mathbb{R}$. Then there exists a time-dependent observable $\{A_t\}_{t\in\mathbb{R}}$ that is a constant of motion plus*

$$\sigma_a^{(t)}(\cdot) := e^{-iaA_t} \cdot e^{iaA_t}\,, \quad a \in \mathbb{R}, t \in \mathbb{R}.$$

Proof. Claims (a), (b) were proved above, while (c) is evidently a subcase of (d) if we set $\sigma_a^{(t)} = \sigma_a$ and $A_t = A$ for any $t \in \mathbb{R}$. So there remains to prove (d). By Theorem 13.5, for any $t \in \mathbb{R}$ we can write $\sigma_a^{(t)}(\cdot) := e^{-iaA'_t} \cdot e^{iaA'_t}$, $a \in \mathbb{R}$, where the self-adjoint A'_t are given by the group $\mathbb{R} \ni \mapsto \sigma_a^{(t)}$ and can be redefined to $A'_t + c(t)I = A_t$ by adding constants $c(t)$. Let us imagine we have made a choice for those operators. By Theorem 13.5(a) for suitable unit complex numbers $\chi(t,a)$ we have

$$\chi(t,a) = e^{-iaA'_t} e^{-itH} e^{iaA'_0} e^{itH}, \tag{13.57}$$

whence $\chi(t,0) = 1$ for every $t \in \mathbb{R}$. Furthermore

$$\chi(t,a)(\psi|\phi) = \left(e^{itH} e^{iaA'_t} \psi \middle| e^{iaA'_0} e^{itH} \phi \right).$$

Choosing, for given $t \in \mathbb{R}$, $\psi \in (D(A'_t))$ and $\phi \in e^{itH}(D(A'_0))$ not orthogonal (the domains are dense because A_t is self-adjoint and e^{itH} unitary), and applying Stone's theorem on the right for the variable a, we obtain the derivative in a of the left-hand side, for every $a \in \mathbb{R}$. At the same time (13.57) imples, for given $t \in \mathbb{R}$:

$$\chi(t,a+a') = e^{-i(a+a')A'_t} e^{-itH} e^{i(a+a')A'_0} e^{itH}$$

$$= e^{-iaA'_t} \left(e^{-ia'A'_t} e^{-itH} e^{ia'A'_0} e^{itH} \right) e^{-itH} e^{iaA'_0} e^{itH}$$

$$= e^{-iaA'_t} \chi(t,a') e^{-itH} e^{iaA'_0} e^{itH} = \chi(t,a')\chi(t,a).$$

For $t \in \mathbb{R}$ given, the map $\mathbb{R} \ni a \mapsto \chi(t,a)$ is differentiable and satisfies $\chi(t,a+a') = \chi(t,a)\chi(t,a')$, so $\frac{\partial \chi(t,a)}{\partial a} = \frac{\partial \chi(t,a)}{\partial a}|_{a=0}\chi(t,a)$. Since $|\chi(t,a)| = 1$, $\chi(t,0) = 1$ for all $t \in \mathbb{R}$, the differential equation is solved by $\chi(t,a) = e^{ic(t)a}$ with $c(t) = -i\frac{\partial \chi(t,a)}{\partial a}|_{a=0} \in \mathbb{R}$. So we have

$$e^{ic(t)a} = e^{-iaA'_t} e^{-itH} e^{iaA'_0} e^{itH},$$

and hence

$$e^{-ia(A'_t + c(t)I)} e^{-itH} = e^{-itH} e^{-iaA'_0}.$$

By (13.57) $e^{ic(0)a} = 1$ for any $a \in \mathbb{R}$, so necessarily $c(0) = 0$. Then the above identity reads

$$e^{-ia(A'_t + c(t)I)} e^{-itH} = e^{-itH} e^{-ia(A'_0 + c(0)I)}.$$

As we said earlier we are free to modify the A'_t by adding constants, so with $A_t := A'_t + c(t)I$ we obtain

$$e^{-iaA_t} e^{-itH} = e^{-itH} e^{-iaA_0}.$$

This ends the proof. □

Remarks 13.34. Suppose the system's Hilbert space splits in coherent sectors under a superselection rule, and assume this rule corresponds to a certain observable Q being defined and taking a precise value in every sector, on every pure state. This is the case of the electric charge, for example. The self-adjoint operator representing Q is a constant of motion, since the evolution prevents the (pure) state to escape the sector where it initially lives. This observation unveils a deep relationship, between superselection rules and constants of motion, that proved extremely relevant in the algebraic formulation of quantum theories [Haa96] we will talk about in Chapter 14. ■

13.3.2 A short detour on Ehrenfest's theorem and related mathematical issues

Before we go on to examine the constants of motion of the Galilean group, we would like to spend some time on a topic related to the evolution of observables. In QM manuals there is a statement of acclaimed heuristic importance, especially in relating QM to its classical limit, known as *Ehrenfest theorem*. The heart of Ehrenfest's theorem is, formally, quite straightforward. Take a quantum system S described on the Hilbert space H_S and an observable or self-adjoint operator A (for simplicity time-independent). Fix a pure state/unit vector ψ and consider its evolution under the operator e^{-itH}. In formal terms, overlooking domains,

$$\frac{d}{dt}\langle A\rangle_{\psi_t} = \frac{d}{dt}\left(e^{-itH}\psi\middle|Ae^{-itH}\psi\right) = i\,(H\psi_t|A\psi_t) - i\,(\psi_t|AH\psi_t)$$

for $\psi_t := e^{-itH}\psi$. This implies the general **Ehrenfest relation**:

$$\frac{d}{dt}\langle A\rangle_{\psi_t} = \langle i[H,A]\rangle_{\psi_t}\,. \tag{13.58}$$

Although to obtain (13.58) we ignored important mathematical details, it is easy to prove (exercise) that the relation is implied by the following three conditions: (i) $A \in \mathfrak{B}(H)$; (ii) $\psi_\tau \in D(H)$ around t – equivalently $\psi \in D(H)$, since $D(H)$ is evolution-invariant; (iii) $\psi_\tau \in D(HA)$ around t. It is far from easy to make assumptions of some help in physical applications that only concern H, A and ψ and valid on a neighbourhood of some t. We can nevertheless weaken (i), (ii), (iii): beside $A \in \mathfrak{B}(H)$, assume only $\psi \in D(H)$, and interpret $\langle i[H,A]\rangle_{\psi_t}$ in (13.58) as a *quadratic form*:

$$\langle i[H,A]\rangle_{\psi_t} := i(H\psi_t|A\psi_t) - i(A\psi_t|H\psi_t)\,.$$

This yields a weaker version of Ehrenfest's theorem:

$$\frac{d}{dt}\langle A\rangle_{\psi_t} = i(H\psi_t|A\psi_t) - i(A\psi_t|H\psi_t)\,. \tag{13.59}$$

Even in this reading Ehrenfest's statement is still too abstract, because practically every observable A of interest in QM is not a bounded operator. In fact, the importance of Ehrenfest's theorem becomes evident precisely when applied to the unbounded operators position and momentum.

Consider, to that end, a system made by a spin-zero particle of mass m, subject to a potential V, in an inertial frame. The Hamiltonian is a self-adjoint extension of the differential operator $H_0 := -\frac{\hbar^2}{2m}\Delta + V$. Suppose we work with $\tau \mapsto \psi_\tau$, which around t belongs to some subdomain of $D(X_i\overline{H_0}) \cap D(\overline{H_0}X_i)$ on which the Hamiltonian $\overline{H_0}$ is differentiable. Then (reintroducing $\hbar$ everywhere):

$$[\overline{H_0}, X_i]\psi = -\frac{\hbar}{2m}\sum_{j=1}^{3}\left[\frac{\partial^2}{\partial x_j^2}, x_i\right]\psi = -\frac{\hbar}{m}\frac{\partial \psi}{\partial x_i},$$

whence (13.58) gives

$$m\frac{d}{dt}\langle X_i\rangle_{\psi_t} = \langle P_i\rangle_{\psi_t}. \tag{13.60}$$

Similarly, working around t with $\tau \mapsto \psi_\tau$ in some domain inside $D(P_i\overline{H_0}) \cap D(\overline{H_0}P_i)$ where $\overline{H_0}$ is differentiable, we obtain

$$[\overline{H_0}, P_i]\psi = -i\left[-V, \frac{\partial}{\partial x_i}\right]\psi = -i\frac{\partial V}{\partial x_i}\psi,$$

so from (13.58):

$$\frac{d}{dt}\langle P_i\rangle_{\psi_t} = -\left\langle \frac{\partial V}{\partial x_i}\right\rangle_{\psi_t}. \tag{13.61}$$

The classical statement of Ehrenfest's theorem consists of the pair (13.60)–(13.61), from which the mean values of position and momentum have a classical-like behaviour. Precisely, assume the gradient of V does not vary much on the spatial reach of the wavefunction $\psi_t(\mathbf{x})$. Then we can estimate the right-hand side of (13.61) by

$$\left\langle \frac{\partial V}{\partial x_i}\right\rangle_{\psi_t} \simeq \int_{\mathbb{R}^3} \overline{\psi_t(\mathbf{x})}\, \frac{\partial V}{\partial x_i}\bigg|_{\langle \mathbf{X}\rangle_{\psi_t}} \psi_t(\mathbf{x})\,dx = \left(\int_{\mathbb{R}^3} \overline{\psi_t(\mathbf{x})}\psi_t(\mathbf{x})dx\right) \frac{\partial V}{\partial x_i}\bigg|_{\langle \mathbf{X}\rangle_{\psi_t}} = \frac{\partial V}{\partial x_i}\bigg|_{\langle \mathbf{X}\rangle_{\psi_t}}.$$

Substituting in (13.61) we get the classical equation:

$$\frac{d}{dt}\langle P_i\rangle_{\psi_t} \simeq -\frac{\partial V}{\partial x_i}\bigg|_{\langle \mathbf{X}\rangle_{\psi_t}}. \tag{13.62}$$

The punchline is that *under Ehrenfest's equations (13.60)–(13.61), the more wave packets cluster around their mean value – under a potential whose force varies slowly on the packet's range – the better the momentum and position mean values obey the evolution laws of classical mechanics.*

Alas, the entire discussion is rather academic, because establishing *physically-sound* mathematical conditions on H_0, for the argument leading to (13.60)–(13.61) to be fully justified, is a largely unsolved problem.

Remark 13.35. (1) Recently, conditions on H and A have been found that realise (13.59) when A is *neither bounded* nor self-adjoint, including the case where A is the position or the momentum. The result we are talking about is the following [FK09].

Theorem 13.36. *On the Hilbert space* H *let* $H : D(H) \to \mathsf{H}$, $A : D(A) \to \mathsf{H}$ *be densely defined and such that:*

(H1) *H is self-adjoint and A Hermitian (hence symmetric);*
(H2) *$D(A) \cap D(H)$ is invariant under $\mathbb{R} \ni t \mapsto e^{-itH}$, for every $t \in \mathbb{R}$;*
(H3) *if $\psi \in D(A) \cap D(H)$ then $\sup_I ||Ae^{-itH}\psi|| < +\infty$ for any bounded interval $I \subset \mathbb{R}$.*

Let $\psi_t := e^{-itH}\psi$. Then for any $\psi \in D(A) \cap D(H)$ the map $t \mapsto \langle A \rangle_{\psi_t}$ is C^1 and

$$\frac{d}{dt}\langle A \rangle_{\psi_t} = i(H\psi_t|A\psi_t) - i(A\psi_t|H\psi_t)\,.$$

As earlier claimed, the above hypotheses subsume the case where A is the position or the momentum on $\mathsf{H} = L^2(\mathbb{R}^n, dx)$, even though proving it is highly non-trivial (cit., Corollary 1.2). For it to happen it is enough that H is the only self-adjoint extension of $H_0 = -\Delta + V$ on $\mathscr{D}(\mathbb{R}^n)$ with V real and $(-\Delta)$-bounded with relative bound $a < 1$, in the sense of Definition 10.41.
(2) From the point of view of physics it is impossible to build an experimental device capable of measuring all possible values of an observable described by an unbounded self-adjoint operator. For the position observable, for instance, it would mean filling the universe with detectors! So we expect any observable represented by the *unbounded* self-adjoint operator A to be – physically speaking – indistinguishable from the observable of the self-adjoint operator $A_N := \int_{\sigma(A)\cap[-N,N]} \lambda dP^{(A)}(\lambda) \in \mathfrak{B}(H)$, with $N > 0$ large but finite. The general form of Ehrenfest's theorem (13.58) applies to such class of observables, if we assume (ii) and (iii), or only $\psi \in D(H)$ to have (13.59). In this case, though, it is not easy to use the formal commutation of position and momentum with a Hamiltonian like $-\frac{\hbar^2}{2m}\Delta + V$, which would bring to (13.60), (13.61). ■

13.3.3 Constants of motion associated to symmetry Lie groups and the case of the Galilean group

Consider a quantum system S with Hilbert space H_S, Hamiltonian H and inertial frame $\mathscr{I}$. Suppose there is a Lie group G with a strongly continous unitary representation $\mathsf{G} \ni g \mapsto U_g$ on H_S, and assume the evolution operator $\mathbb{R} \ni t \mapsto e^{-itH}$ coincides with the representation of a one-parameter subgroup of G (clearly G is a symmetry group for S, since the representation U induces a projective representation of the same group). What we want to prove is that every $T \in \mathsf{T}_e\mathsf{G}$ determines a dynamical symmetry and a constant of motion (explicitly time-dependent, in general). In fact,

Theorem 13.37. *Let S be a quantum system on the Hilbert space H_S, with Hamiltonian H (in some inertial frame). Let $\mathsf{G} \ni g \mapsto U_g$ be a strongly continuous unitary representation on H_S of the n-dimensional Lie group G, and suppose the evolution operator $\mathbb{R} \ni t \mapsto e^{-itH}$ coincides with the representation of a one-parameter subgroup generated by some $-\mathbf{h} \in \mathsf{T}_e\mathsf{G}$:*

$$e^{-itH} = U_{\exp(t\mathbf{h})}\,, \quad t \in \mathbb{R}.$$

(a) *To each* $T \in \mathsf{T}_e\mathsf{G}$ *there correspond a constant of motion* $\{\widehat{T}_t\}_{t\in\mathbb{R}}$, *in general time-dependent, and a relative dynamical symmetry.*
(b) *If* $[\mathbf{h}, T] = 0$ *the constant of motion* $\{\widehat{T}_t\}_{t\in\mathbb{R}}$ *is time-independent.*

Proof. (a) Consider the map from $\mathbb{R}$ to G:

$$\mathbb{R} \ni a \mapsto \exp(t\mathbf{h})\exp(aT)\exp(-t\mathbf{h}).$$

It is certainly a one-parameter subgroup for any given $T \in \mathsf{T}_e\mathsf{G}$ and every $t \in \mathbb{R}$. So if $T_1, \dots T_n$ is a basis of $\mathsf{T}_e\mathsf{G}$, for suitable real functions $c_j = c_j(t)$ we can write

$$\exp(t\mathbf{h})\exp(aT)\exp(-t\mathbf{h}) = \exp(a\sum_{j=1}^{n} c_j(t)T_j).$$

Apply U and pass to the Lie algebra representation $\mathsf{T}_e\mathsf{G} \ni T \mapsto A_U[T] := A_U(T)\restriction_{\mathscr{D}_G}$, where the Gårding space $\mathscr{D}_G$ is invariant and a core for the self-adjoint operators $A_U(T)$ (Chapter 12, in particular Theorem 12.60). Then

$$e^{-itH}e^{-ia\overline{A_U[T]}}e^{itH} = e^{-ia\overline{\sum_{j=1}^{n} c_j(t)A_U[T_j]}}. \tag{13.63}$$

Define self-adjoint operators parametrised by time

$$\widehat{T}_t := \overline{\sum_{j=1}^{n} c_j(t)A_U[T_j]}.$$

Then (13.63) shows $\widehat{T}_t$ is a constant of motion that depends explicitly on time, for in fact (13.63) implies:

$$e^{itH}\widehat{T}_t e^{-itH} = \overline{A_U[T]} = \widehat{T}_0, \quad t \in \mathbb{R}.$$

Again (13.63) shows that the family of symmetries $\sigma_a^{(t)} := e^{-ia\widehat{T}_t} \cdot e^{-ia\widehat{T}_t}$, for any $a \in \mathbb{R}$, is a time-dependent dynamical symmetry. In fact (13.63) forces

$$e^{-ia\widehat{T}_t}e^{-itH} = e^{-itH}e^{-ia\widehat{T}_0}, \quad t \in \mathbb{R},$$

and then Theorem 13.5 proves the claim.
(b) Assuming $[T, \mathbf{h}] = 0$, and using the Baker–Campbell–Hausdorff formula (12.69), (12.70), (12.71), we obtain

$$\exp(\tau\mathbf{h})\exp(aT) = \exp(aT)\exp(\tau\mathbf{h}) \tag{13.64}$$

so long as $|a|, |\tau| < \varepsilon$ with $\varepsilon > 0$ small enough. Those formulas actually hold for any value of $a, \tau \in \mathbb{R}$. To see that it suffices to observe, irrespective of a and τ, that we can write $a = \sum_{r=1}^{N} a_r$ and $\tau = \sum_{r=1}^{N} \tau_r$ so that $|a_r|, |\tau_r| < \varepsilon$ for any r. For example,

$$\begin{aligned}
&\exp(\tau \mathbf{h}) \exp(aT) \\
&\quad = \exp(\tau_N \mathbf{h}) \cdots \exp(\tau_2 \mathbf{h}) \exp(\tau_1 \mathbf{h}) \exp(a_1 T) \exp(a_2 T) \cdots \exp(a_N T) \\
&\quad = \exp(\tau_N \mathbf{h}) \cdots \exp(\tau_2 \mathbf{h}) \exp(a_1 T) \exp(\tau_1 \mathbf{h}) \exp(a_2 T) \cdots \exp(a_N T) \\
&\quad \cdots \\
&\quad = \exp(a_1 T) \exp(\tau_N \mathbf{h}) \cdots \exp(\tau_2 \mathbf{h}) \exp(a_2 T) \cdots \exp(a_N T) \exp(\tau_1 \mathbf{h}) \\
&\quad \cdots \\
&\quad = \exp(a_1 T) \exp(a_2 T) \cdots \exp(a_N T) \exp(\tau_N \mathbf{h}) \cdots \exp(\tau_2 \mathbf{h}) \exp(\tau_1) \\
&\quad = \exp(aT) \exp(\tau \mathbf{h}) .
\end{aligned}$$

Consequently, using U we get

$$e^{-itH} e^{-ia\overline{A_U[T]}} e^{itH} = e^{-ia\overline{A_U[T]}} ,$$

whence the claim. □

To exemplify the general result found above, we revert to the Galilean group and its projective unitary representations seen at the end of the previous chapter. We will show there are 10 first integrals for a system having the restricted Galilean group $S\mathscr{G}$ as symmetry group (described by a unitary representation of a central extension of the universal covering $\widetilde{S\mathscr{G}}$). We consider in particular the spin-zero particle of mass m, and refer to the unitary representation of the central extension $\widehat{\widetilde{S\mathscr{G}}}_m$ of Chapter 12. The Lie algebra is the extension of the Lie algebra of $\widetilde{S\mathscr{G}}$, that has 10 generators $-\mathbf{h}\,, \mathbf{p}_i\,, \mathbf{j}_i\,, \mathbf{k}_i$ i=1,2,3, such that:

(i) $-\mathbf{h}$ generates the subgroup $\mathbb{R} \ni c \mapsto (c, \mathbf{0}, \mathbf{0}, I)$ of **time displacements**;
(ii) the three $\mathbf{p}_i$ generate the Abelian subgroup $\mathbb{R}^3 \ni \mathbf{c} \mapsto (0, \mathbf{c}, \mathbf{0}, I)$ of **space translations**;
(iii) the three $\mathbf{j}_i$ generate the subgroup $SO(3) \ni R \mapsto (0, \mathbf{0}, \mathbf{0}, R)$ of **space rotations**;
(iv) the three $\mathbf{k}_i$ generate the Abelian subgroup $\mathbb{R}^3 \ni \mathbf{v} \mapsto (0, \mathbf{0}, \mathbf{v}, I)$ of **pure Galilean transformations**.

These elements obey the commutation relations (12.106). To pass from the Lie algebra of $\widetilde{S\mathscr{G}}$ to that of $\widehat{\widetilde{S\mathscr{G}}}_m$ we add a generator commuting with the above ones, plus central charges for the commutation relations between $\mathbf{k}_i$ and $\mathbf{p}_j$ equal to the mass m (cf. (12.115) and ensuing discussion). The strongly continuous unitary representation of our concern is the one of $\widehat{\widetilde{S\mathscr{G}}}_m$:

$$\widehat{\widetilde{S\mathscr{G}}}_m \ni (\chi, g) \mapsto \chi \widetilde{Z^{(m)}}_g ,$$

induced by unitary operators $\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)}$ (12.107):

$$\left(\widetilde{Z^{(m)}}_{(c,\mathbf{c},\mathbf{v},U)} \widetilde{\psi}\right)(\mathbf{k}) := e^{i(c\mathbf{v}-\mathbf{c})\cdot(\mathbf{k}-m\mathbf{v})} e^{i\frac{c}{2m}(\mathbf{k}-m\mathbf{v})^2} \widetilde{\psi}\left(R(U)^{-1}(\mathbf{k}-m\mathbf{v})\right) .$$

Notice the Lie group $\widehat{\widetilde{\mathscr{SG}}}_m$ contains the subgroup spanned by $\mathbf{h}$, corresponding to the evolution operator on the system's Hilbert space H_S. Among the commutation relations (12.106) defining the Lie algebra of $\widetilde{\mathscr{SG}}$ (and valid on the central extension $\widehat{\widetilde{\mathscr{SG}}}_m$), we are interested in the ones directly involving $\mathbf{h}$:

$$[\mathbf{p}_i, \mathbf{h}] = \mathbf{0}\,, \quad [\mathbf{j}_i, \mathbf{h}] = \mathbf{0}\,, \quad [\mathbf{k}_i, \mathbf{h}] = -\mathbf{p}_i \quad i = 1,2,3. \tag{13.65}$$

Adapting the proof of Theorem 13.37 to the representation $\widehat{\widetilde{\mathscr{SG}}}_m \ni (\chi, g) \mapsto \chi Z^{(m)}_g$, the first two brackets give

$$e^{-i\tau H} e^{-iaP_i} = e^{-iaP_i} e^{-i\tau H} \tag{13.66}$$

and

$$e^{-i\tau H} e^{-iaL_i} = e^{-iaL_i} e^{-i\tau H}\,. \tag{13.67}$$

Using Theorem 13.5 and Definition 13.29, these tell, in agreement with Theorem 13.37:

(a) The three momentum components and the three orbital angular momentum components are constants of motion (time-independent).

(b) The symmetries generated by these integrals of motion, i.e. the *translations along the axes* and the *rotations about the axes* are (time-independent) dynamical symmetries (see Examples 12.37 and (12.99), respectively, for the explicit action on wavefunctions).

Let us tackle the third bracket in (13.65). A direct use of the Baker–Campbell–Hausdorff formula is not trivial, even if technically possible with a bit of work, also in the general case. To understand what this third identity corresponds to concerning the associated one-parameter subgroups, let us study the matter in the Galilean group. The subgroup generated by $-\mathbf{h}$ is the time displacement:

$$\exp(\tau\mathbf{h}) = (-\tau, \mathbf{0}, \mathbf{0}, I) \quad \tau \in \mathbb{R}\,.$$

The subgroup generated by $\mathbf{k}_j$ is a pure Galilean transformation along the jth axis with unit vector $\mathbf{e}_j$:

$$\exp(a\mathbf{k}_j) = (0, \mathbf{0}, a\mathbf{e}_j, I) \quad a \in \mathbb{R}\,.$$

Immediately, then, the group law (12.101) gives

$$\exp(\tau\mathbf{h})\exp(a\mathbf{k}_j)\exp(-\tau\mathbf{h}) = \exp(a(\tau\mathbf{p}_j + \mathbf{k}_j))\,.$$

Applying the unitary representation these become

$$e^{-i\tau H} e^{-aK_j} e^{i\tau H} = e^{-ia\left(\overline{\tau P_j\restriction_{\mathscr{D}_G} + K_j\restriction_{\mathscr{D}_G}}\right)}\,.$$

Therefore, if we define self-adjoint operators

$$K_{jt} := \overline{tP_j\restriction_{\mathscr{D}_G} + K_j\restriction_{\mathscr{D}_G}} \quad j = 1,2,3,$$

each observable is a constant of motion explicitly dependent on time, and each defines a dynamical symmetry for every $a \in \mathbb{R}$:

$$e^{-iaK_{jt}} e^{-itH} = e^{-itH} e^{-iaK_{j0}} .$$

The dynamical symmetry $e^{-iaK_{jt}}$ thus defines *pure Galilean transformation along* $\mathbf{e}_j$ *at time* t.

Remark 13.38. **(1)** It can be interesting to question about the meaning of the conservation law of K_{jt}, which is not at all obvious. We remind that the boost is defined (see (12.114)) as $K_j = -mX_j$. Choosing $\psi \in \mathscr{D}_G$ and letting it evolve under the evolution operator, $\psi_t := e^{-itH}\psi$, the conservation law for K_{jt} implies:

$$t(\psi_t|P_j\psi_t) - m(\psi_t|X_j\psi_t) = cost,$$

i.e.

$$\langle P_j\rangle_{\psi_t} = m\frac{d}{dt}\langle X_j\rangle_{\psi_t} . \tag{13.68}$$

Hence the mean momentum of the particle is, in some sense, the product of the mass times the velocity, the latter indicating the mean position of the particle. The result is *a priori* not evident, since in QM the momentum is *not* the product of mass and velocity.
(2) Suppose we work with a multi-particle system admitting the Galilean group as symmetry group described by a unitary representation of a central extension associated to the total mass M (see Chapter 12). Identity (13.68) is at present proved in the same way, and hence holds true. But now P_j is the component along $\mathbf{e}_j$ of the *total* momentum, and X_j the $\mathbf{e}_j$-component of the *position vector of the centre of mass*. A similar relationship holds for systems invariant under the Poincaré group, and follows from invariance by pure Lorentz transformations. The term corresponding to the total mass accounts for the energy contributions of the single components (like the kinetic energies of the isolated points making the system), in conformity to equation $M = E/c^2$. ∎

This accounts for 9 first integrals, but we said there are 10 in total.

The attentive reader will have noticed there is still a dynamical symmetry around, and a corresponding conservation law: energy! Namely, the obvious commutation relation $[\mathbf{h}, \mathbf{h}] = \mathbf{0}$ holds on the Lie algebra, or $[H, H] = 0$ at the level of self-adjoint generators, or $[e^{-i\tau H}, e^{-i\tau' H}] = 0$ for the exponentials. The last identity says, in agreement with Theorem 13.37, that applying Theorem 13.5 and Definition 13.6:

- the Hamiltonian is a constant of motion;
- the symmetry generated by $-H$ (the time displacement) is a dynamical symmetry.

The result is completely general and does not depend on having the Galilean group as symmetry group; it suffices that the Hamiltonian exists.

13.4 Compound systems and their properties

We met in Chapter 12 systems composed by subsystems and saw that the overall Hilbert space is a tensor product of Hilbert spaces relative to the subsystems. But this is actually an axiom of the theory. Compound systems bear a host of fascinating non-classical features, which we will review in this section.

13.4.1 Axiom A7*: compound systems*

We are ready to state the seventh axiom of QM, the one about compound quantum systems. For the mathematical contents we refer to the definitions and results of Chapter 10.2.1.

A7. *When a quantum system consists of a finite number N of subsystems, each described on a Hilbert space* H_i*, $i = 1,2,\ldots,N$, the comprehensive system is described on the Hilbert space* $\bigotimes_{i=1}^N \mathsf{H}_i$*. Any observable $A_i : D(A_i) \to \mathsf{H}_i$ on the ith subsystem (including elementary observables defined by orthogonal projectors) is identified in the larger system with the observable* $\overline{I \otimes \cdots \otimes I \otimes A_i \otimes I \otimes \cdots \otimes I}$*.*

Two are the types of compound systems we have already met: those made of *elementary particles with internal structure*, and *multi-particle* systems (elementary particles with or without internal structure). In the first case the Hilbert space is $L^2(\mathbb{R}^3,dx) \otimes \mathsf{H}_0$, where H_0 is *finite-dimensional* and describes the particle's internal degrees of freedom: spin and charges of various sort (cf. Chapter 11). By elementary particle with internal structure we mean that the internal space is finite-dimensional. The literature, referring to systems of elementary particles with space H_0, calls $L^2(\mathbb{R}^3,dx)$ the **orbital space** or **space of orbital degrees of freedom**, and H_0 the **internal space** or **space of internal degrees of freedom**. In case the space of internal freedom degrees describes a (certain type of) charge, we should also keep possible superselection rules into account.

We would like to make a few remarks on the Hamiltonian operator of multi-particle systems, when the single Hilbert spaces are $L^2(\mathbb{R}^3,dx)$ with a fixed inertial frame, and identifying $\mathbb{R}^3$ with the rest space via orthonormal Cartesian coordinates. The Hilbert space of a system on N particles with masses m_1, ..., m_N is the N-fold tensor product of $L^2(\mathbb{R}^3,dx)$. From Example 10.27(1) this product is naturally isomorphic to $L^2(\mathbb{R}^{3N},dx)$. Indicate by $(\mathbf{x}_1,\ldots,\mathbf{x}_N)$ the generic point in $\mathbb{R}^{3N}$, where $\mathbf{x}_k = ((x_k)_1,(x_k)_2,(x_k)_3)$ is the triple of orthonormal Cartesian coordinates of the kth factor of $\mathbb{R}^{3N} = \mathbb{R}^3 \times \cdots \times \mathbb{R}^3$. The natural isomorphism turns (prove it as exercise) the position operator of the kth particle into the multiplication by the corresponding $\mathbf{x}_k = ((x_k)_1,(x_k)_2,(x_k)_3)$, and the momentum into the unique self-adjoint extension of the $\mathbf{x}_k$-derivatives (times $-i\hbar$), for instance on $\mathscr{D}(\mathbb{R}^{3N})$. The Hamiltonians of each particle, assumed free, coincide with the self-adjoint extension, say on $\mathscr{D}(\mathbb{R}^{3N})$, of the corresponding Laplacian $-\Delta_k = \sum_{i=1}^3 \frac{\partial^2}{\partial (x_k)_i^2}$ times $-\hbar^2/(2m_k)$. Relying on Chapter 11.3.8, if the particles undergo interactions described classically by a potential $V = V(\mathbf{x}_1,\ldots,\mathbf{x}_N)$, the Hamiltonian is expected to be some self-adjoint extension of

$$H_0 := \sum_{k=1}^{N} -\frac{\hbar^2}{2m_k}\Delta_k + V(\mathbf{x}_1,\ldots,\mathbf{x}_N)\,.$$

For instance, particles with charges e_k interacting with one another under Coulomb forces and with external charges Q_k are expected to have a self-adjoint extension of

$$H_0 := \sum_{k=1}^{N} -\frac{\hbar^2}{2m_k}\Delta_k + \sum_{k=1}^{N}\frac{Q_k e_k}{|\mathbf{x}_k|} + \sum_{i<j}^{N}\frac{e_i e_j}{|\mathbf{x}_i - \mathbf{x}_j|}$$

as Hamiltonian. As we explained in Chapter 10.4 and Examples 10.51, important results mainly due to Kato imply, under natural assumptions on V, that not only H_0 is essentially self-adjoint on standard domains like $\mathscr{D}(\mathbb{R}^{3N})$ or $\mathscr{S}(\mathbb{R}^{3N})$, but the only self-adjoint extension is *bounded below*, thus making *the system energetically stable*. This happens in particular for the operator with the Coulomb interaction presented above (cf. Examples 10.51).

If the N particles have an internal structure, with internal Hilbert space H_{0k}, the overall system's Hilbert space will be isomorphic to $L^2(\mathbb{R}^{3N}) \otimes_{k=1}^{N} \mathsf{H}_{0k}$, and the possible Hamiltonians are more complicated, usually. We encourage the reader to consult the specialised texts [Mes99, CCP82, Pru81, ReSi80] for examples of this kin.

13.4.2 Entangled states and the so-called "EPR paradox"

A measuring device is not necessarily located at a point in space. On the contrary, if we want to measure quantities defined in space, first and foremost the position of a quantum particle, we must fill space with instruments: particle detectors that measure the position. The process of reduction of the state described by axiom **A3** is "instantaneous". This means that once a device has detected the particle at the point p and at time t, from that instant onwards no other device, as remote in space as we want from the first detector, will be able to detect the particle. The reduction of state therefore seems to be a *nonlocal* process: apparently it implies a "simultaneous" transmission of information between faraway places. This appears to violate the principles of the theory of relativity. In 1935 Einstein, Podolsky and Rosen [Des99, Bon97, Ghi97, Alb94], considering systems of two particles, showed the question can be phrased in physically-operative terms by which the violation seems to materialise effectively [EPR35].

Axiom **A7** describes the possible states of a compound quantum system. Let S be a system made of two subsystems A, B. The Hilbert space of S is $\mathsf{H}_S = \mathsf{H}_A \otimes \mathsf{H}_B$, in the obvious notations. The vectors of $\mathsf{H}_A \otimes \mathsf{H}_B$ are not just of the *factorised* sort $\psi_A \otimes \psi_B$, with one simple tensor product, for there are linear combinations of these products, too, like

$$\Psi = \frac{\psi_A \otimes \psi_B - \psi'_A \otimes \psi'_B}{\sqrt{2}}\,. \tag{13.69}$$

Pure states corresponding to unit vectors of the above form are said **entangled pure states**[6].

Consider the entangled state associated to the Ψ of (13.69), and let us suppose ψ_A and ψ'_A are eigenstates normalised to 1 of some observable G_A with discrete spectrum on system A, respectively corresponding to distinct eigenvalues a and a'. Assume the same for ψ_B, ψ'_B: they are unit eigenstates of an observable G_B with discrete spectrum on system B, with eigenvalues $b \neq b'$.

The discrete-spectrum observables G_A, G_B are, for instance, relative to internal freedom degrees of the systems A and B. They can typically be components of the spin or the polarisation of the particles. In that case the spaces H_A, H_B are also factorised into orbital space and internal space.

Until we measure it, the quantity G_A is not defined on the system, if the latter is in the state given by Ψ; there are two possible values a, a' with probability $1/2$ each. The same pattern is valid for G_B. The minute we measure G_A, reading – say – a (*a priori* unpredictable, in principle), the state of the *total system* changes, in allegiance to axiom **A3**, becoming the pure state of the unit vector

$$\psi_A \otimes \psi_B \,.$$

The crucial point is the following: if the initial state is the entangled state Ψ, the measurement of G_A determines *a measurement of G_B as well*: in the pure state associated to $\psi_A \otimes \psi_B$ the value of G_B is well defined, and equals b in our conventions. Any measurement of G_B can only give b.

Following the famous study of Einstein, Podolsky and Rosen, consider now compound systems of two particles A, B, prepared in the entangled pure state of the vector Ψ of (13.69), that *move away from each other* at great speed (i.e., the state's orbital part is the product of two very concentrated packets that separate rapidly from each other). In principle we can measure G_A and G_B on the respective particles in distant places and at so short lapses that no physical signal, travelling below the speed of light, can be transmitted from one experiment to the other in good time.

If axiom **A3** is to be valid, there should be a correlation between the outcomes: every time the reading of G_A is a (or a'), G_B will give b (respectively, b').

How can system A communicate to system B the outcome of the measurement of G_A in time to produce said correlations, without breaching the cornerstones of relativity?

This is a common situation in classical systems too, and in that case the explanation is very easy: there is no superluminal communication between the systems, for the correlations *preexist* the measurements. For example, let the observed quantities G_A, G_B be some particle "charge" or the like, and suppose the overall system S has charge 0 in the state in which it has been prepared, while the particles could have charge ± 1 corresponding to the aforementioned values a, a' and b', b. Then, *if we reason with classical particles*, we have to conclude one particle has charge 1, the other one -1. If the values of charge are preexistent, i.e. they exist before and independently of the fact we take a measurement to observe the charge, we can rest

[6] Analogously, for mixed states: $\rho \in \mathfrak{S}(\mathsf{H}_A \otimes \mathsf{H}_B)$ is **entangled** if it is *not* of the form $\rho_A \otimes \rho_B$, $\rho_A \in \mathfrak{S}(\mathsf{H}_A)$, $\rho_B \in \mathfrak{S}(\mathsf{H}_B)$.

assured that if a particle has charge 1 when measured, the second one will give -1 when observed, irrespective of where and when charges are measured, because the values are fixed *beforehand*.

The picture described by QM is, however, different: even if the state associated to Ψ has total charge $G = G_A + G_B$ equal 0, the subsystems's charges *are not defined* on the state of Ψ, and become fixed *at the time of the measurement* (of either state). Therefore the situation preexisting the measurement cannot be held responsible for the correlations predicted by QM, if we accept the standard interpretation of QM.

The idea of Einstein, Podolsky, Rosen was that if said correlations were really observed (as required by QM itself), and since defying the assumption of relativistic locality was out of the question, the reason for the correlations was due to a *preexisting* state to measurements. This being indescribable in the framework of the standard formulation of QM would have proved that the standard formulation of QM was, by nature, *incomplete*. (The probabilities used in QM, moreover, would reduce to mere epistemic probabilities).

J. Bell, in a brilliant paper of 1964 [Bel64][Bon97, Ghi97], measuring at least three types of "charges" producing correlations (in reality one measures three spin components for massive particles or polarisation states of photons), proved it is possible to *distinguish experimentally* between two situations, where charges are:

(i) fixed *before* the measurements;

or

(ii) fixed *at the same time* of the measurements.

Bell proved that case (i) occurs only if a series of inequalities on the outcomes hold: these are the celebrated *Bell inequalities*.

It is important to remark that a potential experimental infringement of Bell's inequalities does not automatically validate the standard formulation of QM. Nonlocal correlations, if observed experimentally, could in principle be justified without QM. What is true is that Quantum Mechanics, in contrast to Classical Mechanics, *forecasts* the presence of said correlations and at the same time the *violation* of Bell's inequalities, as we will see in short.

13.4.3 Bell's inequalities and their experimental violation

We will discuss briefly a simplified version of Bell's inequalities as proposed by Wigner. Take two particles A, B of spin $1/2$ produced together, in a region O, in the "singlet state", i.e. in the unique pure state of *zero total spin*. Fix an inertial system where the phenomenon is described. The entangled pure spin state is representable by Ψ_{sing} in the spin space $\mathsf{H}_{Aspin} \otimes \mathsf{H}_{Bspin}$:

$$\Psi_{sing} = \frac{\psi_+^{(n)} \otimes \psi_-^{(n)} - \psi_-^{(n)} \otimes \psi_+^{(n)}}{\sqrt{2}}\,, \tag{13.70}$$

where each H_{Aspin}, H_{Bspin} is isomorphic to $\mathbb{C}^2$, since each particle has spin $s = 1/2$.

Moreover, $\psi_+^{(n)}$ and $\psi_-^{(n)}$ are unit eigenvectors with respective eigenvalues $1/2$, $-1/2$ for the spin operator $S_n := \mathbf{n}\cdot\mathbf{S}$ along $\mathbf{n}$ (unit three-dimensional vector), for the single particle (as usual, $\hbar = 1$). The decomposition (13.70) holds for the singlet state Ψ_{sing}, *irrespective of where the axis* $\mathbf{n}$ *is*.

We suppose the particles part from each other. In other words the state's orbital part will, for example, be a product of wavefunctions, one in the orbital variables of A and one in the orbital variables of B, given by packets concentrated around their centres. The packets move away quickly from O in the chosen frame, so that the packets never overlap when the spin measurements are taken on A and B (we will not discuss the case of identical particles, which is practically the same). To study the correlation of spin measurements that violate locality, actually, it is not even necessary to assume the orbital part has the form we said. It suffices to place the devices measuring spin in faraway regions O_A, O_B (and far from O), and make sure the axis of the spin analyser of A in O_A can be re-oriented during consecutive measurements (see below) fast enough to prevent signals to propagate subluminally from O_A and reach O_B during measurements on the spin of B. This setup was concretely put into practice by Aspect's experiments.

To fix ideas imagine O_B is on the right of O and O_A on the left. The spin measurements in A and B (even two or more consecutive readings along distinct axes for each particle) can be taken, independent of one another, along given independent directions $\mathbf{u},\mathbf{v},\mathbf{w}$, not necessarily orthogonal. Assume at last that N pairs AB in spin singlet state are generated in O, and then each pair is analysed by spin measurements on A, B in O_A, O_B along three given independent unit vectors $\mathbf{u},\mathbf{v},\mathbf{w}$. Suppose that on the N pairs the values of the spin components are fixed *before* measuring in O_A and O_B, the contrary of what the standard formulation of QM predicts. *In order to have zero total spin*, for each pair AB the spin triples $(S_u,S_v,S_w)_A$ for A and $(S_u,S_v,S_w)_B$ for B must have opposite corresponding components. For instance, $(+,-,+)_A$ and $(-,+,-)_B$ are admissible, whereas $(+,+,+)_A$ and $(-,+,-)_B$ are not (from now on we abbreviate $+1/2$ with $+$ and $-1/2$ with $-$). There are 8 possible combinations altogether, tabled below.

Among the N pairs there will be N_1 pairs $(+,+,+)$ for A and $(-,-,-)$ for B *irrespective of whether measured or not*, N_2 pairs $(+,+,-)$ for A and $(-,-,+)$ for B *irrespective of whether measured or not*, and so on. At any rate we will have $N = \sum_{n=1}^{8} N_k$.

	part. A	part. B
N_1	$(+,+,+)$	$(-,-,-)$
N_2	$(+,+,-)$	$(-,-,+)$
N_3	$(+,-,+)$	$(-,+,-)$
N_4	$(+,-,-)$	$(-,+,+)$
N_5	$(-,+,+)$	$(+,-,-)$
N_6	$(-,+,-)$	$(+,-,+)$
N_7	$(-,-,+)$	$(+,+,-)$
N_8	$(-,-,-)$	$(+,+,+)$

With our "classical" hypotheses, every pair among the N examined must belong, after it has been created, in one of the sets, independently of what sort of spin measurements is taken. So let us suppose, for a certain pair, we measure S_u on A finding $+$, and S_v on B finding $+$. Then the pair can only belong to class 3 or 4, and there are N_3+N_4 possibilities out of N that this happens. If we call $p(u+,v+)$ the probability of finding $+$ measuring S_u on A and $+$ measuring S_v on B, we have

$$p(u+,v+) = \frac{N_3+N_4}{N} \,. \tag{13.71}$$

Similarly

$$p(u+,v+) = \frac{N_2+N_4}{N} \,, \quad p(w+,v+) = \frac{N_3+N_7}{N} \,. \tag{13.72}$$

Since $N_2, N_7 \geq 0$:

$$p(u+,v+) = \frac{N_3+N_4}{N} \leq \frac{N_2+N_4}{N} + \frac{N_3+N_7}{N} = p(u+,w+)+p(w+,v+) \,,$$

i.e. **Bell's inequalities** hold:

$$p(u+,v+) \leq p(u+,w+)+p(w+,v+) \,. \tag{13.73}$$

These inequalities hold whatever is the basis of unit vectors (not necessarily orthogonal) $\mathbf{u},\mathbf{v},\mathbf{w}$, if the values of the spin components along them are defined *before* we take the spin measurements and if the total spin of each pair is null. QM's prediction leads to a *violation* of the inequalities if we choose the axes suitably. Compute first $p(u+,v+)$ with the quantum recipe. Suppose the measure on A of S_u is $+$. Measuring S_u on B will give (or has already given) $-$, by (13.70). Anyway, particle B will have spin state represented by $\psi_-^{(u)}$ in the eigenvector basis of S_u. So we can evaluate $p(u+,v+)$ as:

$$p(u+,v+) = \frac{1}{2}\left|\left(\psi_-^{(u)}\middle|\psi_+^{(v)}\right)\right|^2 \,, \tag{13.74}$$

where $1/2$ is the initial probability of having $+$ on A when measuring S_u in state Ψ_{sing}. It is an easy exercise to compute the right-hand side of (13.74) in terms of the angle θ_{uv} between $\mathbf{u}$ and $\mathbf{v}$:

$$p(u+,v+) = \frac{1}{2}\sin^2\left(\frac{\theta_{uv}}{2}\right) \,. \tag{13.75}$$

The other terms in (13.73) are similar, so Bell's inequality (13.73) is equivalent to:

$$\sin^2\left(\frac{\theta_{uv}}{2}\right) \leq \sin^2\left(\frac{\theta_{uw}}{2}\right) + \sin^2\left(\frac{\theta_{wv}}{2}\right) \,. \tag{13.76}$$

It is not hard to see that a smart choice of angles invalidates the inequality. For example $\theta_{uv} = \pi/2$ implies $\sin^2\left(\frac{\theta_{uv}}{2}\right) = 1/2$. Setting $\theta_{uw} = \theta_{wv} = 2\phi$, the inequality becomes

$$\frac{1}{4} \leq \sin^2\phi \,,$$

clearly contradicted by independent axes $\mathbf{u},\mathbf{v},\mathbf{w}$ with $\theta_{uv} = \pi/2$ and $\theta_{uw} = \theta_{wv} = 2\phi \in (\pi/4, \pi/3)$.

Remarks 13.39. Despite the whole theory has unfolded within the non-relativistic formalism, one could already at this juncture raise an important question, unabated when passing to the relativistic formalism. In the quantum commutation of $p(u+,v+)$ we assumed we had measured *first* the spin of A, and *then* the spin of B. If measurements are taken in *causally disjoint* spacetime events – events that cannot be connected by future-directed timelike or spacelike paths – then the chronological order of the events is conventional, and depends on the choice of (inertial) frame, as is well known in special relativity. Thus we can find a frame where B is measured before A. So the question is whether computing $p(u+,v+)$ in this situation – that by the principle of relativity is physically equivalent to the previous one – gives the same result found earlier. Leaving behind the issue of a relativistic formalisation, the probability $p(u+,v+)$ is easily seen not to change, since the particles' spin observables commute. We will return to this kind of problem later. ■

Beginning from 1972 several experiments have been conducted to test the existence of the aforementioned correlations and the truth or falsity of Bell's inequalities (in particular, the decisive experiment was made in 1982 by A. Aspect, J. Dalibard and G. Roger [Bon97, Ghi97]). Experiments have proved, within the acceptability range of experimental errors, that (a) the nonlocal correlations predicted by QM *do exist*, (b) Bell's inequalities are *violated*.

Unless we deny the validity of the above experiments, therefore, and *independent of whether we accept or not the standard formulation of Quantum Mechanics*, we must agree that the correlations anticipated by Quantum Mechanics exist, and the outcomes are fixed at the moment of the measurements.

13.4.4 EPR correlations cannot transfer information

Although we developed QM in its non-relativistic version, the problems posed by the EPR analysis do not substantially change in the relativistic framework. But one question remains unanswered (we retain the conventions and notations of Chapter 13.4.2):

How does system A communicate to system B the outcome of the measurement of G_A, in time to produce the correlations we know of, and without destroying the cornerstones of relativity?

The answer is quite intricate, and by no means conclusive. First we have to say the question is ill posed, because it understates that the outcome of measuring G_A *causes* the outcome of G_B. In spacetime regions where the two measurements are taken (or can be taken) the latter are, in relativistic language, *causally disjoint*: there is no future-directed "spacelike" or "timelike" path in spacetime joining them. Well known from relativity is that there exists an inertial system in which A is measured *before B*, and another one where the situation is opposite: B's measurement *precedes in time* A's. So it makes no sense to say that the outcome of the experiment on B is the consequence, or the cause, of the outcome on B. One could, notwithstanding, resort to the partial conventionality of Einstein's synchronisation procedure in order to dismiss that problem. But despite the conventional choices underpinning special relativity it is known that the correlations between causally disjoint events are "dangerous" in relativistic theories, for they can spawn causal paradoxes: with a chain of causally disjoint events we can put two events in the history of a given system in any chronological order whatsoever. If it were possible to use the correlations of causally disjoint events to transfer information either way, we would be able to communicate with the past (inside the light cone) and thereby obtain causal paradoxes.

It can be proved (see [Bon97] and references for a detailed study) that *by accepting the standard formulation of QM* for systems made by entangled states like (13.69) (but also general entangled mixed states), no piece of information can be transmitted from (event) X, where part of the system is measured, to (event) Y, where the other part is measured, by measuring arbitrary pairs of quantities and exploiting the quantum correlations between the readings. Not only that, but observing the outcomes on one part of the system we cannot establish *whether* on the other part measurements have been taken, if they are being taken as we speak, nor if they will be taken in the future.

Let us examine two ways of transferring information from X to Y via EPR correlations.

- Consider the single pairs of measurements on A, B of the observables G_A, G_B, which we know have correlated outcome. We cannot pass information from X to Y using the correlation, because the outcome, albeit correlated, is completely accidental. It is like having two coins A, B with the remarkable property that each time one shows "heads", the other one gives "tails", independent of the fact they are tossed far away, rapidly, and that A is tossed before or after B in some frame. The coins, though, have a quantum character and it is physically impossible to force one to give a certain result: the outcome of the toss is determined in a probabilistic way and whatever our wish is. Thus the two coins, i.e. our quantum system made by parts A and B, cannot be used as a Morse telegraph of sorts to transmit information between X and Y.
- The second possibility is to consider not the single measurements of G_A and G_B, but a large number thereof, and study the statistical features of the outcome distributions. The statistics of the measurements of G_A migh be different according to whether we measure G_B as well, or whether we measure a new quantity G'_B. In this way, by measuring or not measuring G_B (and measuring G'_B or measuring nothing at all) in Y, we can send an elementary signal to X, of the type "yes" or "no", that we recover by checking experimentally the statistics of A. We claim that this procedure, neither, allows to transfer information, since the statistics relative to G_A is exactly the same in case we also measure G_B (or any other G'_B) or we do not measure G_B. Consider the state $\rho \in \mathfrak{S}(\mathsf{H}_A \otimes \mathsf{H}_B)$ of the system composed by A, B. Suppose $G_A = G^{(A)} \otimes I_B$, with $G^{(A)}$ self-adjoint on H_A, has discrete and finite spectrum $\{g_1^{(A)}, g_2^{(A)}, \ldots, g_n^{(A)}\}$, with eigenspaces $H_{g_k^{(A)}} \subset \mathsf{H}_A \otimes \mathsf{H}_B$ as ranges of the orthogonal projectors $P_k^{(G_A)} := P_k^{G^{(A)}} \otimes I_B$. Similarly, $G^{(B)}$ is self-adjoint on H_B, $G_B = I_A \otimes G^{(B)}$ has spectrum $\{g_1^{(B)}, g_2^{(B)}, \ldots, g_m^{(B)}\}$ discrete and finite, the eigenspaces $H_{g_k^{(B)}} \subset \mathsf{H}_A \otimes \mathsf{H}_B$ are targets of orthogonal projectors $P_k^{(G_B)} := I_A \otimes P_k^{G^{(B)}}$.

If we measure G_B on state ρ reading $g_k^{(B)}$, the post-measurement state is

$$\frac{1}{tr\left(P_k^{(G_B)}\rho P_k^{(G_B)}\right)}P_k^{(G_B)}\rho P_k^{(G_B)}\,.$$

Considering all possible readings of B, if we measure first B and then A (in some frame), the system we want to test on A is the mixture

$$\rho' = \sum_{k=1}^{m}\frac{p_k}{tr\left(P_k^{(G_B)}\rho P_k^{(G_B)}\right)}P_k^{(G_B)}\rho P_k^{(G_B)}$$

where $p_k = tr(P_k^{(G_B)}\rho P_k^{(G_B)})$ is the probability of reading $g_k^{(B)}$ for B. Altogether

$$\rho' = \sum_{k=1}^{m} P_k^{(G_B)}\rho P_k^{(G_B)}\,.$$

Hence the probability of getting $g_h^{(A)}$ for A, when B has been measured (irrespective of the latter's outcome), is:

$$\mathscr{P}(g_h^{(A)}|B) = tr(\rho' P_h^{(G_A)}) = tr\left(\sum_{k=1}^{m} P_k^{(G_B)}\rho P_k^{(G_B)}P_h^{(G_A)}\right)\,.$$

The trace is linear and invariant under cyclic permutations, so

$$\mathscr{P}(g_h^{(A)}|B) = \sum_{k=1}^{m} tr(P_k^{(G_B)}\rho P_k^{(G_B)}P_h^{(G_A)}) = \sum_{k=1}^{m} tr(\rho P_k^{(G_B)}P_h^{(G_A)}P_k^{(G_B)})$$

$$= \sum_{k=1}^{m} tr(\rho P_k^{(G_B)}P_k^{(G_B)}P_h^{(G_A)})\,.$$

In the last passage we used $P_k^{(G_B)}P_h^{(G_A)} = P_h^{(G_A)}P_k^{(G_B)}$ from the structure of the projectors. On the other hand $P_k^{(G_B)}P_k^{(G_B)} = P_k^{(G_B)}$ and $\sum_k P_k^{(G_B)} = I$ by the spectral theorem. Therefore

$$\mathscr{P}(g_h^{(A)}|B) = \sum_{k=1}^{m} tr\left(\rho P_k^{(G_B)}P_h^{(G_A)}\right) = tr\left(\rho\sum_{k=1}^{m} P_k^{(G_B)}P_h^{(G_A)}\right) = tr\left(\rho P_h^{(G_A)}\right)$$

$$= \mathscr{P}(g_h^{(A)})\,.$$

The final result is: *the probability of obtaining $g_h^{(A)}$ from A when the quantity B has been measured (with any possibile outcome) coincides with the probability of obtaining $g_h^{(A)}$ from A without measuring B.*

So even by considering the statistics of outcomes of A, there is no way to transmit information by EPR correlations: when measuring part B of the system, the presence or the absence of the correlations is completely irrelevant if we observe only part A. Therefore, Quantum Mechanics and Special Relativity seem to coexist peacefully. In reality the above discussion turns a blind eye on whether spacetime is classical or relativistic. Apparently, the lesson learned is that the processes of compound quantum systems are not describable in spacetime. Only the *outcomes* of measurements, interpreted as states of macroscopic systems (detectors, meters, *etc*...) can be described in spacetime using events. Spacetime allegedly resembles an *a posteriori* structure in which macroscopic phenomena are recorded, sometimes in relationship to microscopic phenomena. But this is not the only possible way to look at things. The apparent violation of locality due to the "collapse of the state" might in fact be a purely speculative construction, related to an all-too-simplistic model of the measuring procedures. Furthermore, a careful analysis might reveal that spacetime categories carry on being fundamental at the quantum level as well. In this respect see the recent study [Dop09].

13.4.5 The phenomenon of decoherence as a manifestation of the macroscopic world

It must be clear that the point of view outlined in the previous section must be considered as a starting point and not the end of the journey, at least until we understand, experimentally, what a macroscopic/classical system is, what a microscopic/quantum system is, and which are the reasons for switching from one regime to the other.

An interesting perspective for recovering the classical world from the quantum one is based on the notion of *decoherence* [BGJKS00]. We present the main idea quite rapidly (see in particular [Kup00], [Zeh00]). Consider a quantum system S interacting with a quantum system E, the latter seen as the *ambient* where the evolution takes place and that includes measuring instruments and any other object that interacts with S. The evolution is described on the Hilbert space $\mathsf{H} = \mathsf{H}_S \otimes \mathsf{H}_E$ by an operator (unitary and strongly continuous) $\mathbb{R} \ni t \mapsto U_t$. If $\rho(0)$ is a state (mixed in general) of the total system at time $t = 0$, measurements of observables on S at time t are taken using an effective statistical operator $\rho_S(t)$ of the form:

$$\rho_S(t) = tr_E\left(U_t \rho(0) U_t^{-1}\right) , \tag{13.77}$$

where $tr_E(W)$ denotes the **partial trace** with respect to E of the self-adjoint $W \in \mathfrak{B}_1(\mathsf{H}_S \otimes \mathsf{H}_E)$ (we used it tacitly in the previous section as well). $tr_E(W)$ is the unique self-adjoint operator in $\mathfrak{B}_1(\mathsf{H}_S)$ for which

$$\sum_{n\in\mathbb{N}} (\phi \otimes \psi_n | W \phi \otimes \psi_n) = (\phi | tr_E(W)\phi) , \quad \phi \in \mathsf{H}_S$$

in any basis $\{\psi_n\}_{n\in\mathbb{N}} \subset \mathsf{H}_E$. The role of the partial trace (for the subsystem E) is to assign, in a natural way, a state ($tr_E(W)$) to the subsystem (S) with respect to which the trace is not taken, whenever we have a state (W) for the total system ($S+E$). If an observable $A \otimes I_{\mathsf{H}_E}$ is bounded on S, as expected $tr(A\, tr_E(W)) = tr(A \otimes I_{\mathsf{H}_E} W)$.

The evolution given by (13.77), in general, cannot be expressed canonically by a unitary evolution operator acting directly on $\rho_S(0)$. This approach seems to account well for the experimental behaviour of many a system that interact intensely with the ambient (like macromolecules). In certain cases the interaction with the ambient determines a collection $\{P_k\}_{k\in K} \subset \mathfrak{B}(\mathsf{H}_S)$ of pairwise-orthogonal projectors, not dependent on the overall state, for which *almost instantaneously* the state ρ_t satisfies:

$$\rho_S(t) = \sum_{k\in K} P_k \rho_S(t) P_k \,.$$

Any mechanism due to the interaction of S and E that produces this situation is called a **decoherence** process. In practice decoherence corresponds to a *dynamical procedure that generates a superselection rule* for S, whose coherent sectors are the projection spaces of the P_k which give propositions about quantities that are typically considered completely classical. A mechanism of this sort (see [Kup00] and the models therein) could shed light on the reasons why large molecules, for example, have geometric features that vary with continuity and can be described in classical terms. What is more, it could elucidate why certain macroscopic objects behave classically; perhaps it could explain, alternatively, what in the common interpretation of the formalism goes under the name of collapse of the state (which in reality would never occur) even though it is not clear how to justify the apparent violation of locality [BGJKS00]. An elementary physical process leading to the superselection of the mass, once assumed that the spectrum mass is a discrete set or positive reals, was presented in [AnMo12].

13.4.6 Axiom A8*: compounds of identical systems*

The elementary particles of QM are *identical particles*. That they cannot be distinguished is formalised in QM in a precise way by keeping axiom **A7** in account, as we will see in a moment.

First we need a few technical results.

Definition 13.40. *The* **permutation group** *on n elements $\mathscr{P}_n$ is the set of bijective maps $\sigma : \{1,2,\dots,n\} \to \{1,2,\dots,n\}$ (called* **permutations of** *n* **objects***) equipped with composition product.*
In particular, a **permutation of two objects** *is a $\sigma \in \mathscr{P}_n$ that restricts to the identity on a subset of $n-2$ elements of $\{1,2,\dots,n\}$.*

To any $\sigma \in \mathscr{P}_n$ we associate a number $(-1)^\sigma \in \{-1,+1\}$ called its **parity**. If σ is the product of an even number of permutations of two objects then $(-1)^\sigma := 1$, while if the number of permutaions is odd, $(-1)^\sigma := -1$. Despite the number of permutations of two objects appearing in σ is not uniquely determined, the parity is, as one can show.

Consider a Hilbert space H and its n-fold tensor product $\mathsf{H}^{\otimes n} := \bigotimes_{i=1}^n \mathsf{H}$. Any $\sigma \in \mathscr{P}_n$ induces a unitary operator $U_\sigma : \mathsf{H}^{\otimes n} \to H^{\otimes n}$ defined as follows. Pick a basis N for H. By Proposition 10.25 the vectors $\psi_1 \otimes \cdots \otimes \psi_n$, with $\psi_k \in N$, $k = 1,2,\dots,n$, form a basis of $\mathsf{H}^{\otimes n}$. If σ is an arbitrary permutation also $\psi_{\sigma^{-1}(1)} \otimes \cdots \otimes \psi_{\sigma^{-1}(n)}$ will give a basis for $\mathsf{H}^{\otimes n}$. This basis, actually, is precisely the same one we had before acting by σ, up to rearrangements. Define $U_\sigma : \mathsf{H}^{\otimes n} \to \mathsf{H}^{\otimes n}$ as the unique bounded operator satisfying

$$U_\sigma(\psi_1 \otimes \cdots \otimes \psi_n) := \psi_{\sigma^{-1}(1)} \otimes \cdots \otimes \psi_{\sigma^{-1}(n)} , \quad \psi_k \in N, k = 1, 2, \ldots, n.$$

U_σ is unitary for it preserves bases; moreover if $\phi_1, \ldots \phi_n \in \mathsf{H}$ are arbitrary (also not in N), decomposing over the ψ_i and exploiting linearity and continuity gives

$$U_\sigma(\phi_1 \otimes \cdots \otimes \phi_n) := \phi_{\sigma(1)} \otimes \cdots \otimes \phi_{\sigma(n)} .$$

This proves half of the following proposition.

Proposition 13.41. *Consider* $\mathsf{H}^{\otimes n} := \bigotimes_{i=1}^n \mathsf{H}$*, where* H *is a Hilbert space, and the permutation group* $\mathscr{P}_n$ *on n elements.*

(a) *If* $\sigma \in \mathscr{P}_n$ *there exists a unique unitary operator* $U_\sigma : \mathsf{H}^{\otimes n} \to \mathsf{H}^{\otimes n}$ *such that:*

$$U_\sigma(\phi_1 \otimes \cdots \otimes \phi_n) := \phi_{\sigma^{-1}(1)} \otimes \cdots \otimes \phi_{\sigma^{-1}(n)} , \tag{13.78}$$

for any $\phi_1, \ldots, \phi_n \in \mathsf{H}$.
(b) $U : \mathscr{P}_n \ni \sigma \mapsto U_\sigma$ *is a faithful unitary representation of* $\mathscr{P}_n$.

Proof. (a) The claim descends from the arguments preceding the proposition: just define U_σ via a basis and check (13.78) holds for *any* $\phi_1, \ldots, \phi_n \in \mathsf{H}$. Two bounded operators satisfying (13.78) coincide on a basis, hence everywhere (being bounded).
(b) By direct inspection (using the fact that σ^{-1} appears in the right-hand side of (13.78)) $(U_\sigma U_{\sigma'})(\phi_1 \otimes \cdots \otimes \phi_n) = U_{\sigma \circ \sigma'}(\phi_1 \otimes \cdots \otimes \phi_n)$. Linearity and continuity imply $U_\sigma U_{\sigma'} = U_{\sigma \circ \sigma'}$, making $\sigma \mapsto U_\sigma$ a (unitary) representation of $\mathscr{P}_n$. Faithfulness is granted by U's injectivity, since $U_\sigma = I$ implies $\phi_{\sigma^{-1}(1)} \otimes \cdots \otimes \phi_{\sigma^{-1}(n)} = \phi_1 \otimes \cdots \otimes \phi_n$ for any orthonormal $\phi_1, \ldots, \phi_n \in \mathsf{H}$, hence $\sigma^{-1} = id = \sigma$. □

Physically, if $\Psi \in \mathsf{H}^{\otimes n}$ is a pure state of a system made of n identical subsystems, each described on its own Hilbert space H, the pure state of $U_\sigma \Psi$ is naturally interpreted as the state in which the n subsystems have been permuted under σ. The action of U_σ extends to all states $\rho \in \mathfrak{S}(\mathsf{H}^{\otimes n})$ by the transformation that maps ρ to $U_\sigma \rho U_\sigma^{-1}$. As U_σ is unitary, the transformation preserves ρ's positivity and trace ($U_\sigma \rho U_\sigma^{-1}$ is of trace class if ρ is, because trace-class operators form an ideal), so $U_\sigma \rho U_\sigma^{-1} \in \mathfrak{S}(\mathsf{H}^{\otimes n})$ if $\rho \in \mathfrak{S}(\mathsf{H}^{\otimes n})$.

The permutation group's action on states dualises to an action on propositions $P \in \mathfrak{P}(\mathsf{H}^{\otimes n})$ on the system. The dual action, as usual, is given by the transformation mapping P to $U_\sigma^{-1} P U_\sigma$. Since U_σ is unitary, $U_\sigma^{-1} P U_\sigma$ is an orthogonal projector if P is.

By the properties of the trace (Proposition 4.36(c))

$$tr\left(U_\sigma^{-1} P U_\sigma \, \rho\right) = tr\left(P \, U_\sigma \rho U_\sigma^{-1}\right) .$$

As always, then, letting the permutation group act on states or on propositions is physically the same for computing the truth probability of measured propositions.

The natural interpretation of the transformation associating P to $U_\sigma^{-1}PU_\sigma$ is to take P and permute the subsystems with σ. The action of U_σ on propositions induces an action on each PVM $\{P^{(A)}(E)\}_{E\in\mathscr{T}(\mathbb{R})}$ (associated to the observable A) that maps it to a PVM $\{U_\sigma^{-1}P^{(A)}(E)U_\sigma\}_{E\in\mathscr{T}(\mathbb{R})}$. From the spectral theorem we know the latter action corresponds to transforming the observable A into $U_\sigma^{-1}AU_\sigma$. The physical meaning is obvious in the light of previous considerations. Now we are ready to state the axiom for compounds of identical systems.

A8. *If a physical system S consists of $n < +\infty$ identical subsystems, each described on one copy of the Hilbert space H, physically-admissible propositions correspond to the subset in $\mathfrak{P}(\mathsf{H}^{\otimes n})$ of orthogonal projectors invariant under the permutation group (cf. Proposition 13.41). Equivalently: $P \in \mathfrak{P}(\mathsf{H}^{\otimes n})$ makes physical sense on S only if*

$$U_\sigma^{-1}PU_\sigma = P\,, \quad \text{for every } \sigma \in \mathscr{P}_n.$$

Therefore physically-admissible observables A on S are those whose spectral measures satisfy the above condition, i.e.

$$U_\sigma^{-1}AU_\sigma = A\,, \quad \text{for every } \sigma \in \mathscr{P}_n.$$

Just for example, if we work with a compound of two identical particles of mass m, with coordinates $x_i^{(1)}$ and $x_i^{(2)}$, an admissible observable is the ith component of the mean position $(X_i^{(1)} + X_i^{(1)})/2$. Without going into details, using the spectral measures of $X_i^{(1)}$ and $X_i^{(2)}$ we can construct an admissible proposition (an orthogonal projector commuting with every U_σ) corresponding to the statement: "one of the particles has ith coordinate falling within the Borel set E". Conversely, propositions like " particle 1 has ith coordinate falling in the Borel set E" are not admissible.

13.4.7 Bosons and Fermions

At last, we would like to show one consequence of axiom **A8** worthy of mention. Consider the usual system S made of n identical subsystems. Take $\sigma \in \mathscr{P}_n$ and the λ-eigenspace of U_σ inside $\mathsf{H}^{\otimes n}$:

$$(\mathsf{H}^{\otimes n})_\lambda^{(\sigma)} := \{\Psi \in \mathsf{H}^{\otimes n} \mid U_\sigma\Psi = \lambda\Psi\}\,.$$

Note U_σ is unitary, so $|\lambda| = 1$.

Every meaningful proposition must commute with U_σ, so if the system's state $\Psi \in (\mathsf{H}^{\otimes n})_\lambda^{(\sigma)}$ is initially pure, following a measument by the admissible (true) proposition P the state will be described by $P\Psi/||P\Psi||$; this is in $(\mathsf{H}^{\otimes n})_\lambda^{(\sigma)}$ since $U_\sigma \frac{P\Psi}{||P\Psi||} = \frac{U_\sigma P\Psi}{||P\Psi||} = \frac{PU_\sigma\Psi}{||P\Psi||} = \lambda \frac{P\Psi}{||P\Psi||}$. By taking measurements, therefore, we cannot "make the system leave" the space $(\mathsf{H}^{\otimes n})_\lambda^{(\sigma)}$ if it was in a pure state described by a vector in $(\mathsf{H}^{\otimes n})_\lambda^{(\sigma)}$ immediately prior to the measurement. Not even time evolution, at least under time homogeneity, “allows the system to leave” the space $(\mathsf{H}^{\otimes n})^{(\sigma)}_{\lambda}$ if it was, at the initial time, in a pure state in $(\mathsf{H}^{\otimes n})^{(\sigma)}_{\lambda}$. In fact, the Hamiltonian observable H (being admissible) will have commuting spectral measure with U_σ. Consequently

$$e^{-itH}U_\sigma = \int_{\sigma(H)} e^{-ih}\, dP^{(H)}(h) U_\sigma = U_\sigma \int_{\sigma(H)} e^{-ih}\, dP^{(H)}(h) = U_\sigma e^{-itH} .$$

If $U_\sigma \Psi = \lambda\Psi$, then $U_\sigma \Psi_t = U_\sigma e^{-itH}\Psi = e^{-itH}U_\sigma\Psi = e^{-itH}\lambda\Psi = \lambda\Psi_t$, so $\Psi_t \in (\mathsf{H}^{\otimes n})^{(\sigma)}_{\lambda}$, for any time $t \in \mathbb{R}$. In case the evolution operator is not the exponential of the Hamiltonian (lack of time homogeneity), under suitable assumptions one can still prove the same result. This happens if, for instance, the evolution operator is given by the Dyson series (see Proposition 13.19) for a special class of time-dependent Hamiltonian observables. We have the following result.

Proposition 13.42. *Suppose a compound system S is made of $n < +\infty$ identical subsystems, each described on the same Hilbert space H, and at some time t_0 the system is in a pure state represented by a vector in $(\mathsf{H}^{\otimes n})^{(\sigma)}_{\lambda}$ for some $\sigma \in \mathscr{P}_n$. Then the evolution (in regime of time homogeneity), or a measurement, leaves the system in a pure state represented by a vector in $(\mathsf{H}^{\otimes n})^{(\sigma)}_{\lambda}$.*

The experimental evidence not only confirms this fact, but shows that pure states of a compound of identical particles in 4 dimensions (three for space plus time) are determined by vectors in *two subspaces only*, built intersecting the $(\mathsf{H}^{\otimes n})^{(\sigma)}_{\lambda}$. (Moreover, mixtures are incoherent superpositions of pure states in those two subspaces). To explain that fact we need a few comments.

Consider a permutation $\delta \in \mathscr{P}_n$ of two elements, so $\delta \circ \delta = id$ and $U_\delta U_\delta = I$. As U_δ is unitary, U_δ is self-adjoint. Hence U_δ is an observable, actually a constant of motion (exercise). Not just that: $\sigma(U_\delta) \subset \{-1, 1\}$ because $\sigma(U_\delta)$ is contained in $\mathbb{R}$ (U_δ is self-adjoint) and also in the unit circle in $\mathbb{C}$ (U_δ is unitary). Therefore $\sigma(U_\delta) = \sigma_p(U_\delta)$ because the spectrum consists of one or two isolated points. It is easy to prove $\sigma_p(U_\delta) = \{-1, 1\}$. In fact, if δ swaps the kth and jth elements, every vector of $\mathsf{H}^{\otimes n}$ of the form

$$(\psi_1 \otimes \cdots \otimes \psi_k \otimes \cdots \otimes \psi_j \otimes \cdots \otimes \psi_n) \pm (\psi_1 \otimes \cdots \otimes \psi_j \otimes \cdots \otimes \psi_k \otimes \cdots \otimes \psi_n)$$

is an eigenvector of U_δ with eigenvaule ± 1. From this follows, for any $\sigma \in \mathscr{P}_n$, that U_σ admits the eigenvalues (possibly coinciding) 1 and $(-1)^\sigma$. It is enough to write $U_\sigma = U_{\delta_1} \cdots U_{\delta_m}$, where the σ_i are permutations of two elements. The intersections

$$(\mathsf{H}^{\otimes n})^{(\sigma)}_{+} := \cap_{i=1}^{m} (\mathsf{H}^{\otimes n})^{(\delta_i)}_{+1} \quad \text{and} \quad (\mathsf{H}^{\otimes n})^{(\sigma)}_{-} := \cap_{i=1}^{m} (\mathsf{H}^{\otimes n})^{(\delta_i)}_{-1}$$

are eigenspaces for U_σ with respective eigenvalues $+1$ and $(-1)^\sigma$, since $U_\sigma = U_{\delta_1} \cdots U_{\delta_m}$.

$\mathsf{H}^{\otimes n}$ has two physically-interesting closed subspaces, obtained from the intersections of all spaces of type $(\mathsf{H}^{\otimes n})^{(\sigma)}_{+}$ and $(\mathsf{H}^{\otimes n})^{(\sigma)}_{-}$, respectively, as $\sigma \in \mathscr{P}_n$ varies. These are the **totally symmetric product**

$$(\mathsf{H}^{\otimes n})_{+} := \{\Psi \in \mathsf{H}^{\otimes n} \mid U_\sigma \Psi = \Psi, \forall \sigma \in \mathscr{P}_n\}$$

and the **totally skew-symmetric product**

$$(\mathsf{H}^{\otimes n})_- := \{\Psi \in \mathsf{H}^{\otimes n} \mid U_\sigma \Psi = (-1)^\sigma \Psi, \forall \sigma \in \mathscr{P}_n\}.$$

Their physical relevance lies in that every known compound of identical particles has pure states described by vectors either in $(\mathsf{H}^{\otimes n})_+$ or in $(\mathsf{H}^{\otimes n})_-$. Precisely, particles of the first type, called **Bosons** (whose mixtures are incoherent superpositions of pure states in $(\mathsf{H}^{\otimes n})_+$), have integer spin; particles of the second type (whose mixtures are incoherent superpositions of pure states in $(\mathsf{H}^{\otimes n})_-$), called **Fermions**, have semi-integer spin. This phenomenon is often referred to as the *spin statistical correlation*. Within the non-relativistic formulation of QM there is no proof for this physical constraint on the structure of states, nor for the relationship to the value of spin. In it we can only show, using Proposition 13.42, that if a system of particles has a Fermionic behaviour, or a Bosonic behaviour, at time t_0, it will maintain the behaviour so long it is described by pure states. In the non-relativistic formulation there are states, compatible with Proposition 13.42, that are neither symmetric not skew-symmetric. One says, in jargon, such systems obey a *parastatistics*. Particles of this sort have never been observed.

Many authors (mainly W. Pauli) obtained within the relativistic formulation of QM – more precisely the *Relativistic Quantum Field Theory* on 4-dimensional Minkowski spacetime – a famous theorem, aptly called *spin statistical correlation theorem* [StWi00]. It proves that the restriction on pure states and the spin statistical correlation observed experimentally are consequences of the theory's invariance under the Poincaré group rather than the Galilean group. In three-dimensional spacetime models there are compounds of identical particles that do not abide by Fermi's statistics, not Bose's one. These are the so-called *anions*, useful in explaining phenomena like the *fractional quantum Hall effect* [Ste08].

In conclusion we mention that when we deal with compounds of *infinitely many* identical subsystems described on H, the natural Hilbert spaces to develop the theory are the subspaces of the Fock space (Example 10.27(3)):

$$\mathscr{F}_+(\mathsf{H}) := \bigoplus_{n=0}^{+\infty} (\mathsf{H}^{n\otimes})_+ \quad \text{and} \quad \mathscr{F}_-(\mathsf{H}) := \bigoplus_{n=0}^{+\infty} (\mathsf{H}^{n\otimes})_- ,$$

called **Bosonic Fock space** and **Fermionic Fock space** generated by H. As usual we assumed $(\mathsf{H}^{0\otimes})_\pm := \mathbb{C}$, and that the unique pure state determined by $(\mathsf{H}^{0\otimes})_\pm$ is the *vacuum state* of the system. Within this framework lives *quantum field theory*, for which fields are "replaced" by systems of infinitely many identical Bosonic or Fermionic particles.

Exercises

13.1. Consider a mixed state $\rho \in \mathfrak{S}(\mathsf{H})$ and an orthogonal sum $\mathsf{H} = \oplus_{k \in K} \mathsf{H}_k$, with K finite or countable, associated to orthogonal projectors $\{P_k\}_{k \in K}$.

Using the strong topology define

$$\rho' := \text{s-}\sum_k P_k \rho P_k \,.$$

Prove ρ' is well defined and $\rho' \in \mathfrak{S}(\mathsf{H})$.

Hint. $P_k P_h = 0$ if $k \neq h$, s-$\sum_k P_k = I$, and $||\rho^2|| \leq 1$. This allows to prove the series converges strongly, using known properties of series of orthogonal vectors.
That ρ' is positive and $||\rho'|| \leq 1$ follows by the construction and the similar properties of ρ. Using a basis N of H, union of bases for each summand H_k, with Proposition 4.29 one proves ρ' is of trace class, plus $tr\rho' = tr\rho = 1$.

13.2. In relationship to Chapter 13.4.4, where the probability of measuring $g_k^{(A)}$ for G_A on part A of a quantum system is proven to be independent of the fact that G_B is measured on part $B \neq A$, prove that the result is valid for arbitrary observables (even with continuous and unbounded spectrum). Assume the device measuring G_B gives as possible readings a countable disjoint family of Borel sets $E_k^{(G_B)}$ whose union is $\sigma(G_B)$.

13.3. Referring to (13.78) prove that $U_\sigma U_{\sigma'} = U_{\sigma\circ\sigma'}$.

Solution. By linearity and exploiting the fact that the operators are bounded and everywhere defined, it is sufficient proving that

$$U_\sigma(U_{\sigma'}(\phi_1 \otimes \ldots \otimes \phi_n)) = U_{\sigma\circ\sigma'}(\phi_1 \otimes \ldots \otimes \phi_n)\,.$$

Let us establish that identity. If $\sigma, \sigma' \in \mathscr{P}_n$ then:

$$U_\sigma(U_{\sigma'}(\phi_1 \otimes \ldots \otimes \phi_n)) = U_\sigma(\phi_{\sigma'^{-1}(1)} \otimes \ldots \otimes \phi_{\sigma'^{-1}(n)})\,.$$

Re-defining $u_i := \phi_{\sigma'^{-1}(i)}$ so that $u_{\sigma^{-1}(j)} := \phi_{\sigma'^{-1}(\sigma^{-1}(j))}$, one finds

$$U_\sigma(U_{\sigma'}(\phi_1 \otimes \ldots \otimes \phi_n)) = u_{\sigma^{-1}(1)} \otimes \ldots \otimes u_{\sigma^{-1}(n)}$$

$$= \phi_{\sigma'^{-1}\circ\sigma^{-1}(1)} \otimes \ldots \otimes \phi_{\sigma'^{-1}\circ\sigma^{-1}(n)} = \phi_{(\sigma\circ\sigma')^{-1}(1)} \otimes \ldots \otimes \phi_{(\sigma\circ\sigma')^{-1}(n)}$$

$= U_{\sigma\circ\sigma'}(\phi_1 \otimes \ldots \otimes \phi_n)$ as wanted.

13.4. Consider a compound of n identical particles in $\mathsf{H}^{\otimes n}$. Prove that under axiom **A8**, if $\delta \in \mathscr{P}_n$ is a permutation on two elements, then U_δ is a constant of motion.

13.5. Prove that $(\mathsf{H}^{\otimes n})_+$ and $(\mathsf{H}^{\otimes n})_-$ are orthogonal, that $\mathsf{H}^{\otimes 2} = (\mathsf{H}^{\otimes 2})_+ \oplus (\mathsf{H}^{\otimes 2})_-$ if $n = 2$, and that the previous fact is false already for $n = 3$.

This page intentionally left blank

14

Introduction to the Algebraic Formulation of Quantum Theories

I would like to make a confession which may seem immoral:
I do not believe absolutely in Hilbert spaces any more.
von Neumann, letter to Birkhoff
about the mathematical formulation of QM (1935)

In the last chapter of the book we offer a short presentation of the algebraic formulation of quantum theories, and we will state and prove a central theorem about the so-called *GNS construction*. We will discuss how to treat the notion of quantum symmetry in this framework, by showing that an algebraic quantum symmetry can be implemented (anti)unitarily in GNS representations of states invariant under the symmetry.

As general references, mostly concerned with the algebraic formulation of quantum field theories, we recall [Emc72], [Haa96], [Ara09], [Rob04], and the more recent [Str05a, Str11] on the algebraic formulation of QM. On the mathematical side, detailed and critical studies on the present material are [BrRo02] and [KaRi97].

We will routinely resort to the definitions and notions of Definition 3.48 throughout the chapter.

14.1 Introduction to the algebraic formulation of quantum theories

The fundamental Theorem 11.22 of Stone–von Neumann is stated in the jargon of theoretical physics as follows:

All irreducible representations of the Weyl algebra with a finite, and fixed, number of freedom degrees are unitarily equivalent,

or

All irreducible representations of the CCRs with a finite, and fixed, number of freedom degrees are unitarily equivalent.

The expression *unitarily equivalent* refers to the existence of a Hilbert-space isomorphism S, and the finite number of degrees of freedom is the dimension of the symplectic space X on which the Weyl algebra is built.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7_14, © Springer-Verlag Italia 2013

What happens then in infinite dimensions? Let us keep irreducibility, and suppose we pass from X finite-dimensional – parametrising, e.g., the coordinates of a point-particle in phase space – to X infinite-dimensional – describing a suitable solution space to free Bosonic field equations, say. Then the Stone–von Neumann theorem no longer holds. Theoretical physicists would say that

There exist non-equivalent CCR representations with an infinite number of freedom degrees.

What happens in this stuation, in practice, is that one can find strongly continuous irreducible representations π_1, π_2, on respective (separable) Hilbert spaces $\mathsf{H}_1, \mathsf{H}_2$, of the Weyl *-algebra $\mathfrak{A} := \mathscr{W}(\mathsf{X}, \sigma)$ (here thought of as C^*-algebra, with no change in the results) associated to the physical system under exam (a quantised Bosonic field, typically), that admit *no* isomorphism $S : \mathsf{H}_1 \to \mathsf{H}_2$ satisfying:

$$S\pi_1(a)\,S^{-1} = \pi_2(a)\,, \quad \text{for any } a \in \mathfrak{A}.$$

Pairs of this kind are called *(unitarily) non-equivalent*. Jumping from X being finite-dimensional to infinite-dimensional corresponds to passing from Quantum Mechanics to quantum field theory (possibly relativistic, and on curved spacetime). In these situations (but not only), the existence of non-equivalent representations has often to do with *spontaneous symmetry breaking*. The presence of non-equivalent representations of one single physical system (the pair (X, σ)) shows that a formulation in a fixed Hilbert space is fully inadequate, and we must free ourselves of the structure of Hilbert space in order to lay the foundations of quantum theories in broader generality.

This programme has been developed by and large, starting from the pioneering work of von Neumann himself, and is nowadays called *algebraic formulation of quantum (field) theories*. Within this framework it was possible to formalise, for example, field theories in curves spacetime in relationship to the quantum phenomenology of black-hole thermodynamics.

14.1.1 Algebraic formulation and the GNS theorem

The algebraic formulation prescinds, anyway, from the nature of the quantum system and may be stated for systems with finitely many freedom degress as well [Str05a]. The viewpoint falls back on two assumptions [Haa96, Ara09, Str05a, Str11] (which somehow generalise the results of Chapter 7.4.6).

AA1. *A physical system S is described by its* **observables***, viewed now as self-adjoint elements in a certain C^*-algebra $\mathfrak{A}_S$ with unit $\mathbb{I}$ associated to S.*

AA2. *An* **algebraic state** *on $\mathfrak{A}_S$ is a linear functional $\omega : \mathfrak{A}_S \to \mathbb{C}$ such that:*

$$\omega(a^*a) \geq 0 \quad \forall a \in \mathfrak{A}_S, \qquad \omega(\mathbb{I}) = 1\,,$$

that is, positive and normalised to 1.

Remarks 14.1. We can assign $\mathfrak{A}_S$ to S irrespective of any reference frame, provided we assume that (active) transformations of the various frames are given by automorphisms of $\mathfrak{A}_S$. If the algebra depends on the frame, the time evolution *with respect to the given frame* is described by a one-parameter group of *-automorphisms $\{\alpha_t\}_{t\in\mathbb{R}}$, where $\alpha_t : \mathfrak{A}_S \to \mathfrak{A}_S$ is a *-homomorphism for any $t \in \mathbb{R}$, α_0 is the identity and $\alpha_t \circ \alpha_{t'} = \alpha_{t+t'}$ for $t,t' \in \mathbb{R}$. It is natural to demand weak continuity in t: for every state ω on $\mathfrak{A}_s$, the map $\mathbb{R} \ni t \mapsto \omega(\alpha_t(a))$ is continuous for every $a \in \mathfrak{A}_S$.

In case the algebra of observables is independent of the reference frame, $\mathfrak{A}_S$ is actually thought of as a *net of algebras*, i.e. described by a function $\mathscr{O} \mapsto \mathfrak{A}_S(\mathscr{O})$ mapping to an algebra $\mathfrak{A}_S(\mathscr{O})$ any regular bounded region $\mathscr{O}$ in spacetime (stretching both in the space and time directions). From this point of view time evolution is replaced by causal relations between algebras localised at spacetime regions that are causally related, in particular when one belongs to the other's future. The above approach, thoroughly discussed for the first time in the crucial paper of Haag and Kastler [HaKa64], is the modern stepping stone to develop algebraic field theory in the local and covariant formulation. ■

We have to remark that $\mathfrak{A}_S$ is not seen as a concrete C^*-algebra of operators on a given Hilbert space, but remains an abstract C^*-algebra. Physically, $\omega(a)$ is the *expectation value* of the observable $a \in \mathfrak{A}$ in state ω.

There have been a host of attempts to account for assumptions **AA1** and **AA2** in full generality (see the study of [Emc72], [Ara09] and [Str05a, Str11]), and especially the work of I. E. Segal based on so-called *Jordan algebras*). Yet none seems to be definitive [Stre07]. The most evident justification of an algebraic approach lies in its powerfulness [Haa96].

The set of algebraic states on $\mathfrak{A}_S$ is a convex subset in the dual $\mathfrak{A}'_S$ of $\mathfrak{A}_S$: if ω_1 and ω_2 are positive and normalised linear functionals, $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$ is clearly still the same for any $\lambda \in [0,1]$.

Hence, just as we saw for the standard formulation, we can define *pure algebraic states* as extreme elements of the convex body.

Definition 14.2. *An algebraic state* $\omega : \mathfrak{A} \to \mathbb{C}$ *on the* C^**-algebra with unit* $\mathfrak{A}$ *is called a* **pure algebraic state** *if it is extreme in the set of algebraic states. An algebraic state that is not pure is called* **mixed**.

Later we will show that the space of states is not empty and compact in the *-weak topology. Consequently, pure states exist.

Surprisingly, most of the entire abstract apparatus introduced, given by a C^*-algebra and a set of states, admits elementary Hilbert space representations when a reference algebraic state is fixed. This is by virtue of a famous procedure that Gelfand, Najmark and Segal came up with, and that we prepare to present [Haa96, Ara09, Str05a].

Theorem 14.3 (GNS theorem). *Let* $\mathfrak{A}$ *be a* C^**-algebra with unit* $\mathbb{I}$ *and* $\omega : \mathfrak{A} \to \mathbb{C}$ *a positive linear functional with* $\omega(\mathbb{I}) = 1$*. Then*
(a) *there exist a triple* $(\mathsf{H}_\omega, \pi_\omega, \Psi_\omega)$*, where* H_ω *is a Hilbert space,* $\pi_\omega : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_\omega)$ *a* $\mathfrak{A}$*-representation over* H_ω *and* $\Psi_\omega \in \mathsf{H}_\omega$*, such that:*

(i) *Ψ_ω is cyclic for π_ω: $\pi_\omega(\mathfrak{A})\Psi_\omega$ is dense in H_ω;*
(ii) *$(\Psi_\omega|\pi_\omega(a)\Psi_\omega) = \omega(a)$ for every $a \in \mathfrak{A}$.*

(b) *If (H, π, Ψ) satisfies (i) and (ii), there exists a unitary operator $U : \mathsf{H}_\omega \to \mathsf{H}$ such that $\Psi = U\Psi_\omega$ and $\pi(a) = U\pi_\omega(a)U^{-1}$ for any $a \in \mathfrak{A}$.*

Proof. (a) For a start we will build the Hilbert space. We will refer to the elementary theory of Hilbert spaces of Chapter 3. Let us define the quadratic form $\langle x, y\rangle_\omega := \omega(x^*y)$, $x, y \in \mathfrak{A}$. This is a Hermitian semi-inner product by the requests made on ω, so for the seminorm $p_\omega(x) := \sqrt{\langle x, x\rangle_\omega}$ the Schwarz inequality

$$\omega(x^*y) \leq \sqrt{\omega(x^*x)}\,\sqrt{\omega(y^*y)} \tag{14.1}$$

holds for ω. Call **Gelfand ideal** the set $\mathscr{I}_\omega := \{x \in \mathfrak{A} \mid p_\omega(x) = 0\}$. Since p_ω is a seminorm and by (14.1) $\mathscr{I}_\omega$ is a subspace of $\mathfrak{A}$. Actually $\mathscr{I}_\omega$ is a left ideal in $\mathfrak{A}$, i.e. $yx \in \mathscr{I}_\omega$ for any $x \in \mathscr{I}_\omega$, $y \in \mathfrak{A}$. In fact by (14.1):

$$0 \leq p_\omega(yx)^4 = \omega(x^*y^*yx)^2 \leq \omega(y^*yxx^*y^*y)\omega(x^*x) = 0\,.$$

Hence we may define the vector space $\mathscr{D}_\omega := \mathfrak{A}/\mathscr{I}_\omega$, quotient of the vector space $\mathfrak{A}$ by the ideal $\mathscr{I}_\omega$. The elements of $\mathscr{D}_\omega$ are thus cosets $[x]$ for the equivalence relation $\mathfrak{A}$: $x \sim y \Leftrightarrow x - y \in \mathscr{I}_\omega$. The vector space structure is naturally inherited by $\mathfrak{A}$, and makes $\alpha[x] + \beta[y] := [\alpha x + \beta y]$ meaningful, for any $\alpha, \beta \in \mathbb{C}$, $x, y \in \mathfrak{A}$. That $\mathscr{I}_\omega$ is a subspace guarantees the structure is well defined. Since $\mathscr{I}_\omega$ is also the left ideal of zeroes of the seminorm associated to the semi-product $\langle\ ,\ \rangle_\omega$, as we just showed,

$$([x]\,|[y]\,)_\omega := \langle x, y\rangle_\omega \quad \forall x, y \in \mathfrak{A}\,, \tag{14.2}$$

is a well-defined Hermitian inner product on $\mathscr{D}_\omega$. Introduce the Hilbert space H_ω, completion of $\mathscr{D}_\omega$ for said inner product, which we continue to indicate with (14.2) on the entire H_ω. The representation π_ω is defined in the natural way on the dense subspace $\mathscr{D}_\omega = \mathfrak{A}/\mathscr{I}_\omega \subset \mathsf{H}_\omega$ as:

$$(\pi_\omega(a))([b]) := [ab]\,.$$

$\mathscr{D}_\omega$ is by construction invariant under every $\pi_\omega(a)$. At last, let $\Psi_\omega := [\mathbb{I}]$, so that as $a \in \mathfrak{A}$ vary, the set of vectors $\pi_\omega(a)\Psi_\omega = [a]$ fills the space $\mathscr{D}_\omega$, dense in H_ω by construction. Therefore Ψ_ω is cyclic, as needed. It is easy to see that, by construction, $\mathfrak{A} \ni a \mapsto \pi_\omega(a)$ is linear on the dense domain $\mathscr{D}_\omega$ (hence it has an adjoint) and satisfies, for any $a, b, c \in \mathfrak{A}$, $\mu \in \mathbb{C}$:

(i) $\pi_\omega(a)\pi_\omega(b) = \pi_\omega(ab)$;
(ii) $\pi_\omega(a) + \pi_\omega(b) = \pi_\omega(a+b)$;
(iii) $\mu\pi_\omega(a) = \pi_\omega(\mu a)$;
(iv) $(\pi_\omega(b)\Psi_\omega\,|\pi_\omega(a)\pi_\omega(c)\Psi_\omega\,)_\omega = (\pi_\omega(a^*)\pi_\omega(b)\Psi_\omega\,|\pi_\omega(c)\Psi_\omega\,)_\omega$.

The last fact, equivalent to

$$\pi_\omega(a)^*\restriction_{\mathscr{D}_\omega} = \pi_\omega(a^*)\,,$$

follows by

$$(\pi_\omega(b)\Psi_\omega|\pi_\omega(a)\pi_\omega(c)\Psi_\omega)_\omega = ([b]|[ac])_\omega = \omega(b^*ac) = \omega((a^*b)^*c)$$

$$= ([a^*b]|[c])_\omega = (\pi_\omega(a^*)\pi_\omega(b)\Psi_\omega|\pi_\omega(c)\Psi_\omega)_\omega .$$

By construction, for $a \in \mathfrak{A}$:

$$\omega(a) = \omega(\mathbb{I}^* a \mathbb{I}) = ([\mathbb{I}]|[a][\mathbb{I}])_\omega = (\Psi_\omega|\pi_\omega(a)\Psi_\omega)_\omega . \tag{14.3}$$

To finish (a) it is enough to prove that every operator $\pi_\omega(a) : \mathscr{D}_\omega \to \mathsf{H}_\omega$ is bounded, so it extends uniquely to a bounded operator on H_ω, because $\mathscr{D}_\omega \subset \mathsf{H}_\omega$ is dense. We will call the extended operators with the same names $\pi_\omega(a)$. Thus properties (i)–(iv) are still valid, by continuity. In particular, the operators being bounded, (iv) implies $\pi_\omega(a)^* = \pi_\omega(a^*)$, so the map $\pi_\omega : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_\omega)$ is a *-representation.

To prove the boundedness of the $\pi_\omega(a)$, we begin by showing ω is continuous. We will only assume that the linear functional ω is positive, without using $\omega(\mathbb{I}) = 1$. If $h \in \mathfrak{A}$ is normal, since $|\sigma(h)| \leq ||h||$ by the features of the spectral radius, Theorem 8.36(c) gives $\sigma(h \pm ||h||\mathbb{I}) \geq 0$. By Theorem 8.25, $h \pm ||h||\mathbb{I} = c^*c$, so the positivity and linearity of ω allows to say $\omega(h) \pm ||h||\omega(\mathbb{I}) \geq 0$, meaning $|\omega(h)| \leq \omega(\mathbb{I})||h||$. In turn this implies ω is a bounded linear functional. In fact, if $y \in \mathfrak{A}$ is any element, y^*y is self-adjoint and so normal. Using the above result gives immediately $|\omega(y^*y)| \leq \omega(\mathbb{I})||y^*y||$. Finally, (14.1) with $x = \mathbb{I}$ says

$$|\omega(y)|^4 \leq \omega(\mathbb{I})^2 \, ||y^*y||^2 = (\omega(\mathbb{I})||y||)^2 ,$$

hence $||\omega|| \leq \omega(\mathbb{I})$. On the other hand, from $|\omega(\mathbb{I})| = \omega(\mathbb{I})$ and $||\mathbb{I}|| = 1$ we have

$$||\omega|| = \omega(\mathbb{I}) .$$

In our case, as $\omega(\mathbb{I}) = 1$, we obtain $||\omega|| = 1$.

If $\omega(x^*x) > 0$ we can repeat the argument for the linear functional

$$\mathfrak{A} \ni z \mapsto \rho(z) := \frac{\omega(x^*zx)}{\omega(x^*x)} ,$$

by construction linear, positive and such that $\rho(\mathbb{I}) = 1$; therefore $||\rho|| = \rho(\mathbb{I}) = 1$. We conclude that the state ω satisfies

$$\omega(x^*y^*yx) \leq ||y^*y||\omega(x^*x) ,$$

holding also for $\omega(x^*x) = 0$ because the Cauchy-Schwarz inequality forces $0 \leq \omega(x^*y^*yx) = \omega((x^*y^*y)x) \leq \sqrt{\omega((x^*y^*y)^*(x^*y^*y))}\,\sqrt{\omega(x^*x)}$. Consequently

$$||(\pi_\omega(y))([x])||_\omega = ||[yx]||_\omega = \sqrt{\omega(x^*y^*yx)} \leq \sqrt{||y^*y||}\,\sqrt{\omega(x^*x)} \leq ||y||\,||[x]||_\omega ,$$

and so $||\pi_\omega(y)|| \leq ||y||$. This ends part (a).

(b) Just asking $U\pi_\omega(a)\Psi_\omega := \pi(a)\Psi$ for any $a \in \mathfrak{A}$ determines a densely-defined isometric operator, which we called U. This is well defined because, for $\pi_\omega(a)\Psi_\omega = \pi_\omega(a')\Psi_\omega$, $\pi(a)\Psi = \pi(a')\Psi$, for in fact

$$||\pi(a-a')\Psi||^2 = \omega((a-a')^*(a-a')) = ||\pi_\omega(a-a')\Psi_\omega||^2_\omega .$$

U is isometric for the same reason:

$$||U\pi_\omega(a)\Psi_\omega||^2 = ||\pi(a)\Psi||^2 = \omega(a^*a) = ||\pi_\omega(a)\Psi_\omega||^2_\omega .$$

Hence we can extend U to H to a continuous isometric operator with the same name. Similarly, let us construct an isometric operator $V : \mathsf{H} \to \mathsf{H}_\omega$ as the unique continuous extension of $V\pi(a)\Psi = \pi_\omega(a)\Psi_\omega$. By continuity, and using the density of $\pi(\mathfrak{A})\Psi$, follows $UV\Phi = \Phi$ for every $\Phi \in \mathsf{H}$. Therefore U is onto, beside isometric, so unitary. $\Psi = U\Psi_\omega$ and $\pi(a) = U\pi_\omega(a)U^{-1}$, $a \in \mathfrak{A}$, are obvious by construction. □

The GNS theorem shows that given a algebraic state, the observables of $\mathfrak{A}$ are still represented by (bounded) self-adjoint operators on a Hilbert space H_ω, where the expectation value of ω takes the usual form $(\Psi_\omega|\pi(a)\Psi_\omega)$ with respect to a reference vector Ψ_ω. The latter vector allows to recover the whole Hilbert space by means of the representation π_ω itself, as we said in the GNS theorem (a), part (i). The reader should notice that, however, not all algebraic states on $\mathfrak{A}$ are represented by positive trace class operators in H_ω as we shall discuss shortly.

The representation π_ω need not be injective, i.e. faithful. From the proof we see immediately

Proposition 14.4. *The GNS representation $\pi_\omega : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_\omega)$ of the algebraic state ω on the C^*-algebra with unit $\mathfrak{A}$ is faithful if and only if the Gelfand ideal of ω is trivial: $\mathscr{I}_\omega = \{0\}$. Equivalently, $\omega(a^*a) > 0$ for any $a \in \mathfrak{A} \setminus \{0\}$.*

Algebraic states with trivial Gelfand ideal are called **faithful**.

A technical result that was proved in passing, during the proof, and that is useul in itself, is the following.

Theorem 14.5 (Continuity of positive functionals). *If ω is a positive functional on the C^*-algebra $\mathfrak{A}$ with unit $\mathbb{I}$, ω is continuous and $||\omega|| = \omega(\mathbb{I})$.*

There is a useful technical corollary to the GNS theorem that deserves being stated and proved.

Corollary 14.6. *Let ω be an algebraic state on the C^*-algebra $\mathfrak{A}$ with unit, and $(\mathsf{H}_\omega, \pi_\omega, \Psi_\omega)$ an associated GNS triple.*

(a) *If $\psi : \mathfrak{A} \to \mathbb{C}$ is positive and $\psi \leq \omega$ ($\omega - \psi$ is positive), there exists a unique $T \in \mathfrak{B}(\mathsf{H}_\omega)$ such that*

$$\psi(b^*a) = (\pi_\omega(b)\Psi_\omega|T\pi_\omega(a)\Psi_\omega)_\omega \quad \forall a,b \in \mathfrak{A} .$$

Moreover $0 \leq T \leq I$ and $T \in \pi_\omega(\mathfrak{A})'$ (T commutes with each $\pi_\omega(a)$, $a \in \mathfrak{A}$).
(b) *Conversely, if $0 \leq T \leq I$ and $T \in \pi_\omega(\mathfrak{A})'$, then $\psi(a) := (\Psi_\omega|T\pi_\omega(a)\Psi_\omega)_\omega$, for every $a \in \mathfrak{A}$, is a positive functional with $\psi \leq \omega$.*

Proof. (a) Take ψ as in the assumptions. Since

$$|\psi(b^*a)|^2 \leq \psi(b^*b)\psi(a^*a) \leq \omega(b^*b)\omega(a^*a) = ||[b]||_\omega \, ||[a]||_\omega \, ,$$

setting $\psi'([b],[a]) := \psi(b^*a)$, Riesz's theorem warrants the existence of $T \in \mathfrak{B}(\mathsf{H}_\omega)$ with $\psi'([b],[a]) = ([b]|T[a])_\omega$. In other terms $\psi(b^*a) = (\pi_\omega[b]\Psi_\omega|T\pi_\omega(a)\Psi_\omega)$. Furthermore, by construction:

$$([b]\,|(T\pi_\omega(a) - \pi_\omega(a)T)[c]\,)_\omega = \psi(b^*ac) - \psi((a^*b)^*c) = \psi(b^*ac) - \psi(b^*ac) = 0\,.$$

(b) is immediate. □

Remark 14.7. (1) The cyclic vector Ψ_ω is a unit vector, by (a) (ii) in the GNS theorem, since $a = \mathbb{I}$ and $\omega(\mathbb{I}) = 1$.
(2) Irrespective of the way one proves the GNS theorem, the *-representation of C^*-algebras with unit π_ω must be continuous, because of Theorem 8.22, and must also satisfy $||\pi_\omega(a)|| \leq ||a||$ for any $a \in \mathfrak{A}$. In addition, the same theorem implies π_ω is isometric ($||\pi_\omega(a)|| = ||a||$ for any $a \in \mathfrak{A}$) precisely when it is faithful (one-to-one).
(3) We saw in Chapter 7.4.6 that if we restrict to the C^*-algebra $\mathfrak{B}_\infty(\mathsf{H})$ of compact operators on a Hilbert space H, algebraic states on it are exactly the positive operators of trace class with unit trace. Now the C^*-algebra has no unit, because in infinite dimensions the identity operator is never compact. Algebraic states wanted $||\omega|| = 1$ replacing $\omega(\mathbb{I}) = 1$. But these two, for C^*-algebras with unit, are equivalent, by Theorem 14.5. ■

If ω is an algebraic state on $\mathfrak{A}$, every statistical operator on the Hilbert space of a GNS representation of ω – i.e. every positive, trace-class operator with unit trace $T \in \mathfrak{B}_1(\mathsf{H}_\omega)$ – determines an algebraic state

$$\mathfrak{A} \ni a \mapsto tr(T\pi_\omega(a))\ ,$$

evidently. This is true, in particular, for $\Phi \in \mathsf{H}_\omega$ with $||\Phi||_\omega = 1$, in which case the above definition reduces to

$$\mathfrak{A} \ni a \mapsto (\Phi|\pi_\omega(a)\Phi)_\omega\,.$$

To this end we have

Definition 14.8. *If ω is an algebraic state on the C^*-algebra with unit $\mathfrak{A}$, every algebraic state on $\mathfrak{A}$ obtained either from a density operator or a unit vector, in a GNS representation of ω, is said* **normal state** *of ω. Their set $Fol(\omega)$ is the* **folium** *of the algebraic state ω.*

Note that in order to determine $Fol(\omega)$ one can use a fixed GNS representation of ω. In fact, as the GNS representation of ω varies, normal states do not change, as implied by part (b) of the GNS theorem.

The folium of a state ω of the algebra of observables $\mathfrak{A}$ can be naïvely thought of as *the set of algebraic states arising from the action of observables of $\mathfrak{A}$ on ω, possibly through a limiting process.*

By the GNS theorem, namely, every unit vector $\Phi \in \mathsf{H}_\omega$ is the limit of $\pi_\omega(b_n)\Psi_\omega$ as $n \to +\infty$, provided we choose $b_n \in \mathfrak{A}$ suitably. Hence, the GNS theorem implies that the algebraic state associated to Φ, an element of $Fol(\omega)$, can be always computed as

$$\omega_\Phi(a) = (\Phi|\pi_\omega(a)\Phi)_\omega = \lim_{n\to+\infty} \omega(b_n^* a b_n)\,.$$

The other algebraic states in the folium of ω are determined by positive, trace-class operators $T \in \mathfrak{B}(\mathsf{H}_\omega)$ with unit trace. Decomposing T spectrally as infinite convex combination $T = s\text{-}\sum_i p_i(\Phi_i|\ )\Phi_i$, we can eventually write $\omega_T(a) = \sum_i p_i \omega_{\Phi_i}(a)$, and fall back into the previous case.

In case $\mathfrak{A}$ is a von Neumann algebra of operators on H, normal states are as follows (recall that there is already a natural representation of $\mathfrak{A}$, the one over $\mathfrak{A}$ itself).

Definition 14.9. *Looking at a von Neumann algebra $\mathfrak{R} \subset \mathfrak{B}(\mathsf{H})$ on the Hilbert space H as a C^*-algebra, a* **normal state** *of $\mathfrak{R}$ is an algebraic state ω that can be written as $\omega(A) = tr(\rho_\omega A)$ for some positive $\rho_\omega \in \mathfrak{B}_1(\mathsf{H})$ with unit trace, and for every $A \in \mathfrak{B}(\mathsf{H})$.*

It can be proved that any given C^*-algebra with unit always admits states (hence a convex set of states). We will prove this fact within Lemma 14.22. We can ask whether pure states exist, i.e. if the set of states of a C^*-algebra with unit contains extreme elements. The answer is yes, and one proves that every algebraic state can be obtained as a limit of a sequence of a convex combination of pure states, in the *-weak topology.

Theorem 14.10. *The set $S(\mathfrak{A})$ of algebraic states of a C^*-algebra with unit $\mathfrak{A}$ is bounded, and a convex, compact subset of $\mathfrak{A}'$ in the *-weak topology. Moreover $S(\mathfrak{A})$ coincides with the *-weak closure of the convex hull of pure states (which is therefore non-empty).*

Proof. By Theorem 14.5 the convex set $S(\mathfrak{A})$ is contained in the closed unit ball inside the dual of $\mathfrak{A}$. The latter is *-weakly compact, by Theorem 2.76 of Banach–Alaoglu. As the set of states is closed in that topology (the proof is straightforward), it is also compact in the dual of $\mathfrak{A}$ and convex. The Krein–Milman Theorem 2.77 guarantees the set of extreme algebraic states is not empty, and the closure of its convex hull is $S(\mathfrak{A})$. □

14.1.2 Pure states and irreducible representations

We devote this section to an important relationship between pure algebraic states and irreducible representations of the algebra of observables: *ω is pure if and only if the representation π_ω is irreducible*. To prove it we need the following lemma.

Lemma 14.11. *An algebraic state ϕ on a C^*-algebra with unit $\mathfrak{A}$ is pure if and only if $\phi = \psi_1 + \psi_2$ for positive functionals $\psi_i : \mathfrak{A} \to \mathbb{C}$, $i = 1,2$, implies $\psi_i = \lambda_i \phi$ for some $\lambda_1, \lambda_2 \in \mathbb{C}$.*

Proof. If ϕ is not pure, it is not extreme in the set of of algebraic states, so $\phi = \frac{1}{2}\phi_1 + \frac{1}{2}\phi_2$ for $\phi_1 \neq \phi \neq \phi_2$. Defining $\psi_i := \frac{1}{2}\phi_i$, we see that $\phi = \psi_1 + \psi_2$, where $\psi_1 \neq \lambda_1 \phi$ irrespective of λ_1. Let us assume ϕ is pure, conversely. First, if $\lambda \in (0,1)$ and $\phi = \lambda\phi_1 + (1-\lambda)\phi_2$ for some states ϕ_i, then $\phi = \phi_1 = \phi_2$. So assume such ϕ satisfies $\phi = \psi_1 + \psi_2$, for ψ_1, ψ_2 positive functionals. We claim $\psi_i = \lambda_i \phi$ for some numbers λ_i.

If $\psi_i(\mathbb{I}) = 0$ for $i = 1$ or $i = 2$, then $\psi_i = 0$ by Theorem 14.5, and the conclusion would follow trivially. Then we suppose $\psi(\mathbb{I}) \neq 0$, $i = 1,2$. Define $\phi_i(a) := \psi_i(\mathbb{I})^{-1}\psi_i(a)$. Then ϕ_i is a state and $\phi = \lambda\phi_1 + (1-\lambda)\phi_2$, with $\lambda = \psi_1(\mathbb{I})$ and $1-\lambda = \phi(\mathbb{I}) - \psi_1(\mathbb{I}) = \psi_2(\mathbb{I})$. By what we said above $\phi_1 = \phi_2 = \phi$, hence $\psi_i = \psi_i(\mathbb{I})\phi, i = 1,2$. □

Now the announced result can be stated.

Theorem 14.12 (Characterisation of pure algebraic states). *Let ω be an algebraic state on the C^*-algebra with unit $\mathfrak{A}$ and $(\mathsf{H}_\omega, \pi_\omega, \Psi_\omega)$ a corresponding GNS triple. Then ω is pure if and only if π_ω is irreducible.*

Proof. By Schur's lemma (see esp. Remark 11.17), π_ω is irreducible iff $\pi_\omega(\mathfrak{A})' = \{cI\}_{c\in\mathbb{C}}$. By Corollary 14.6, $\pi_\omega(\mathfrak{A})' = \{cI\}_{c\in\mathbb{C}}$ iff $0 \leq \psi \leq \omega$ implies $\psi = c\omega$ for some $c \in \mathbb{C}$. But $0 \leq \psi \leq \omega$ iff $\omega = \psi + (\omega - \psi)$, $\psi \geq 0$ and $\omega - \psi \geq 0$. Thus we conclude π_ω is irreducible iff $\omega = \psi_1 + \psi_2$, $\psi_i \geq 0$, hence $\psi_i = \lambda_i \omega$ for some choice of λ_i. The previous lemma tells π_ω is irreducible iff ω is pure. □

Now we have two important consequences that relate pure states to irreducible representations of a C^*-algebra with unit.

Corollary 14.13. *Let ω be a pure state on the C^*-algebra with unit $\mathfrak{A}$ and $\Phi \in \mathsf{H}_\omega$ a unit vector. Then*

(a) *the functional*

$$\mathfrak{A} \ni a \mapsto (\Phi|\pi_\omega(a)\Phi)_\omega ,$$

defines a pure algebraic state and $(\mathsf{H}_\omega, \pi_\omega, \Phi)$ is a GNS triple for it. In that case, GNS representations of algebraic states given by non-zero vectors in H_ω are all unitarily equivalent.

(b) *Unit vectors $\Phi, \Phi' \in \mathsf{H}_\omega$ give the same (pure) algebraic state if and only if $\Phi = c\Phi'$ for some $c \in \mathbb{C}$, $|c| = 1$, i.e. if and only if Φ and Φ' belong to the same ray.*

Proof. (a) Consider the closed space $\mathsf{M}_\Phi := \overline{\pi_\omega(\mathfrak{A})\Phi}$; we will show it coincides with H_ω. By construction $\pi(a)\mathsf{M}_\Phi \subset \mathsf{M}_\Phi$ for $a \in \mathfrak{A}$, so M_Φ is closed and π_ω-invariant. As the representation is irreducible, necessarily $\mathsf{M}_\Phi = \mathsf{H}_\omega$ or $\mathsf{M}_\Phi = \{0\}$. The latter case is impossible because $\pi_\omega(\mathbb{I})\Phi = \Phi \neq 0$. Now the claim is clear by construction, because $(\mathsf{H}_\omega, \pi_\omega, \Phi)$ satisfies the GNS assumptions for a triple of an algebraic state given by Φ as above, which is pure because the GNS representation is irreducible. The last statement is obvious since all GNS representations can be constructed as above. The unitary transformation between two such is always the identity operator.

(b) If $\Phi = c\Phi'$ the two vectors give the same pure algebraic state. If, conversely, two unit vectors determine the same pure algebraic state, i.e. $(\Phi|\pi_\omega(a)\Phi)_\omega = (\Phi'|\pi_\omega(a)\Phi')_\omega$ for every $a \in \mathfrak{A}$, then we decompose $\Phi = c\Phi' + \Psi$ with Ψ orthogonal to Φ'. In this way

$$(\Phi|\pi_\omega(a)\Phi)_\omega = |c|^2(\Phi'|\pi_\omega(a)\Phi')_\omega + c(\Psi|\pi_\omega(a)\Phi')_\omega + \overline{c}(\Phi'|\pi_\omega(a)\Psi)_\omega\,,$$

whence

$$(1-|c|^2)(\Phi'|\pi_\omega(a)\Phi')_\omega = c(\Psi|\pi_\omega(a)\Phi')_\omega + \overline{c}(\Phi'|\pi_\omega(a)\Psi)_\omega\,.$$

Choose $a = \mathbb{I}$, so that $|c| = 1$. Back to $\Phi = c\Phi' + \Psi$, we obtain $\Psi = 0$ because $1 = ||\Phi'||^2 = |c|^2 + ||\Psi||^2$. □

Corollary 14.14. *If $\mathfrak{A}$ is a C^*-algebra with unit, every irreducible representation $\pi : \mathfrak{A} \to \mathsf{H}$ is a GNS representation of a state pure.*

Proof. Let $\Psi \in \mathsf{H}$ be a unit vector. As the representation is irreducible, $\pi(\mathfrak{A})\Psi$ is dense in H. It is easy to see that (H, π, Ψ) is a GNS triple for $\omega(\cdot) = (\Psi|\pi(\cdot)\Psi)$. The latter state is pure by irreducibility. □

Examples 14.15. **(1)** For commutative C^*-algebras with unit the following characterisation of pure states holds.

Proposition 14.16. *If $\mathfrak{A}$ is a commutative C^*-algebra with unit, a state $\omega : \mathfrak{A} \to \mathbb{C}$ is pure if and only if it is multiplicative: $\omega(ab) = \omega(a)\omega(b)$ for any $a, b \in \mathfrak{A}$.*

Proof. ω pure implies π_ω irreducible, but $\pi_\omega(a)$ commutes with every other $\pi_\omega(b)$ since $\mathfrak{A}$ is commutative. By Schur's lemma $\pi_\omega(\mathfrak{A}) = \{cI \,|\, c \in \mathbb{C}\}$. Using the GNS theorem gives $\omega(ab) = \omega(a)\omega(b)$. Conversely if ω is multiplicative, by the GNS theorem we can write $(\pi_\omega(a^*)\Psi_\omega|\pi_\omega(b)\Psi_\omega)_\omega = (\pi_\omega(a^*)\Psi_\omega|\Psi_\omega)_\omega(\Psi_\omega|\pi_\omega(b)\Psi_\omega)_\omega$, so Ψ_ω, alone, is a basis of H_ω, because $\pi_\omega(\mathfrak{A})\Psi_\omega$ is dense in H_ω. Therefore H_ω has dimension 1, and all its operators are numbers; in particular, $\pi_\omega(\mathfrak{A})' = \{cI \,|\, c \in \mathbb{C}\}$, which means π_ω is irreducible by Schur's lemma. □

(2) The next example does not originate in QM. Take a compact Hausdorff space X and the commutative C^*-algebra with unit $C(\mathsf{X})$ of $\mathbb{C}$-valued continuous maps on X, equipped with the usual pointwise algebraic operations, involution given by complex conjugation and norm $||\;||_\infty$. If μ denotes a Borel probability measure on X, then

$$\omega_\mu : C(\mathsf{X}) \ni f \mapsto \int_{\mathsf{X}} f d\mu$$

defines an algebraic state on $C(\mathsf{X})$. The GNS theorem then gives a triple $(\mathsf{H}_\mu, \pi_\mu, \Psi_\mu)$ where: $\mathsf{H}_\mu = L^2(\mathsf{X}, d\mu)$, $(\pi_\omega(f)\psi)(x) := f(x)\psi(x)$ for every $x \in \mathsf{X}$, $\psi \in \mathsf{H}_\mu$ and $f \in C(\mathsf{X})$. The cyclic vector Ψ_ω coincides with the constant map 1 on X.

It can be checked that pure states are the Dirac measures δ_x concentrated at points $x \in \mathsf{X}$. In this sense probability measures can be understood as "thick" points. ■

Remarks 14.17. Consider, in the standard (not algebraic) formulation, a physical system S described on the Hilbert space H_S and a mixed state $\rho \in \mathfrak{S}(\mathsf{H})$. The map $\omega_\rho : \mathfrak{B}(\mathsf{H}) \ni A \mapsto tr(\rho A)$ defines an algebraic state on the C^*-algebra $\mathfrak{B}(\mathsf{H}_S)$. By the GNS theorem, there exist another Hilbert space H_ρ, a representation $\pi_\rho : \mathfrak{B}(\mathsf{H}_S) \to \mathfrak{B}(\mathsf{H}_\rho)$ an a unit vector $\Psi_\rho \in \mathsf{H}_\rho$ such that

$$tr(\rho A) = (\Psi_\rho | \pi_\rho(A) \Psi_\rho)$$

for $A \in \mathfrak{B}(\mathsf{H}_S)$. Thus it seems that the initial mixed state has been transformed into a pure state! How is this fact explained?

The answer follows from Theorem 14.12: Ψ_ρ does not correspond to any vector $U^{-1}\Psi_\rho$ in H_S under a unitary transformation $U : \mathsf{H}_S \to \mathsf{H}_\rho$ with $UAU^{-1} = \pi_\rho(A)$. In fact the representation $\mathfrak{B}(\mathsf{H}_S) \ni A \mapsto A \in \mathfrak{B}(\mathsf{H}_S)$ is irreducible, whereas π_ρ cannot be irreducible because the state of ρ is not an extreme point in the space of non-algebraic states, and so it cannot be extreme in the larger space of algebraic states.

This example should clarify that the correspondence pure (algebraic) states vs. state vectors, automatic in the standard formulation, holds in Hilbert spaces of GNS representations of pure algebraic states, but in general not for mixed algebraic states.

■

14.1.3 Hilbert space formulation vs algebraic formulation

Withholding the point of view adopted up to Chapter 13 included, in which one starts from a given Hilbert space H_S, the C^*-algebra $\mathfrak{A}_S$ of observables associated to a system S can be, in the limit situation, the whole space of bounded operators $\mathfrak{B}(\mathsf{H}_S)$. A choice that makes more physical sense is to define the algebra of observables as a C^*-subalgebra with unit in $\mathfrak{B}(\mathsf{H}_S)$, typically having the structure of a *von Neumann algebra* $\mathfrak{R}_S$ (Example 3.44(3)), generated by the PVMs of the system's observables (in the sense of Remark 3.41(2)). Making this a von Neumann algebra implies strong closure, thus allowing to integrate spectral meaures, at least in bounded measurable functions, and still obtain elements of the algebra. Taking *bounded* operators is no major restriction from the physics' point of view. Any observable A represented by an unbounded self-adjoint operator, namely, is physically the same as the sequence of observables represented by bounded self-adjoint operators $A_n := \int_{(-n,n]} \lambda dP^{(A)}(\lambda)$, $n = 1, 2, \ldots$. For the time being we will assume $\mathfrak{A}_S = \mathfrak{R}_S = \mathfrak{B}(\mathsf{H}_S)$, and later return to the general case, when superselection rules are turned on.

Clearly every state $\rho \in \mathfrak{S}(\mathsf{H}_S)$ determines a (normal) algebraic state on the C^*-algebra $\mathfrak{B}(\mathsf{H}_S)$ by setting $\omega_\rho(A) := tr(\rho A), A \in \mathfrak{B}(\mathsf{H}_S)$. From what we said, a state in $\mathfrak{S}(\mathsf{H}_S)$ is pure iff it is algebraically pure in the C^*-algebra $\mathfrak{B}(\mathsf{H}_S)$. The set of algebraic states on $\mathfrak{B}(\mathsf{H}_S)$ coming from positive trace-class operators with unit trace does not exhaust all algebraic states on $\mathfrak{B}(\mathsf{H}_S)$, but only a small part of them.

Nevertheless, viewing the C^*-algebra of observables as a specific C^*-algebra of operators on a Hilbert space (possibly the entire algebra of bounded operators) in the general framework of the algebraic formulation would be a backslide in the theory, for it would lead to assume theoretically the existence of a privileged Hilbert space where states are described. This would rule out, for systems with infinitely many degrees, a host of states corresponding to non-unitarily equivalent representations, which do exist and have a meaning.

In the general case observables are therefore taken to form an abstract C^*-algebra $\mathfrak{A}$; the Hilbert space representation is fixed only *after* a state ω has been given, and is the Hilbert space H_ω of the GNS construction. At that point, in the Hilbert space the C^*-algebra may be enlarged to a von Neumann algebra (still C^*), simply by taking the $\pi_\omega(\mathfrak{A})''$ generated by $\pi_\omega(\mathfrak{A})$. Notice that as $\pi_\omega(\mathfrak{A})''$ is closed in the weak, strong and uniform topologies, there are elements in $\pi_\omega(\mathfrak{A})''$ that are *not* limits in $\pi(\mathfrak{A})$ in the uniform topology (coinciding with the topology of $\mathfrak{A}$ under π_ω). These elements do not correspond to elements of $\mathfrak{A}$, and cannot be considered, in this sense, "true observables of the system", independent of the choice of state. In particular elementary propositions like: "the reading of a falls in the Borel set E" are not usually thinkable as elements of $\mathfrak{A}$, i.e. observables. These should correspond to maps $\chi_E(a)$, where the function of the self-adjoint a is defined via continuous functional calculus under the representation $\Phi_a : C(\sigma(a)) \to \mathfrak{A}$ of Theorem 8.36. But $\chi_E \notin C(\sigma(a))$ in general. We can make sense of these observables only *after* having fixed a state, working in its *GNS* representation. At this juncture the abstract formulation appears to part evidently from the elementary formulation, which is based on a preexisting Hilbert space and the fundamental nature of elemenary propositions about observable readings.

The process of reduction of the state, that follows the outcome of a measurement, should be treated likewise. Take an observable $a \in \mathfrak{A}$, suppose the system is in the pre-measurement state ω, and let the (ideal) reading of a fall in the Borel set E. After the measurement the state is

$$\omega_E : \mathfrak{A} \ni b \mapsto \frac{(\Psi_\omega | P_E \pi_\omega(b) P_E \Psi_\omega)}{(P_E \Psi_\omega | P_E \Psi_\omega)} ,$$

where P_E is the PVM element of the self-adjoint operator $\pi_\omega(a)$ corresponding to the Borel set E.

It would, actually, be possible to narrow down the gap between the two formulations in the following manner. From Chapter 8 we know that the integral of a bounded, measurable map in a PVM on the Hilbert space H is defined using the uniform topology, which is the natural topology of the C^*-algebra $\mathfrak{B}(\mathsf{H})$. Hence one could always ask the C^*-algebra $\mathfrak{A}$ of observables of a physical system be generated by the $p \in \mathfrak{A}$ that have the same features of orthogonal projectors in Hilbert spaces: $p = pp$ and $p^* = p$. These elements correspond to orthogonal projectors in the Hilbert space of any GNS representation of $\mathfrak{A}$. Therefore one could choose the elements p, using GNS representations of physically meaningful states, so to obtain the PVMs of the relevant observables, whence also the (bounded measurable) maps of those observables.

In a general setup it is further reasonable to suppose the C^*-algebra $\mathfrak{A}$ whose self-adjoint elements represent observables is *simple*.

Definition 14.18. *A C^*-algebra $\mathfrak{A}$ is* **simple** *if its only closed two-sided ideals that are invariant under the involution are $\mathfrak{A}$ and $\{0\}$.*

The reason for wanting simple algebras is that then every non-trivial representation (whether GNS or not), on whichever Hilbert space, is faithful, i.e. injective hence isometric, as the next proposition proves.

Proposition 14.19. *If* $\mathfrak{A}$ *is a simple* C^**-algebra with unit and* $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ *a non-zero representation on the Hilbert space* H*, then* π *is faithful (one-to-one) and isometric.*

Proof. The null space of $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ is a two-sided ideal in $\mathfrak{A}$ that is closed (π is continuous by Theorem 8.22) and invariant under the involution, as is immediate to see. As the trivial representation maps everything to the zero operator, our representation must be injective because $Ker(\pi) = \{0\}$, and so isometric by Theorem 8.22. □

This means every operator representation of a simple C^*-algebra with unit faithfully represents the algebra, quite literally.

Sometimes the C^*-structure is too rigid, whereas a *-algebra with unit is better tailored to described observables. This is the case when one studies Bosonic quantum fields without using Weyl C^*-algebras. The key part of the GNS theorem is still valid. In fact, we have the following version of the GNS theorem, whose proof is an easy consequence of the above.

Theorem 14.20 (GNS theorem for *-algebras with unit). *Let* $\mathfrak{A}$ *be a* **-algebra with unit* $\mathbb{I}$ *and* $\omega : \mathfrak{A} \to \mathbb{C}$ *a positive linear functional with* $\omega(\mathbb{I}) = 1$*. Then*

(a) *there exists a quadruple* $(\mathsf{H}_\omega, \mathscr{D}_\omega, \pi_\omega, \Psi_\omega)$ *made of a Hilbert space* H_ω*, a subspace* $\mathscr{D}_\omega \subset \mathsf{H}_\omega$*, a linear map* $\pi_\omega : \mathfrak{A} \to \mathfrak{L}(\mathscr{D}_\omega, \mathsf{H}_\omega)$ *and an element* $\Psi_\omega \in \mathscr{D}_\omega$*, such that:*

(i) $\mathscr{D}_\omega$ *is* $\pi_\omega(a)$*-invariant for every* $a \in \mathfrak{A}$*, since* $\mathscr{D}_\omega = \pi_\omega(\mathfrak{A})\Psi_\omega$ *;*
(ii) Ψ_ω *is cyclic for* π_ω*:* $\mathscr{D}_\omega$ *is dense in* H_ω *;*
(iii) $\pi_\omega : \mathfrak{A} \to \pi_\omega(\mathfrak{A})$ *is an algebra homomorphism satisfying:* $\pi_\omega(\mathbb{I}) = I$ *and* $\pi_\omega(a^*) = \pi_\omega(a)^* \restriction_{\mathscr{D}_\omega}$*,* $a \in \mathfrak{A}$*;*
(iv) $(\Psi_\omega | \pi(a)\Psi_\omega) = \omega(a)$*,* $a \in \mathfrak{A}$*.*

(b) *If* $(\mathsf{H}, \mathscr{D}, \pi, \Psi)$ *fulfills (i)–(iv), there exists a unitary operator* $U : \mathsf{H}_\omega \to \mathsf{H}$ *such that* $\Psi = U\Psi_\omega$*,* $\mathscr{D} = U\mathscr{D}_\omega$ *and* $\pi(a) = U\pi_\omega(a)U^{-1}$ *for any* $a \in \mathfrak{A}$*.*

Now the function π_ω is not (necessarily) continuous. The operators $\pi_\omega(a)$ do not belong in $\mathfrak{B}(\mathsf{H}_\omega)$, in general. Every operator $\pi_\omega(a)$ is *closable* in $\mathscr{D}_\omega$ by Theorem 5.10(b), since the domain contains the dense subspace $\mathscr{D}_\omega$, making $\pi_\omega(a)^*$ densely defined.

A general quantum theory, formulated algebraically, seeks to find, among the immense collection of algebraic states on a C^*-algebra of observables in a given physical system, those states with some meaning. We refer to the aforementioned readings for an in-depth study of such a wide-ranging topic. We shall return to this point at the end of the next section, although usually it is physics that suggests the choice of some privileged state ω. For instance, the referene state of quantum field theories in the lack of gravity (Minkowski's flat spacetime) and without interactions is the so-called *vacuum state*, which corresponds to the absence of particles associated to the field in question, and is invariant under the Poincaré group. The picture changes abruptly when "turning on gravity", i.e. introducing curvature on the spacetime: the absence of the Poincaré symmetry, in general, does not allow to select one's favourite (algebraic) state uniquely, but rather an entire class of states, most of the time what are known as *Hadamard states* [Wald94]. These enable to make sense of renormalisation, and also define important observables such as the energy-momentum tensor (cf. [Mor03], for example).

14.1.4 Superselection rules and Fell's theorem

We want to emphasise how the algebraic formulation permits to handle situations – necessary on physical grounds, as we said – in which *non-unitarily equivalent* representations of the same algebra $\mathfrak{A}_S$ of observables of a given system S coexist. Such representations are associated to pair of distinct algebraic states giving inequivalent GNS representations.

Recall that given a C^*-algebra with unit $\mathfrak{A}$ representing the observables, pure states determine every irreducible representation (all GNS representations, as we saw). We may decompose the set of pure states, i.e. of irreducible representations, in equivalence classes under the relation:

$$\omega_1 \sim \omega_2 \quad \text{if and only if} \quad \pi_{\omega_1} \simeq \pi_{\omega_2} .$$

These classes have a meaning in relationship to *superselection rules* (see Chapter 11.1, 7.4.5), as Haag noticed.

To go into the matter we need to take a step back. We return to the standard formulation in the Hilbert space, though revisited under the algebraic light, and consider a quantum theory that admits superselection rules: these require an observable Q (like the electric charge) to be always defined, with arbitrary value q, on pure normal states. We will assume for a moment that the possible values are countable, so to have closed, pairwise orthogonal coherent sectors H_{Sq_i} in the separable Hilbert space H_S. The H_{Sq_i} are the eigenspaces of Q with eigenvalues q_i. The algebra of (bounded) observables (see Remarks 7.47, 11.2) is the von Neumann algebra $\mathfrak{A}_S := \mathfrak{R}_S$ generated by the orthogonal projectors in $\mathfrak{P}(\mathsf{H}_S)$ (hence all bounded operators) that commute with the projectors P_{q_k} onto the H_{Sq_k}. In other words $\mathfrak{R}_S := (\{P_{q_k} \,|\, k \in \mathbb{N}\}')'' = \{P_{q_k} \,|\, k \in \mathbb{N}\}'$. Clearly $\mathfrak{R}_S$ has a non-trivial centre, that contains P_{q_k}, and states are now viewed as normal algebraic states on $\mathfrak{R}_S$. Therefore each coherent sector is invariant under every physically admissible observables, and on every sector there will be a representation of the observable algebra obtained by restricting observables to the closed invariant space. Every value q_k that Q can take gives a coherent sector H_{Sq_k}. Distinct choices of q_k produce *unitarily inequivalent* representations. In fact, if $q_1 \neq q_2$ on H_{Sq_1} and H_{Sq_2}, Q is represented by different multiples of the identity $q_1 I$ and $q_2 I$; thus $Uq_1IU^{-1} = q_1 I \neq q_2 I$ whichever unitary $U : \mathsf{H}_{Sq_1} \to \mathsf{H}_{Sq_2}$ we take. If more than one superselection rule is activated, Wightman [Wigh95] conjectured[1] that the rules are associated to pairwise compatible observables Q_j in the centre of $\mathfrak{R}_S$. Assuming Wightman is right, the Hilbert space splits in an orthogonal sum of coherent sectors common to all superselecion rules. On each sector all charges of the superselection rules are defined simultaneously, and on the single sector we have a representation of all the observables of the system. These representations are mutually unitarily inequivalent, and they are also taken to be *irreducible representations* of the observable algebra in case they describe the physical system completely. In the standard description we have used that the Hilbert space $\mathsf{H}_S = \oplus_{k\in\mathbb{N}}\mathsf{H}_{Sk}$ is somehow "too big", because not all vectors $\psi \in \mathsf{H}_S$ define physical pure (normal, algebraic) states. Only those belonging to a single sector H_{Sk} do so (this is easily proved), and every irreducible algebra representation is in one coherent sector only. Furthermore, there exist several representations for the same normal state. The vector $\psi = \sum_{k\in\mathbb{N}} \psi_k \in \mathsf{H}_S$, with $\psi_k \in \mathsf{H}_{Sk}$, carries the same amount of information of $\psi' = \sum_{k\in\mathbb{N}} e^{i\alpha_k}\psi_k$, with $\alpha_k \in \mathbb{R}$ fixed arbitrarily, and of the incoherent superposition $\rho_\psi = \sum_{i\in\mathbb{N}} \psi_i(\psi_i|\cdot)$. In fact: $(\psi|P\psi) = (\psi'|P\psi') = tr(\rho_\psi P)$ for every $P \in \mathfrak{P}(\mathsf{H}_S) \cap \mathfrak{R}_S$ (P commutes with coherent projectors) and the identity extends to $(\psi|A\psi) = (\psi'|A\psi') = tr(\rho_\psi A)$, for $A \in \mathfrak{R}_S$, by the spectral theorem. Each normal state, seen as statistical operator, decomposes as convex combination (in general infinite, with respect to the uniform topology) of pure normal states, by the spectral decomposition theorem for compact operators. For the sake of completeness we note that we could consider the situation where the centre of $\mathfrak{R}_S$ contains observables with continuous spectrum, and then one would speak of *continuous superselection rules* [Giu00]. If so, things get more complicated because H_S is no longer an orthogonal sum of coherent subpaces (a *direct integral* is necessary). Still, by the Krein–Milman Theorem 2.77, we can write normal states of $\mathfrak{R}_S$ as combinations (infinite, in the *-weak topology) of extreme algebraic states. That said, though, these extreme elements are usually not normal but just algebraic states, if we view $\mathfrak{R}_S$ as a C^*-algebra. In this sense – concerning the notion of state – continuous superselection rules lead naturally to the algebraic formulation.

Let us start from the algebraic formulation, based on a C^*-algebra of observables and the general notion of algebraic state – thus freeing ourselves from Hilbert spaces and von Neumann algebras as the characterising structures of a physical system. The picture now is suddenly more straighforward, for the use of C^*-algebras eschews convoluted argumentations and technical complications. Extending Wightman's assumption, in the algebraic formalism, superselection rules are accounted for by observables Q in the **centre** of the C^*-algebra $\mathfrak{A}_S$ of observables, i.e. the subalgebra of elements commuting with all of $\mathfrak{A}_S$. Every pure algebraic state ω, corresponding to an irreducible representation of the algebra of observables, must inevitably select a

[1] In *non-Abelian* gauge theories, like quantum chromodynamics, every admissible state must be invariant under a unitary representation of the gauge group. Self-adjoint operators generating the Lie algebra cannot be simultaneously defined on admissible states, because they do not commute with one another by hypothesis. Then self-adjoint operators do *not* represent observables, and the selection of admissible states is not due to Wightman's superselection rule.

value of Q in the GNS representation by Schur's lemma, as $\pi_\omega(Q)$ commutes with all elements. That is to say, $\pi_\omega(Q) = qI$ for some $q \in \mathbb{R}$ (now the values may be uncountable, since the separable Hilbert space is not unique). Exactly as before, two pure algebraic states ω, ω' with distinct $q \neq q'$ produce inequivalent GNS representations, so there is no unitary operator $U : \mathsf{H}_\omega \to \mathsf{H}_{\omega'}$ such that $U\pi_\omega(a)U^{-1} = \pi_{\omega'}(a)$ for each $a \in \mathfrak{A}_S$ (this identity is false for $a = Q$).

In general we expect that families of non-equivalent pure states (i.e. of inequivalent irreducible representations) can be labelled by distinct values of a charge of sorts, corresponding to a central observable. Eventually, the existence of superselection charges might be the reason for the existence of inequivalent irreducible representations of the C^*-algebra of observables $\mathfrak{A}_S$. It is worth observing the new superselection rules can anyway show up in a specific GNS representation of $\mathfrak{A}_S$ associated to a state ω, in case we think the algebra, in such representation, as the von Neumann algebra $\pi_\omega(\mathfrak{A}_S)''$: this is larger than $\pi_\omega(\mathfrak{A}_S)$, so in general it has a non-trivial centre even if $\mathfrak{A}_S$ does not. (See [Prim00] for this point, in particular concerning the interpretation of central observables of $\pi_\omega(\mathfrak{A}_S)''$ as classical observables.)

We suggest to consult [Haa96] to find an exhaustive treatise on superselection rules in the algebraic formalism. Let us just make one general comment. We saw how the space of pure states decomposes in disjoint families of states giving inequivalent representations, and the states of a same family can be viewed as state vectors on one Hilbert space. So we would like to know, given a pure state ω, if it is possible, experimentally speaking, to say which family it belongs to. The answer is not simple, as shown by a theorem proved by Fell: in the case of pure states on a C^*-algebra with unit, this says that pure states in a given family are dense in the set of all pure states for the *-weak topology. Let us explain why this abstract fact is relevant. In the real world we can conduct only a finite (arbitrarily large) number of experiments. Suppose we can measure N observables $a_1, a_2, \ldots, a_N$. The accuracy is finite, so the true value α_i of the reading $\omega(a_i)$ of a_i is given up to $\varepsilon_i > 0$:

$$|\omega(a_i) - \alpha_i| < \varepsilon_i\,, \quad i = 1, 2, \ldots, N\,.$$

Now observe that the numbers α_i and ε_i determine a neighbourhood in the space of states with respect to the *-weak topology. Fell's result implies that it is not possible to establish to which family a given pure state ω belongs using an arbitrarily large, but finite, number of measurements with arbitrarily small, yet finite, errors.

One way to simplify the problem [Haa96] is to choose *a priori* a family of pure states with some *ad hoc* criterion. Supposing, for instance, that the algebra is spatially localisable, we may assume that outside a certain region the physical system is absent. Then all states of interest are those that outside a given and arbitrarily large region resemble the vacuum state, when we measure on it observables localised outside the region.

14.1.5 Proof of the Gelfand–Najmark theorem, universal representations and quasi-equivalent representations

The GNS construction has a purely mathematical consequence known as *Gelfand–Najmark theorem* (stated in Chapter 8), that proves every C^*-algebra with unit can be realised as a C^*-algebra of operators on a Hilbert space, even if not uniquely. To prove the result we need a few technical lemmas.

Lemma 14.21. *Let $\mathfrak{A}$ be a C^*-algebra with unit $\mathbb{I}$. Any bouned linear functional $\phi : \mathfrak{A} \to \mathbb{C}$ with $\phi(\mathbb{I}) = ||\phi||$ is positive.*

Proof. We will make use of Theorem 8.25, and, a usual, $r(c)$ will denote the spectral radius of c. Without loss of generality we assume $\phi(\mathbb{I}) = 1$. Let $a \in \mathfrak{A}$ be positive and set $\phi(a) = \alpha + i\beta$, with $\alpha, \beta \in \mathbb{R}$. We have to show $\alpha \geq 0$ and $\beta = 0$. For small $s \geq 0$: $\sigma(\mathbb{I} - sa) = \{1 - st \,|\, t \in \sigma(a)\} \subset [0,1]$, since $\sigma(a) \subset [0, +\infty)$; hence $||\mathbb{I} - sa|| = r(\mathbb{I} - sa) \leq 1$. Therefore $1 - s\alpha \leq |1 - s(\alpha + i\beta)| = |\phi(\mathbb{I} - sa)| \leq 1$, so $\alpha \geq 0$. Now define $\beta_n := a - \alpha\mathbb{I} + in\beta\mathbb{I}$, $n = 1, 2, \ldots$. Then

$$||b_n||^2 = ||b_n^* b_n|| = ||(a - \alpha\mathbb{I})^2 + n^2\beta^2\mathbb{I}|| \leq ||a - \alpha\mathbb{I}||^2 + n^2\beta^2 .$$

Consequently

$$(n^2 + 2n + 1)\beta^2 = |\phi(b_n)|^2 \leq ||a - \alpha\mathbb{I}||^2 + n^2\beta^2 \quad n = 1, 2, \ldots$$

and then $\beta = 0$. □

Lemma 14.22. *Let $\mathfrak{A}$ be a C^*-algebra with unit and $a \in \mathfrak{A}$.*

(a) *If $\alpha \in \sigma(a)$ there exists a state $\phi : \mathfrak{A} \to \mathbb{C}$ such that $\phi(a) = \alpha$.*
(b) *If $a \neq 0$, there exists a state $\phi : \mathfrak{A} \to \mathbb{C}$ with $\phi(a) \neq 0$.*
(c) *If $a = a^*$, there exists a state $\phi : \mathfrak{A} \to \mathbb{C}$ such that $|\phi(a)| = ||a||$.*

Proof. (a) For any complex numbers β, γ we have $\alpha\beta + \gamma \in \sigma(\beta a + \gamma\mathbb{I})$, so $|\alpha\beta + \gamma| \leq ||\beta a + \gamma\mathbb{I}||$. Hence asking $\phi(\beta a + \gamma\mathbb{I}) := \alpha\beta + \gamma$ defines (unambiguously) a linear functional on the subspace $\{\beta a + \gamma\mathbb{I} \,|\, \beta, \gamma \in \mathbb{C}\}$ such that $\phi(a) = \alpha$, $\phi(\mathbb{I}) = 1$ and $||\phi|| = 1$. By a corollary to the Hahn–Banach theorem, we can extend ϕ to a continuous linear functional on $\mathfrak{A}$ satisfying $||\phi|| = \phi(\mathbb{I}) = 1$. The previous lemma guarantees that the functional is a state on $\mathfrak{A}$ with $\phi(a) = \alpha$.
(b) If $a = a^*$ and $a \neq 0$, then $\sigma(a) \neq \{0\}$, for otherwise the properties of the spectral radius of self-adjoint elements would imply $||a|| = r(a) = 0$. Then the state ϕ of part (a) satisfies $\phi(a) \neq 0$ for $\alpha \in \sigma(a) \setminus \{0\}$. Consider when $a \neq a^*$, $a \neq 0$. Then we can decompose $a = b + ic$ with $b = b^*$, $c = c^*$. If $\phi(a) = 0$ for any state $\phi : \mathfrak{A} \to \mathbb{C}$, we would have $0 = \phi(a) = \phi(b) + i\phi(c)$ for any ϕ. But the GNS theorem implies $\phi(d) = \overline{\phi(d)}$ for $d = d^*$. Hence $\phi(b) = \phi(c) = 0$ for any ϕ. Since c and d are self-adjoint, the proof's starting argument forces $b = c = 0$ so $a = 0$. As this was excluded, there must exist a state with $\phi(a) \neq 0$.
(c) In the case examined, since $||a|| = \sup\{|\lambda| \,|\, \lambda \in \sigma(a)\}$ and $\sigma(a)$ is compact in $\mathbb{R}$, there must be an element $\Lambda \in \sigma(a)$ with $|\Lambda| = ||a||$. Using part (a) with $\alpha = \Lambda$ proves the claim. □

Now we are ready to state and prove the Gelfand–Najmark theorem.

Theorem 14.23 (Gelfand–Najmark). *For any C^*-algebra with unit $\mathfrak{A}$ there exist a Hilbert space H and an (isometric) *-isomorphism $\phi : \mathfrak{A} \to \mathfrak{B}$, where $\mathfrak{B} \subset \mathfrak{B}(\mathsf{H})$ is a C^*-subalgebra of $\mathfrak{B}(\mathsf{H})$.*

Proof. For every $x \in \mathfrak{A} \setminus \{0\}$ let us fix a state $\phi_x : \mathfrak{A} \to \mathbb{C}$ with $\phi_x(x) \neq 0$. This state exists by part (b) of the above lemma. Consider the collection of GNS triples $(\mathsf{H}_x, \pi_x, \Psi_x)$ associated to each ϕ_x, and the Hilbert sum

$$\mathsf{H} := \bigoplus_{x \in \mathfrak{A}\setminus\{0\}} \mathsf{H}_x .$$

In this way the elements of H are $\psi = \oplus_{x\in\mathfrak{A}\setminus\{0\}} \psi_x := \{\psi_x\}_{x\in\mathfrak{A}\setminus\{0\}}$ such that:

$$\sum_{x\in\mathfrak{A}\setminus\{0\}} ||\psi_x||_x^2 < +\infty . \tag{14.4}$$

On H we have an inner product making it a Hilbert space:

$$(\psi|\psi') = \sum_{x\in\mathfrak{A}\setminus\{0\}} (\psi_x|\psi'_x)_x .$$

Define the map $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ by imposing:

$$\pi(0) := 0 \quad \text{and} \quad (\pi(a)\psi)_x := \pi_x(a)\psi_x \quad \text{for } \psi \in \mathsf{H},\, a \in \mathfrak{A} \setminus \{0\} .$$

It is not hard to see π is a *-homomorphism of C^*-algebras with unit mapping $\mathfrak{A}$ to $\mathfrak{B}(\mathsf{H})$. In particular, $||\pi(a)|| \leq ||a||$, as prescribed by Theorem 8.22. In fact if (14.4) holds, since Theorem 8.22 gives $||\pi_x(a)|| \leq ||a||$, we obtain

$$||\pi(a)\psi||^2 = \sum_{x\in\mathfrak{A}\setminus\{0\}} ||\pi_x(a)\psi_x||_x^2 \leq ||a||^2 \sum_{x\in\mathfrak{A}\setminus\{0\}} ||\psi_x||_x^2 = ||a||^2||\psi||^2 < +\infty .$$

To end the proof it suffices to show π is isometric. By Theorem 8.22(a) that is equivalent to injectivity. Suppose $\pi(a) = 0$, so $\pi_x(a)\psi_x = 0$ for any $x \in \mathfrak{A} \setminus \{0\}$, $\psi_x \in \mathsf{H}_x$. In particular $\phi_x(a) = (\Psi_x|\pi_x(a)\Psi_x) = 0$, so choosing $x = a$ gives $\phi_a(a) = 0$. But this is not possible if $a \neq 0$. Therefore $a = 0$ and π is one-to-one, so isometric. □

The Gelfand–Najmark theorem enables us to introduce an extremely useful technical tool called the *universal representation* of a C^*-algebra with unit.

Let $\mathfrak{A}$ be a C^*-algebra with unit and denote by $S(\mathfrak{A}) \subset \mathfrak{A}'$ its convex set of algebraic states, by $(\mathsf{H}_\omega, \pi_\omega, \Psi_\omega)$ the GNS representation of state $\omega \in S(\mathfrak{A})$. Consider the Hilbert sum $\bigoplus_{\omega\in S(\mathfrak{A})} \mathsf{H}_\omega$. Its elements $\oplus_{\omega\in S(\mathfrak{A})} \psi_\omega := \{\psi_\omega\}_{\omega\in S(\mathfrak{A})}$ satisfy

$$\sum_{\omega\in S(\mathfrak{A})} ||\psi_\omega||_\omega^2 < +\infty . \tag{14.5}$$

The space $\bigoplus_{\omega\in S(\mathfrak{A})} \mathsf{H}_\omega$ is a Hilbert space for the inner product

$$(\psi|\psi') = \sum_{\omega\in S(\mathfrak{A})} (\psi_\omega|\psi'_\omega)_\omega .$$

The **universal representation** of $\mathfrak{A}$ is the representation:

$$\Pi : \mathfrak{A} \to \mathfrak{B}\left(\bigoplus_{\omega \in S(\mathfrak{A})} \mathsf{H}_\omega\right) \quad \text{given by} \quad \Pi\left(\bigoplus_{\omega \in S(\mathfrak{A})} \psi_\omega\right) := \bigoplus_{\omega \in S(\mathfrak{A})} \pi_\omega(a)\psi_\omega\,.$$

Definition 14.24. *Let $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ be a representation of the *-algebra $\mathfrak{A}$ on the Hilbert space H. A* **subrepresentation** *of π is a representation of the form $\pi\restriction_{\mathsf{H}_0} : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_0)$, where the subspace $\mathsf{H}_0 \subset \mathsf{H}$ is closed and π-invariant.*

Clearly any GNS representation of a C^*-algebra with unit is a subrepresentation of the universal representation. Then the next easy, but useful, fact holds.

Proposition 14.25. *The universal representation of any given C^*-algebra with unit is faithful and isometric.*

Proof. That a representaion is faithful implies, by Theorem 8.22, that it is isometric. Faithfulness descends immediately from the fact that Π, as subrepresentation, contains the representation π used in the proof of the Gelfand–Najmark theorem (the latter is injective). □

Eventually we mention a result on the structure of the folium of an algebraic state. First, a notation and an important definition.

Notation 14.26. Let $\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H})$ be a representation of the *-algebra $\mathfrak{A}$, and n a cardinal number. Then we denote by $n\pi$ the representation on $\bigoplus_{i=1}^n \mathsf{H}_i$, $\mathsf{H}_i := \mathsf{H}$ defined by

$$n\pi(a)(\oplus_{i=1}^n \psi_i) := \oplus_{i=1}^n \pi(a)\psi_i \quad \text{for any } a \in \mathfrak{A},\ \psi_i \in \mathsf{H}\,.$$ ■

Definition 14.27. *Two representations $\pi_1 : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_1)$, $\pi_2 : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}_2)$ of the same *-algebra $\mathfrak{A}$ are called* **quasi-equivalent**, *written*

$$\pi_1 \approx \pi_2\,,$$

if they are unitarily equivalent up to multiplicities. Equivalently, there exist cardinals n_1, n_2 such that $n_1\pi_1 \simeq n_2\pi_2$.

For example (indicating $(A \oplus B)(u \oplus v) := Au \oplus Bv$)

$$\pi : \mathfrak{A} \to \mathfrak{B}(\mathsf{H}) \quad \text{and} \quad \pi_1 : \mathfrak{A} \ni a \mapsto \pi(a) \oplus U\pi(a)U^{-1} \in \mathfrak{B}(\mathsf{H} \oplus \mathsf{H}')$$

are quasi-equivalent if $U : \mathsf{H} \to \mathsf{H}'$ is a unitary operator. Unitarily equivalent representations are obviously quasi-equivalent. And quasi-equivalence is an equivalence relation. About this (see [Haa96] and [BrRo02, vol. 1]) we have

Proposition 14.28. *Let $\mathfrak{A}$ be a C^*-algebra with unit and $\omega : \mathfrak{A} \to \mathbb{C}$ an algebraic state with GNS representation π_ω.*

(a) *If π_1 and π_2 are representations of $\mathfrak{A}$, $\pi_1 \approx \pi_2$ if and only if the von Neumann algebras $\pi_1(\mathfrak{A})''$, $\pi_2(\mathfrak{A})''$ are *-isomorphic as *-algebras, and the *-isomorphism restricts to a *-isomorphism from $\pi_1(\mathfrak{A})$ to $\pi_2(\mathfrak{A})$.*

(b) *The GNS representations of* $\mathfrak{A}$ *generated by states in* $Fol(\omega)$ *are quasi-equivalent. In particular, if* $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$, *with* $\lambda \in (0,1)$ *and* $\omega_1 \neq \omega$, *the GNS representation of* ω_1 *is unitarily equivalent to a GNS subrepresentation of* ω.
(c) *If* π *is a representation of* $\mathfrak{A}$ *and* $\pi \approx \pi_\omega$, *then* π *is a GNS representation of a state in* $Fol(\omega)$.

14.2 Example of a C^*-algebra of observables: the Weyl C^*-algebra

This section is devoted to the simplest non-trivial C^*-algebra of observables used in physics. We are talking about the Weyl C^*-algebra involved in the description of several systems, among which non-interacting Bosonic quantum systems. Almost all systems that are describable using a Weyl C^*-algebra can also be described by weakening the observables' structure to a *-algebra. Yet Weyl C^*-algebras are mathematically attractive, motivating our interest.

14.2.1 Further properties of Weyl *-algebras $\mathscr{W}(\mathsf{X},\sigma)$

Keeping in mind Chapter 11.3.4, let (X,σ) be a symplectic space: a pair consisting of a real vector space X of any even dimension (possibly infinite), henceforth non-trivial, and a weakly non-degenerate symplectic form $\sigma : \mathsf{X}\times\mathsf{X} \to \mathbb{R}$. Let $\mathscr{W}(\mathsf{X},\sigma)$ denote the Weyl *-algebra of (X,σ) introduced in Definition 11.25. We know (Theorem 11.26) it is defined up to *-isomorphisms. We wish to explain that it is possible, and in a unique way, to enlarge $\mathscr{W}(\mathsf{X},\sigma)$ to a C^*-algebra called the *Weyl C^*-algebra associated to* (X,σ). More precisely, we will exhibit on $\mathscr{W}(\mathsf{X},\sigma)$ a unique norm satisfying the C^* property: $||a^*a|| = ||a||^2$. The Weyl C^*-algebra is then defined to be the completion of $\mathscr{W}(\mathsf{X},\sigma)$ for that norm. In order to prove all this we need a few preliminary facts that form the contents of the section. Various procedures exist, and distinct (equivalent) formulations, that prove the ensuing properties (see [BrRo02] in particular). We will essentially follow the approach of [BGP07].

Lemma 14.29. *Let* X *be a non-trivial real vector space,* $\sigma : \mathsf{X}\times\mathsf{X} \to \mathbb{R}$ *a weakly non-degenerate symplectic form, and consider a Weyl* **-algebra* $\mathscr{W}(\mathsf{X},\sigma)$ *associated to the system.*

(a) *There exists a norm* $||\,||$ *on* $\mathscr{W}(\mathsf{X},\sigma)$ *satisfying the* C^* *property:* $||a^*a|| = ||a||^2$ *for any* $a \in \mathscr{W}(\mathsf{X},\sigma)$.
(b) *If* $\psi \in \mathsf{X}$, *the generator* $W(\psi)$ *is unitary, so for the above norm* $||W(\psi)|| = 1$.
(c) *If* $\psi,\phi \in \mathsf{X}$, $\psi \neq \phi$, *in the above norm*

$$||W(\psi) - W(\phi)|| = 2\,,$$

so $\mathscr{W}(\mathsf{X},\sigma)$ *is not separable.*
(d) *If we set, for any* $a \in \mathscr{W}(\mathsf{X},\sigma)$:

$$||a||_c := \sup\{p(a)\,|\,p : \mathscr{W}(\mathsf{X},\sigma) \to [0,+\infty) \textit{ is a } C^* \textit{ norm}\}\,,$$

then $||\,||_c$ *is a* C^* *norm.*

Proof. (a) Let us focus again on the construction of $\mathscr{W}(\mathsf{X},\sigma)$ of Theorem 11.26(a). Consider the complex Hilbert space $\mathsf{H} := L^2(\mathsf{X},\mu)$ where μ is the counting measure on X. For $u \in \mathsf{X}$ the operators $W(u) \in \mathfrak{B}(L^2(\mathsf{X},\mu))$, $(W(u)\psi)(v) := e^{i\sigma(u,v)}\psi(u+v)$ for $\psi \in L^2(\mathsf{X},\mu)$, $v \in \mathsf{X}$, define a Weyl *-algebra associated to (X,σ): $\mathscr{W}(\mathsf{X},\sigma) \subset \mathfrak{B}(\mathsf{H})$. The norm $|| \; ||$ of $\mathfrak{B}(\mathsf{H})$ satisfies the C^* property. Starting from a different representation $\mathscr{W}'(\mathsf{X},\sigma)$, $|| \; ||$ induces a C^* norm on $\mathscr{W}'(\mathsf{X},\sigma)$ by means of the *-isomorphism $\alpha : \mathscr{W}(\mathsf{X},\sigma) \to \mathscr{W}'(\mathsf{X},\sigma)$ of Theorem 11.26(c).
(b) From the Weyl relations $W(\psi)W(\psi)^* = W^*(\psi)W(\psi) = \mathbb{I}$, so $W(\psi)$ is unitary. The C^* property implies $||W(\psi)|| = 1$.
(c) Let us complete $\mathscr{W}(\mathsf{X},\sigma)$ with respect to the norm $|| \; ||$ of (a), so to obtain a C^*-algebra. By Weyl's relations we have $W(\chi)W(\phi-\psi)W(\chi)^{-1} = e^{-i\sigma(\chi,\phi-\psi)}W(\phi-\psi)$. Since $W(\phi-\psi)$ is unitary, $\sigma(\phi-\psi) \subset \{z \in \mathbb{C} \,|\, |z| = 1\}$. By definition of spectrum

$$\sigma(W(\chi)W(\phi-\psi)W(\chi)^{-1}) = \sigma(W(\phi-\psi)) = e^{-i\sigma(\chi,\phi-\psi)}\sigma(W(\phi-\psi))\,.$$

Since $\psi \neq \phi$, $\sigma(\chi,\phi-\psi)$ covers the whole $\mathbb{R}$ as χ varies in X. Hence $\sigma(W(\phi-\psi)) = \{z \in \mathbb{C} \,|\, |z| = 1\}$. Therefore $\sigma(e^{i\sigma(\psi,\phi)}W(\phi-\psi) - \mathbb{I})$ is the unit circle in $\mathbb{C}$ centred at -1, so if r is the spectral radius, $r(e^{i\sigma(\psi,\phi)}W(\phi-\psi)) = 2$. But $e^{i\sigma(\psi,\phi)}W(\phi-\psi)$ is normal: $2 = r(e^{i\sigma(\psi,\phi)}W(\phi-\psi)) = ||e^{i\sigma(\psi,\phi)}W(\phi-\psi) - \mathbb{I}||$. Using the norm's C^* property ans the generators' unitarity, the Weyl identities imply that $||W(\phi) - W(\psi)||^2$ equals

$$||(W(\phi)^* - W(\psi)^*)(W(\phi) - W(\psi))|| = ||e^{i\sigma(\psi,\phi)}W(\phi-\psi) - \mathbb{I}|| = 4\,.$$

There are uncountably many elements $\psi \in \mathsf{X}$ ($\mathsf{X} \neq \{0\}$ by assumption), so $\mathscr{W}(\mathsf{X},\sigma)$ is not separable: if $S \subset \mathsf{X}$ were dense, there would be an element of S inside the ball of radius $1/2$ centred at each $W(\psi)$, but said balls do not intersect, so S cannot be countable.
(d) Every property of a norm, plus the C^* property $||a^*a||_c = ||a||_c^2$, hold by direct inspection. The only thing left is to show that the supremum defining $||a||_c$ is finite. To this end, on $\mathscr{W}(\mathsf{X},\sigma)$ we have a norm (not C^* in general): $||\sum_i a_i W(\psi_i)||_0 := \sum_i |a_i|$. As every $W(\psi)$ has unit norm with respect to any C^* norm p, as seen in (b), we have $p(a) \leq ||a||_0 < +\infty$. Therefore the least upper bound in $||a||_c$ is smaller than $||a||_0$, hence finite. □

Lemma 14.30. *Let* (X,σ) *be a non-trivial weakly non-degenerate real symplectic space,* $\mathscr{W}(\mathsf{X},\sigma)$ *a Weyl* **-algebra associated to* (X,σ)*. Denote by* $C\mathscr{W}(\mathsf{X},\sigma)$ *the* C^* *completion of* $\mathscr{W}(\mathsf{X},\sigma)$ *in the norm* $|| \; ||_c$ *of Lemma 14.29(d).*
Then $C\mathscr{W}(\mathsf{X},\sigma)$ *is simple: it does not admit two-sided closed ideals invariant under the involution other than* $\{0\}$ *and* $C\mathscr{W}(\mathsf{X},\sigma)$ *itself.*

Proof. Write $\mathfrak{A}$ for the C^*-algebra with unit obtained by completion of $\mathscr{W}(\mathsf{X},\sigma)$ under $|| \; ||_c$. Suppose $I \subset \mathfrak{A}$ is a closed, two-sided ideal that is *-invariant. Then $I_0 := I \cap \{cW(0) \,|\, c \in \mathbb{C}\}$ is a complex subspace of $\{cW(0) \,|\, c \in \mathbb{C}\}$ identified with $\mathbb{C}$. Hence $I_0 = \{0\}$ or $I_0 = \{cW(0) \,|\, c \in \mathbb{C}\}$. In the latter case I would then contain $\mathbb{I}$, so it would coincide with $\mathfrak{A}$. So assume $I_0 = \{0\}$ and consider the map:

$$P : \mathscr{W}(\mathsf{X}, \sigma) \to \{cW(0) \,|\, c \in \mathbb{C}\},$$

with

$$P\left(\sum_{\phi \in F \subset \mathsf{X}} W(\phi)\right) = a_0 W(0) \text{ in case } F \subset \mathsf{X} \text{ is finite}.$$

We claim P is bounded, and that it extends continuously to an operator, P, defined on $\mathfrak{A}$. To do so let us realise $\mathscr{W}(\mathsf{X}, \sigma)$ in the C^*-algebra of operators $\mathfrak{B}(L^2(\mathsf{X}, \mu))$, as in the proof of Lemma 14.29(a). Call $\delta_0 \in L^2(\mathsf{X}, \mu)$ the map $\delta_0(0) = 1$, $\delta_0(\phi) = 0$ for $\phi \neq 0$. For $a = \sum_{\phi \in F \subset \mathsf{X}} a_\phi W(\phi)$ and $\psi \in \mathsf{X}$ we have

$$(a\delta_0)(\psi) = \left(\sum_{\phi \in F \subset \mathsf{X}} a_\phi W(\phi)\delta_0\right)(\psi) = \sum_{\phi \in F \subset \mathsf{X}} a_\phi e^{i\sigma(\phi,\psi)/2}\delta_0(\phi + \psi)$$
$$= a_{-\psi} e^{i\sigma(-\psi,\psi)/2} = a_{-\psi}.$$

Consequently

$$(\delta_0 | a\delta_0)_{L^2(\mathsf{X},\mu)} = \sum_{\psi \in \mathsf{X}} \overline{\delta_0(\psi)}(a\delta_0)(\psi) = (a\delta_0)(0) = a_0.$$

In addition, $||\delta_0|| = 1$, so

$$||P(a)||_c = ||a_0 W(0)||_c = |a_0| = |(\delta_0 | a\delta_0)_{L^2}| \leq ||a||_{op} \leq ||a||_c,$$

proving P extends to a bounded operator on $\mathfrak{A}$.

Take now $a \in I \subset \mathfrak{A}$ and fix $\varepsilon > 0$. Write

$$a = a_0 W(0) + \sum_{j=1}^{n} a_j W(\phi_j) + r,$$

where the ϕ_j are all distinct and $||r||_c < \varepsilon$. For $\psi \in \mathsf{X}$ we have

$$I \ni W(\psi) a W(-\psi) = a_0 W(0) + \sum_{j=1}^{n} a_j e^{-i\sigma(\psi,\phi_j)/2} W(\phi_j) + r(\psi),$$

since

$$||r(\phi)||_c = ||W(\psi) r W(-\psi)||_c \leq ||r||_c < \varepsilon.$$

Choosing ψ_1 and ψ_2 so that $e^{-i\sigma(\psi_1,\phi_n)} = -e^{-i\sigma(\psi_2,\phi_n)}$, then adding two elements

$$a_0 W(0) + \sum_{j=1}^{n} a_j e^{-i\sigma(\psi_1,\phi_j)/2} W(\phi_j) + r(\psi_1) \in I$$

and

$$a_0 W(0) + \sum_{j=1}^{n} a_j e^{-i\sigma(\psi_2,\phi_j)/2} W(\phi_j) + r(\psi_2) \in I$$

gives

$$a_0 W(0) + \sum_{j=1}^{n-1} a'_j W(\phi_j) + r_1 \in I\,,$$

where $||r_1||_c = \frac{1}{2}||r(\psi_1) + r(\psi_2)||_c < (\varepsilon + \varepsilon)/2 = \varepsilon$. We can repeat the argument, and eventually obtain, for some r_n with $||r_n||_c < \varepsilon$:

$$a_0 W(0) + r_n \in I\,.$$

As $\varepsilon > 0$ is arbitrary and I closed, we conclude $P(a) = a_0 W(0) \in I_0$, so $a_0 = 0$. With $\psi \in \mathsf{X}$ and $a = \sum_\phi a_\phi W(\phi) \in I$ arbitrary, we similarly have $W(\psi)a \in I$, whence $P(W(\psi)a) = 0$. This means $a_{-\psi} = 0$ for any $\psi \in \mathsf{X}$, so $a = 0$. Therefore $I = \{0\}$, ending the proof. □

Now to the key theorem on a given symplectic space's Weyl C^*-algebras.

Theorem 14.31. *Let (X, σ) be a non-trivial weakly non-degenerate real symplectic space, and consider a Weyl *-algebra $\mathscr{W}(\mathsf{X}, \sigma)$ associated to (X, σ).*

(a) *There exist a unique norm on $\mathscr{W}(\mathsf{X}, \sigma)$ satisfying the C^* property:*

$$||a^* a|| = ||a||^2 \quad \textit{for any } a \in \mathscr{W}(\mathsf{X}, \sigma).$$

(b) *Let $C\mathscr{W}(\mathsf{X}, \sigma)$ be the C^*-algebra completion of $\mathscr{W}(\mathsf{X}, \sigma)$ for the C^* norm of (a). If $\mathscr{W}'(\mathsf{X}, \sigma)$ is another Weyl *-algebra associated to the same space (X, σ) and $||\ ||'$ the unique C^* norm, call $CW'(\mathsf{X}, \sigma)$ the corresponding C^*-algebra with unit.*
Then there is a unique isometric *-isomorphism $\gamma : C\mathscr{W}(\mathsf{X}, \sigma) \to C\mathscr{W}'(\mathsf{X}, \sigma)$ such that:

$$\gamma(W(\psi)) = W'(\psi) \quad \textit{for any } \psi \in \mathsf{X},$$

where $W(\psi)$, $W'(\psi)$ are generators of the Weyl *-algebras $\mathscr{W}(\mathsf{X}, \sigma)$, $\mathscr{W}'(\mathsf{X}, \sigma)$.

Proof. (a) By Theorem 11.26(c) it is known that two Weyl *-algebras $\mathscr{W}(\mathsf{X}, \sigma)$, $\mathscr{W}'(\mathsf{X}, \sigma)$ on the same symplectic space are *-isomorphic under some $\alpha : \mathscr{W}(\mathsf{X}, \sigma) \to \mathscr{W}'(\mathsf{X}, \sigma)$ that is totally determined by $\alpha(W(\psi)) = W'(\psi)$, $\psi \in \mathsf{X}$. Equip $\mathscr{W}(\mathsf{X}, \sigma)$, $\mathscr{W}'(\mathsf{X}, \sigma)$ with C^* norms $||\ ||$, $||\ ||'$. Then $||a||_1 = ||\alpha(a)||'$ is a C^* norm on $\mathscr{W}(\mathsf{X}, \sigma)$, other than $||\ ||$ in general. By definition of $||\ ||_c$ we have $||\alpha(a)||' \leq ||a||_c$, so α extends to a *-homomorphism of C^*-algebras:

$$\widetilde{\alpha} : \overline{\mathscr{W}(\mathsf{X}, \sigma)}_{||\ ||_c} \to \overline{\mathscr{W}(\mathsf{X}', \sigma')}_{||\ ||'}\,.$$

The kernel of $\widetilde{\alpha}$ is a closed *-invariant two-sided ideal, hence trivial by the previous lemma. In conclusion $\widetilde{\alpha}$ is one-to-one and an isometry by Theorem 8.22(a). Suppose now $\mathscr{W}(\mathsf{X}, \sigma) = \mathscr{W}'(\mathsf{X}, \sigma)$, so $||\ || = ||\ ||'$ too. Then α has to be the identity, extending to the identity $\widetilde{\alpha}$, and also isometric by the above argument. So, $||\ ||_c = ||\ ||' = ||\ ||$ is the only C^* norm on $\mathscr{W}(\mathsf{X}, \sigma)$.
(b) We have to prove that the *-isomorphism $\alpha : \mathscr{W}(\mathsf{X}, \sigma) \to \mathscr{W}'(\mathsf{X}, \sigma)$, determined by $\alpha(W(\psi)) = W'(\psi)$, $\psi \in \mathsf{X}$, extends to a *-isomorphism between the C^*-algebras $C\mathscr{W}(\mathsf{X}, \sigma)$ and $C\mathscr{W}'(\mathsf{X}, \sigma)$. The same argument used above (now we do

know $|| \, || = || \, ||_c$) shows that α extends to an injective *-homomorphism $\gamma : C\mathscr{W}(\mathsf{X},\sigma) \to C\mathscr{W}'(\mathsf{X},\sigma)$. On the other hand we can swap $\mathscr{W}(\mathsf{X},\sigma)$ and $\mathscr{W}'(\mathsf{X},\sigma)$, and extend $\alpha' : \mathscr{W}'(\mathsf{X},\sigma) \to \mathscr{W}(\mathsf{X},\sigma)$, determined by $\alpha'(W'(\psi)) = W(\psi)$, $\psi \in \mathsf{X}$, to $\gamma' : C\mathscr{W}'(\mathsf{X},\sigma) \to C\mathscr{W}(\mathsf{X},\sigma)$. By construction $\alpha'\alpha = id_{\mathscr{W}(\mathsf{X},\sigma)}$, $\alpha\alpha' = id_{\mathscr{W}'(\mathsf{X},\sigma)}$, which continue to hold, by continuity, when extended to $\gamma'\gamma = id_{C\mathscr{W}(\mathsf{X},\sigma)}$, $\gamma\gamma' = id_{C\mathscr{W}'(\mathsf{X},\sigma)}$. Therefore γ is onto, as well, and thus a *-isomorphism. □

14.2.2 The Weyl C^*-algebra $C\mathscr{W}(\mathsf{X},\sigma)$

By keeping Theorem 14.31 into account, we can define Weyl C^*-algebras.

Definition 14.32. *Let X be a non-trivial real vector space equipped with a weakly non-degenerate symplectic form $\sigma : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$. The* **Weyl C^*-algebra** $C\mathscr{W}(\mathsf{X},\sigma)$ *associated to (X,σ) is a C^*-algebra with unit generated by non-zero elements $W(\psi)$, $\psi \in \mathsf{X}$, satisfying the Weyl relations:*

$$W(\psi)W(\psi') = e^{-\frac{i}{2}\sigma(\psi,\psi')}W(\psi+\psi')\,, \quad W(\psi)^* = W(-\psi)\,, \quad \psi,\psi \in \mathsf{X}\,.$$

This notion is well defined, and as consequence of Theorem 14.31 we obtain the following result. It shows that the Weyl C^*-algebra is unique up to *-isomorphisms.

Theorem 14.33. *Let (X,σ) be a non-trivial weakly non-degenerate real symplectic space, $C\mathscr{W}(\mathsf{X},\sigma)$ a Weyl C^*-algebra associated to it.*

(a) *If $C\mathscr{W}'(\mathsf{X},\sigma)$ is a second Weyl C^*-algebra associated to (X,σ), there exists a unique (isometric) *-isomorphism $\gamma : C\mathscr{W}(\mathsf{X},\sigma) \to C\mathscr{W}'(\mathsf{X},\sigma)$ such that*

$$\gamma(W(\psi)) = W'(\psi) \quad \textit{for any } \psi \in \mathsf{X},$$

where $W(\psi)$, $W'(\psi)$ generate the Weyl *-algebras $\mathscr{W}(\mathsf{X},\sigma)$, $\mathscr{W}'(\mathsf{X},\sigma)$ respectively.
(b) *$C\mathscr{W}(\mathsf{X},\sigma)$ is simple: there are no non-trivial closed, *-invariant two-sided ideals.*
(c) *If $C\mathscr{W}(\mathsf{X}',\sigma')$ is the Weyl C^*-algebra associated to the weakly non-degenerate symplectic space (X',σ') and $f : \mathsf{X} \to \mathsf{X}$ a symplectic homomorphism, there is a unique injective and isometric *-homomorphism $\gamma_f : C\mathscr{W}(\mathsf{X},\sigma) \to C\mathscr{W}(\mathsf{X}',\sigma')$ such that:*

$$\gamma_f(W(\psi)) = W'(f(\psi)) \quad \textit{for any } \psi \in \mathsf{X}.$$

$\gamma_f(C\mathscr{W}(\mathsf{X},\sigma))$ is a C^*-subalgebra with unit of $C\mathscr{W}(\mathsf{X}',\sigma')$.

Proof. Items (a) and (b) were actually proven with Theorem 14.31. Let us see to (c). By Theorem 11.26(f) there is one injective *-homomorphism $\alpha_f : \mathscr{W}(\mathsf{X},\sigma) \to \mathscr{W}(\mathsf{X}',\sigma')$ such that $\alpha_f(W(\psi)) = W'(f(\psi))$, $\psi \in \mathsf{X}$. The C^* norm $|| \, ||'$ on $C\mathscr{W}(\mathsf{X}',\sigma')$ induces a C^* norm on $\mathscr{W}(\mathsf{X},\sigma)$, $||a|| = ||\alpha_f(a)||'$. By uniqueness of the C^* norm on a Weyl *-algebra, the latter coincides with the original norm of $C\mathscr{W}(\mathsf{X},\sigma)$. Hence α_f is isometric and continuous, and extends continuously to an isometric (so injective) *-homomorphism $\gamma_f : C\mathscr{W}(\mathsf{X},\sigma) \to C\mathscr{W}(\mathsf{X}',\sigma')$. That $\gamma_f(C\mathscr{W}(\mathsf{X},\sigma))$ is a C^*-subalgebra with unit in $C\mathscr{W}(\mathsf{X}',\sigma')$ follows from Theorem 8.22(b). □

Remarks 14.34. If $\mu : \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ is a real inner product fulfilling

$$\frac{1}{4}|\sigma(\psi,\phi)|^2 \leq \mu(\psi,\psi)\mu(\phi,\phi)\,, \quad \text{for any } \psi,\phi \in \mathsf{X},$$

it can be proved there exists a unique algebraic state ω_μ on $C\mathscr{W}(\mathsf{X},\sigma)$ such that:

$$\omega_\mu(W(\psi)) = e^{-\frac{1}{2}\mu(\psi,\psi)}\,.$$

This type of states are called *Gaussian* or *quasi-free*, and play a big role in physical theories. The GNS representations of a quasi-free state generates Hilbert spaces with totally symmetric Fock structure (Bosonic Fock spaces). ■

Example 14.35. Take Minkowski's spacetime, with coordinates $(t,\mathbf{x}) \in \mathbb{R} \times \mathbb{R}^3$, and consider there the *Klein–Gordon equation*:

$$-\frac{1}{c^2}\frac{\partial^2 \phi}{\partial t^2} + \Delta_{\mathbf{x}}\phi - \frac{m^2c^2}{\hbar^2}\phi = 0\,,$$

where c is the speed of light and $m > 0$ the mass of the particles associated to the Bosonic field ϕ. Indicate with X the vector space of real smooth solutions ϕ such that $\mathbb{R}^3 \ni \mathbf{x} \mapsto \phi(t,\mathbf{x})$ have compact support for every $t \in \mathbb{R}$. This space admits a weakly non-degenerate symplectic form:

$$\sigma(\phi,\phi') := \int_{\mathbb{R}^3} \left(\phi(t,\mathbf{x})\frac{\partial}{\partial t}\phi'(t,\mathbf{x}) - \phi'(t,\mathbf{x})\frac{\partial}{\partial t}\phi(t,\mathbf{x})\right) dx\,.$$

For given solutions X, one can prove that the symplectic form does not depend on the choice of $t \in \mathbb{R}$ by the nature of the Klein–Gordon equation itself. The C^*-algebra $C\mathscr{W}(\mathsf{X},\sigma)$ is the algebra of observables of the Klein–Gordon quantum field ϕ, and can be taken as the starting point for the procedure of "second quantisation" of Bosonic fields. In this case an algebraic state of paramount importance is the so-called *Minkowski vacuum*, i.e. the *Gaussian state* (see Remark 14.34) determined by a special μ that takes spatial Fourier transforms of solutions at time $t = 0$. This particular state represents the absence of particles, and is invariant under the Poincaré group. In the GNS representation of the state, the Weyl generators have the form $\pi_{\omega_\mu}(W(\phi)) = e^{i\Phi(\phi)}$. The self-adjoint operator $\Phi(\phi)$ is called *operator of the field of second quantisation.* ■

14.3 Introduction to Quantum Symmetries within the algebraic formulation

In this section we briefly discuss how quantum symmetries are dealt with in the algebraic formulation [Haa96, Str05b]. After recalling the basic notions, we will prove two theorems about the (anti)unitary representation of symmetries on the space of the GNS representation of an invariant algebraic state. The strategy allows to describe precisely, in mathematical terms, the concept of the *spontaneous breaking of symmetry*.

14.3.1 The algebraic formulation's viewpoint on quantum symmetries

Consider a quantum system S described by the C^*-algebra with unit $\mathfrak{A}_S$ of observables. Said better, observables are the self-adjoint elements of $\mathfrak{A}_S$. A quantum symmetry α should be seen as a *-automorphism $\alpha : \mathfrak{A}_S \to \mathfrak{A}_S$, i.e. as a bijective *-homomorphism (hence isometric), or as a **-anti-automorphism*.

Definition 14.36. *If $\mathfrak{A}$ is a C^*-algebra with unit $\mathbb{I}$, a* ***-anti-automorphism** *is a bijective, antilinear isometry $\alpha : \mathfrak{A} \to \mathfrak{A}$ such that $\alpha(\mathbb{I}) = \mathbb{I}$, $\alpha(a^*) = \alpha(a)^*$ and $\alpha(ab) = \alpha(a)\alpha(b)$, for any $a,b \in \mathfrak{A}$.*

A "psychological explanation" for the above definition of algebraic symmetry emerges from the notion of quantum symmetry when the theory is formulated over a Hilbert space and recalling the theorems of Wigner and Kadison. Fix a GNS triple $(\mathsf{H}_\omega, \pi_\omega, \Psi_\omega)$, suppose the GNS representation $\pi_\omega : \mathfrak{A}_S \to \mathfrak{B}(\mathsf{H})$ is injective (always the case if $\mathfrak{A}_S$ is simple, as we said in Chapter 14.1.3), and represent the symmetry on H_ω by the operator U, unitary or anti-unitary. Then we can set

$$\alpha(a) := \pi_\omega^{-1}\left(\gamma^*(\pi_\omega(a))\right), \quad a \in \mathfrak{A}_S,$$

where, mimicking the previous section's definition, the action γ of the symmetry U on observables is:

$$\gamma^*(A) := U^{-1}AU\,.$$

α is well defined, and gives a *-automorphism or *-anti-automorphism provided $U^{-1}\cdot U$ maps observables (seen as operators) to observables, as is only natural to suppose.

Here is the formal definition.

Definition 14.37. *Let S be a physical system described by the C^*-algebra with unit $\mathfrak{A}_S$ of observables. An* **(algebraic) quantum symmetry** *of S is a *-automorphism or *-anti-automorphism $\alpha : \mathfrak{A}_s \to \mathfrak{A}_s$.*

This natually begs a question: given a symmetry α and an algebraic state ω, under which assumptions is α representable by a unitary, or anti-unitary, operator on the Hilbert space H_ω of the GNS representation of ω? The next theorem is a big step forward in this direction.

Theorem 14.38. *Let α be an algebraic quantum symmetry of system S, described by the C^*-algebra with unit $\mathfrak{A}_S$ of observables. Suppose ω is an* α**-invariant** *algebraic state on $\mathfrak{A}_S$, :*

$$\omega(\alpha(a)) = \omega(a) \quad \text{for } a \in \mathfrak{A}_S \text{ with } a = a^*. \tag{14.6}$$

If $(\mathsf{H}_\omega, \pi_\omega, \Psi_\omega)$ is the GNS triple of ω, there exists only one operator $U_\alpha : \mathsf{H}_\omega \to \mathsf{H}_\omega$, unitary or anti-unitary according to whether α is linear or antilinear, such that:

$$U_\alpha \Psi_\omega = \Psi_\omega \quad \text{and} \quad U_\alpha^{-1}\pi_\omega(a)U_\alpha = \pi_\omega(\alpha(a)), \quad a \in \mathfrak{A}_S. \tag{14.7}$$

Remarks 14.39. Since any $a \in \mathfrak{A}_S$ can be written $a = a_1 + ia_2^*$ with a_1, a_2 self-adjoint, and $\omega(a^*) = \overline{\omega(a)}$ for any state ω (straightforward from the GNS theorem), ω being α-invariant is the same as imposing $\omega(\alpha(a)) = \omega(a)$ if α is linear, or $\omega(\alpha(a)) = \overline{\omega(a)}$, if α is antilinear, for any $a \in \mathfrak{A}_S$. ■

Proof of Theorem 14.38. The idea is to define U_α first on the dense space $\pi(\mathfrak{A}_S)\Psi_\omega$ by

$$U\pi_\omega(a)\Psi_\omega = \pi(\alpha^{-1}(a))\Psi_\omega\,, \tag{14.8}$$

and then extend it continuously to H_ω. The definition is unambiguous if $\pi_\omega(a)\Psi_\omega = \pi_\omega(a')\Psi_\omega$ implies $\pi(\alpha^{-1}(a))\Psi_\omega = \pi(\alpha^{-1}(a'))\Psi_\omega$, i.e. if $\pi_\omega(b)\Psi_\omega = 0$ implies $\pi(\alpha^{-1}(b))\Psi_\omega = 0$. But this is true by the GNS theorem and ω's invariance:

$$\begin{aligned}||\pi(\alpha^{-1}(b))\Psi_\omega||^2 &= (\pi(\alpha^{-1}(b))\Psi_\omega|\pi(\alpha^{-1}(b))\Psi_\omega) = (\Psi_\omega|\pi(\alpha^{-1}(b^*))\pi(\alpha(b))\Psi_\omega)\\ &= (\Psi_\omega|\pi(\alpha^{-1}(b^*b))\Psi_\omega) = \omega(\alpha^{-1}(b^*b)) = \omega(b^*b) = (\Psi_\omega|\pi(b^*b)\Psi_\omega)\\ &= ||\pi(b)\Psi_\omega||^2\,.\end{aligned}$$

By construction U, as in (14.8), is linear or antilinear depending on how α is. Moreover, it is isometric/anti-isometric, if α is a *-isomorphism/anti-isomorphism respectively. Hence it is continuous, as above computations show that

$$||U\pi(b)\Psi_\omega||^2 = ||\pi(b)\Psi_\omega||^2\,.$$

We extend U by continuity to H_S, since $\pi(\mathfrak{A}_S)\Psi_\omega$ is dense in H, and obtain a linear/antilinear operator $U_\alpha : \mathsf{H}_\omega \to \mathsf{H}_\omega$ preserving norms. U is onto as inverse of the analogous uniquely-defined extension of

$$U^{-1}\pi_\omega(a)\Psi_\omega = \pi(\alpha(a))\Psi_\omega\,. \tag{14.9}$$

Therefore $U_\alpha : \mathsf{H}_\omega \to \mathsf{H}_\omega$ is well defined, unitary/anti-unitary if α is linear/antilinear, and satisfies (14.7). The first condition is trivially true setting $b = \mathbb{I}$ in (14.8). As for the second one, put $a = bc$ in (14.9):

$$U_\alpha^{-1}\pi_\omega(b)\pi_\omega(c)\Psi_\omega = \pi(\alpha(b))\pi_\omega(\alpha(c))\Psi_\omega\,.$$

Using (14.9):

$$U_\alpha^{-1}\pi_\omega(b)U_\alpha\pi_\omega(\alpha(c))\Psi_\omega = \pi(\alpha(b))\pi_\omega(\alpha(c))\Psi_\omega\,.$$

That is to say, if $\Phi \in \pi_\omega(\alpha(\mathfrak{A}_S))\Psi_\omega = \pi_\omega(\mathfrak{A}_S)\Psi_\omega$:

$$U_\alpha^{-1}\pi_\omega(b)U\Phi = \pi(\alpha(b))U_\alpha\Phi\,.$$

Since $\pi_\omega(\mathfrak{A}_S)\Psi_\omega$ is dense in H_ω, the second identity in (14.7) holds. Uniqueness of U_α is patent by construction, because if V satisfies (14.7) (V replacing U_α) it must satisfy (14.9) as well (V replacing U), which fact fixes it. □

Remarks 14.40. A system S may admit an algebraic symmetry α that is *not* representable unitarily (or anti-unitarily) on the Hilbert space of the theory (e.g., a GNS representation of a reference algebraic state ω). If so, the symmetry α is said to have been **broken spontaneously** by the representation employed. The phenomenon of **spontaneous symmetry breaking** is hugely important in particle physics and statistical mechanics [Str05b]. ■

14.3.2 (Topological) symmetry groups in the algebraic formalism

We want to show, concisely, how Theorem 14.38, proven in Chapter 14.3.1 in the algebraic formalism, generalises naturally to the situation where the algebraic symmetry α is replaced by an algebraic symmetry group.

So take a quantum system S described, in the algebraic formalism, by the C^*-algebra with unit $\mathfrak{A}_S$, whose self-adjoint elements are the system's observables. Suppose there is a representation $\alpha : \mathsf{G} \ni g \mapsto \alpha_g$ of the group G in terms of *-automorphisms α_g of $\mathfrak{A}_S$. If ω is an invariant algebraic state, Theorem 14.38 guarantees every α_g is representable by a unitary operator U_{α_g} on the Hilbert space H_ω of the GNS representation of ω. We will show that this correspondence produces automatically a unitary right representation of G, without the need to redefine the phases of the unitary U_g. This representation is also strongly continuous under a certain hypothesis. In the sequel we will refer to Remark 12.22(4) and the definition of *right representation* (Definition 12.23).

Theorem 14.41. *Let S be a quantum system described, in the algebraic formalism, by the C^*-algebra with unit $\mathfrak{A}_S$, and let G be a group with a representation*

$$\alpha : \mathsf{G} \ni g \mapsto \alpha_g$$

by *-automorphisms α_g of $\mathfrak{A}_S$. Suppose ω is a G-invariant algebraic state on S represented by α:

$$\omega(\alpha_g(a)) = \omega(a) \quad \text{for any } g \in \mathsf{G},\ a \in \mathfrak{A}_S \text{ with } a = a^*. \tag{14.10}$$

(a) *If $U_{\alpha_g} : \mathsf{H}_\omega \to \mathsf{H}_\omega$ is the unitary operator associated to the *-automorphism α_g by Theorem 14.38, for any $g \in \mathsf{G}$ the map*

$$\mathsf{G} \ni g \mapsto U_{\alpha_g} \tag{14.11}$$

is a unitary right representation of G.
(b) *If G is a topological group and $\mathsf{G} \ni g \mapsto \omega(a^* \alpha_g(a))$ is continuous for any given $a \in \mathfrak{A}_S$, the representation (14.11) is strongly continuous.*

Proof. Consider the operators U_{α_g} defined by Theorem 14.38. By assumption, since $U_{\alpha_h}\Psi_\omega = \Psi_\omega$:

$$U_{\alpha_g}U_{\alpha_h}\pi_\omega(a)\Psi_\omega = U_{\alpha_g}U_{\alpha_h}\pi_\omega(a)U_{\alpha_h}^{-1}\Psi_\omega = U_{\alpha_g}\pi_\omega(\alpha_h^{-1}(a))\Psi_\omega$$

$$= U_{\alpha_g}\pi_\omega(\alpha_h^{-1}(a))U_{\alpha_g}^{-1}\Psi_\omega = \pi_\omega(\alpha_g^{-1}(\alpha_h^{-1}(a)))\Psi_\omega = \pi_\omega(\alpha_{hg}^{-1}(a))\Psi_\omega$$

$$= U_{\alpha_{hg}}\pi_\omega(a)\Psi_\omega\,.$$

As $\pi_\omega(\mathfrak{A}_S)\Psi_\omega$ is dense in H_S, we have $U_{\alpha_g}U_{\alpha_h} = U_{\alpha_{hg}}$. Similarly we can prove $U_{\alpha_g}^{-1} = U_{\alpha_{g^{-1}}}$ and $U_{\alpha_e} = I$. In other terms (14.11) is a unitary right representation of G.

Now assume G is a topological group and $\mathsf{G} \ni g \mapsto \omega(a^*\alpha_g(a))$ continuous for every $a \in \mathfrak{A}_S$. By the GNS theorem, and the fact that $U_{\alpha_g}\Psi_\omega = \Psi_\omega$, this implies $\mathsf{G} \ni g \mapsto (\pi_\omega(a)\Psi_\omega|U_{\alpha_g}\pi_\omega(a)\Psi_\omega)$ is a continuous function. But $a \in \mathfrak{A}_S$ is generic, so we have proved that for every Φ in the dense space $\pi_\omega(\mathfrak{A}_S)\Psi_\omega$, $\mathsf{G} \ni g \mapsto (\Phi|U_{\alpha_g}\Phi)$ is continuous. Using that U_{α_g} is unitary, it is easy to see that, consequently:

$$||U_{\alpha_g}\Phi - U_{\alpha_{g'}}\Phi||^2 = ||U_{\alpha_{gg'^{-1}}}\Phi - \Phi||^2 \to 0$$

for $g \to g'$, and any Φ in the dense subspace $\pi_\omega(\mathfrak{A}_S)\Psi_\omega$. Therefore $\mathsf{G} \ni g \mapsto U_{\alpha_g}$ is strongly continuous on the dense space $\pi_\omega(\mathfrak{A}_S)\Psi_\omega$. This generalises to the generic vector $\Psi \in \mathsf{H}_S$ as follows. For every $\varepsilon > 0$ we can find $\Phi \in \pi_\omega(\mathfrak{A}_S)\Psi_\omega$ so that $||\Psi - \Phi|| < 2\varepsilon/3$. For such Φ, there is a neighbourhood $I_{g'}$ of g' in G such that $||U_{\alpha_g}\Phi - U_{\alpha_{g'}}\Phi|| < \varepsilon/3$ if $g \in I_{g'}$. Hence for $g' \in \mathsf{G}$ and any $\varepsilon > 0$ there is a neighbourhood $I_{g'}$ of g such that:

$$||U_{\alpha_g}\Psi - U_{\alpha_{g'}}\Psi|| \leq ||U_{\alpha_g}\Psi - U_{\alpha_g}\Phi|| + ||U_{\alpha_g}\Phi - U_{\alpha_{g'}}\Phi|| + ||U_{\alpha_{g'}}\Phi - U_{\alpha_{g'}}\Psi||$$

$$= ||\Psi - \Phi|| + ||U_{\alpha_g}\Phi - U_{\alpha_{g'}}\Phi|| + ||\Phi - \Psi|| < \varepsilon$$

for $g \in I_{g'}$. This ends the proof. □

Remark 14.42. **(1)** Obtaining a *right representation* simply depends on the fact Theorem 14.38 associates to the algebraic symmetry α the unitary operator U_α such that $\pi_\omega(\alpha(a)) = U_\alpha^{-1}\pi_\omega(a)U_\alpha$: we are reading the action of α as *dual action* on observables of a Wigner (or Kadison) symmetry, for which the emphasis is put on state vectors rather than observables.The symmetry acts as $U_\alpha\Psi$ on states. We could also decide to make the algebraic symmetry α act by $\pi_\omega(\alpha(a)) = U_\alpha\pi_\omega(a)U_\alpha^{-1}$, calling U_α the previous U_α^{-1}. Under the above theorem, we would now obtain $\mathsf{G} \ni g \mapsto U_{\alpha_g}$ in the form we are more accustomed to. In the algebraic formulation, where the focus is more on observables than on states, one normally follows this second road.
(2) In case G is the topological group $\mathbb{R}$, and $\{\alpha_t\}_{t\in\mathbb{R}}$ satisfies part (b) for the invariant state ω, Stone's theorem warrants that the one-parameter unitary group $\{U_t\}_{t\in\mathbb{R}}$ representing $\mathbb{R}$ ($\pi_\omega(\alpha_t(a)) = U_t^{-1}\pi_\omega(a)U_t$) admits a self-adjoint generator $H : D(H) \to \mathsf{H}_\omega$ defined on the dense domain $D(H) \subset \mathsf{H}_\omega$, for which $U_t = e^{itH}$. Then we may think $\{\alpha_t\}_{t\in\mathbb{R}}$ as a one-parameter group of *-automorphisms that describes the evolution of the system with the parameter t as time.

By Stone's theorem, the first condition in (14.7) implies $\Psi_\omega \in D(H)$ and $H\Psi_\omega = 0$, so $0 \in \sigma_p(H)$. If $\sigma(H) \subset [0, +\infty)$ and $dim(Ker(H)) = 1$, ω is called a **ground state** for the time evolution $\{\alpha_t\}_{t\in\mathbb{R}}$. ■

This page intentionally left blank

Appendix A

Order relations and groups

A.1 Order relations, posets, Zorn's lemma

A relation $\geq$ on an arbitrary set X is called a **partial order (relation)** if it is *reflexive* ($x \geq x$, $\forall x \in \mathsf{X}$), *transitive* ($x \geq y \geq z \Rightarrow x \geq z$, $\forall x, y, z \in \mathsf{X}$) and *skew-symmetric* ($x \geq y \geq x \Rightarrow x = y$, $\forall x, y \in \mathsf{X}$). The pair $(\mathsf{X}, \geq)$ is then said a **partially ordered set** (shortened to **poset**).

An equivalent writing of $a \geq b$ is $b \leq a$.

The partial order $\geq$ is a **total order** if, further, either $x \geq y$ or $y \geq x$ for any $x, y \in \mathsf{X}$.

If $(\mathsf{X}, \geq)$ is a partially ordered set:

(i) $Y \subset \mathsf{X}$ is **upper bounded** (resp. **lower bounded**) if it admits an **upper bound** (**lower bound**), i.e. $x \in \mathsf{X}$ such that $x \geq y$ ($y \geq x$) for any $y \in Y$;
(ii) an element $x_0 \in \mathsf{X}$ for which there exists *no* element $x \neq x_0$ in X such that $x \geq x_0$ is **maximal** in X. (Note that for us a maximal element in X may not be an upper bound in X).

If $(\mathsf{X}, \geq)$ is a poset, a subset $Y \subset \mathsf{X}$ is (**totally**) **ordered** if the relation $\geq$, restricted to $Y \times Y$, is a total order.

Recall that Zorn's lemma is an equivalent statement to the Axiom of Choice (also known as Zermelo's axiom).

Theorem A.1 ("Zorn's lemma"). *If any ordered subset in a poset* $(\mathsf{X}, \geq)$ *is upper bounded,* X *admits a maximal element.*

Useful among the various notions on posets $(\mathsf{X}, \geq)$ are those of supremum and infimum:

(i) a is called **least upper bound** (or **supremum**, or just **sup**) of the set $A \subset \mathsf{X}$, written $a = \sup A$, if a is an upper bound of A and any other upper bound a' of A satisfies $a \leq a'$;
(ii) a is called **greatest lower bound**, (or **infimum** or **inf**) of $A \subset \mathsf{X}$, written $a = \inf A$, if a is a lower bound for A and any other lower bound a' satisfies $a' \leq a$;

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7, © Springer-Verlag Italia 2013

It is immediate to see that any subset $A \subset \mathsf{X}$ has at most one least upper bound and one greatest lower bound.

A.2 Round-up on group theory

A group is an algebraic structure $(\mathsf{G}, \circ)$ consisting in a set G and an operation $\circ : \mathsf{G} \times \mathsf{G} \to \mathsf{G}$ (the composition law, often called **product**) satisfying three properties:

(1) $\circ$ is *associative*

$$g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3 \,, \quad \text{for any} \quad g_1, g_2, g_3 \in \mathsf{G} \,;$$

(2) there exists an element $e \in \mathsf{G}$, called **identity** or **neutral** element, such that

$$e \circ g = g \circ e = g \,, \quad \text{for any} \quad g \in \mathsf{G} \,;$$

(3) each element $g \in \mathsf{G}$ admits an **inverse**, i.e.

$$\text{for any} \quad g \in \mathsf{G} \quad \text{there exists} \quad g^{-1} \in \mathsf{G} \ \text{ such that } \ g \circ g^{-1} = g^{-1} \circ g = e \,.$$

The identity and the inverse to a given element are easily seen to be unique.

A group $(\mathsf{G}, \circ)$ is **commutative** or **Abelian** if $g \circ g' = g' \circ g$ for any $g, g' \in \mathsf{G}$; otherwise it is **noncommutative** or **non-Abelian**.

A subset $\mathsf{G}' \subset \mathsf{G}$ in a group is a **subgroup** if it becomes a group with the product of G restricted to $\mathsf{G}' \times \mathsf{G}'$. A subgroup N in a group G is **normal** if it is invariant under **conjugation**, i.e. for any $n \in N$ and $g \in \mathsf{G}$ the conjugate element $g \circ n \circ g^{-1}$ belongs to N.

If N is a normal subgroup in G, then G/N denotes the quotient, i.e. the set of equivalence classes in G with respect to the equivalence relation $g \sim g' \Leftrightarrow g = ng'$ for some $n \in N$. It is easy to prove that G/N inherits a natural group structure from G.

The **centre** Z of G is the commutative subgroup of G made by elements z that commute with every element of G. In other words, $z \in Z \Leftrightarrow z \circ g = g \circ z$ for any $g \in \mathsf{G}$.

If $(\mathsf{G}_1, \circ_1)$ and $(\mathsf{G}_2, \circ_2)$ are two groups, a **(group) homomorphism** from G_1 to G_2 is a map $h : \mathsf{G}_1 \to \mathsf{G}_2$ that *preserves the groups' structures*, i.e.:

$$h(g \circ_1 g') = h(g) \circ_2 h(g') \quad \text{for any} \quad g, g' \in \mathsf{G}_1 \,.$$

With the obvious notation it is clear that $h(e_1) = e_2$ and $h(g^{-1_1}) = (h(g))^{-1_2}$ for any $g \in \mathsf{G}_1$.

The **kernel** $Ker(h) \subset \mathsf{G}$ of a homomorphism $h : \mathsf{G} \to \mathsf{G}'$ is the pre-image under h of the identity e' of G', i.e. the set of elements g such that $h(g) = e'$. Notice $Ker(h)$ is a normal subgroup. Clearly h is one-to-one if and only if its kernel contains the identity of G only. It turns out that the image $h(\mathsf{G})$ of a homomorphism $h : \mathsf{G} \to \mathsf{G}'$ is a subgroup of G' isomorphic to $\mathsf{G}/Ker(h)$.

A **group isomorphism** is a *bijective* group homomorphism. An isomorphism $h : \mathsf{G} \to \mathsf{G}$ is an **automorphism** of G. The set $Aut(\mathsf{G})$ of automorphisms of G is itself a group under composition of maps.

If G_1, G_2 are groups, the **direct product** $\mathsf{G}_1 \otimes \mathsf{G}_2$ is a group with the following structure. The elements of $\mathsf{G}_1 \otimes \mathsf{G}_2$ are pairs (g_1, g_2) of the *Cartesian product* of the *sets* G_1, G_2. The composition law is

$$(g_1, g_2) \circ (f_1, f_2) := (g_1 \circ_1 f_1, g_2 \circ_2 f_2) \quad \forall (g_1, g_2), (f_1, f_2) \in \mathsf{G}_1 \times \mathsf{G}_2 .$$

The neutral element is obviously (e_1, e_2), where e_1, e_2 are the identities of G_1, G_2. Moreover, G_1 and G_2 can be identified with normal subgroups of $\mathsf{G}_1 \otimes \mathsf{G}_2$.

The ensuing generalisation of the notion of product plays a big role in physical applications. Let $(\mathsf{G}_1, \circ_1)$, $(\mathsf{G}_2, \circ_2)$ be groups and suppose that for any $g_1 \in \mathsf{G}_1$ there is a group isomorphism $\psi_{g_1} : \mathsf{G}_2 \to \mathsf{G}_2$ such that:

(i) $\psi_{g_1} \circ \psi_{g_1'} = \psi_{g_1 \circ_1 g_1'}$;
(ii) $\psi_{e_1} = id_{\mathsf{G}_2}$;

where $\circ$ is the composition of functions and e_1 the neutral element in G_1. (Equivalently, $\psi_g \in Aut(\mathsf{G}_2)$ for any $g \in \mathsf{G}_1$, and the map $\mathsf{G}_1 \ni g \mapsto \psi_g$ is a group homomorphism from G_1 to $Aut(\mathsf{G}_2)$.) We can endow the Cartesian product $\mathsf{G}_1 \times \mathsf{G}_2$ with a group structure simply by defining the composite of $(g_1, g_2), (f_1, f_2) \in \mathsf{G}_1 \times \mathsf{G}_2$ as

$$(g_1, g_2) \circ_\psi (f_1, f_2) := (g_1 \circ_1 f_1,\, g_2 \circ_2 \psi_{g_1}(f_2)) .$$

The operation is well defined, so $(\mathsf{G}_1 \otimes_\psi \mathsf{G}_2, \circ_\psi)$ is a group called the **semidirect product** of G_1 and G_2 by ψ. The order of the factors in the product is clearly relevant.

Looking at the semidirect product $(\mathsf{G} \otimes_\psi \mathsf{N}, \circ_\psi)$ we could prove N is a normal subgroup of $\mathsf{G} \otimes_\psi \mathsf{N}$, and

$$\psi_g(n) = g \circ_\psi n \circ_\psi g^{-1} \quad \text{for any } g \in \mathsf{G},\, n \in \mathsf{N}.$$

There is also a converse of sorts. Consider a group $(\mathsf{H}, \circ)$, let G be a subgroup of H and N a normal subgroup. Assume $N \cap G = \{e\}$, e being the identity of H. Suppose also $\mathsf{H} = GN$, meaning that for any $h \in \mathsf{H}$ there exist $g \in G$ and $n \in N$ such that $h = gn$. Then one can prove that the pair (g, n) is uniquely determined by h, and H is isomorphic to the semidirect product $G \otimes_\psi N$ with

$$\psi_g(n) := g \circ h \circ g^{-1} \quad \text{for any } g \in G,\, n \in N.$$

If now V is a vector space (real or complex), $GL(\mathsf{V})$ denotes the group of bijective linear maps $f : \mathsf{V} \to \mathsf{V}$ with the usual composition law. $GL(\mathsf{V})$ is called the **(general) linear group** of V.

If $\mathsf{V} := \mathbb{R}^n$ or $\mathbb{C}^n$ then $GL(\mathsf{V})$ is denoted by $GL(n, \mathbb{C})$ or $GL(n, \mathbb{R})$, respectively.

Let us define linear representations of a group. Take $(\mathsf{G}, \circ)$ a group and V a vector space. A **(linear) representation** of G on V is a homomorphism $\rho : \mathsf{G} \to GL(\mathsf{V})$.

A representation $\rho : \mathsf{G} \to GL(\mathsf{V})$ is called:

(1) **faithful** if it is injective;
(2) **free** if the subgroup made of elements h_v such that $\rho(h_v)v = v$ is trivial for any $v \in \mathsf{V} \setminus \{0\}$, i.e. it contains only the neutral element of G;
(3) **transitive** if, for any $v, v' \in \mathsf{V} \setminus \{0\}$ there exists $g \in \mathsf{G}$ with $v' = \rho(g)v$;
(4) **irreducible** if no proper subspace $S \subset \mathsf{V}$ exists that is **invariant** under the action of $\rho(\mathsf{G})$, i.e. $\rho(g)S \subset S$ for any $g \in \mathsf{G}$.

In case V is a Hilbert or Banach space and ρ defines *bounded operators on the entire* V, the representation is said irreducible if there are no *closed* $\rho(\mathsf{G})$-invariant subspaces in V.

Appendix B

Elements of differential geometry

Let $n, m = 1, 2, \ldots$, $k = 0, 1, \ldots$ be fixed integers and $\Omega \subset \mathbb{R}^n$ an open non-empty set. A map $f : \Omega \to \mathbb{R}^m$ is **of class** C^k (or simply C^k), written $f \in C^k(\Omega; \mathbb{R}^n)$, if all partial derivatives of the components of f are continuous up to order k included. Conventionally, $C^k(\Omega) := C^k(\Omega; \mathbb{R})$.

A function $f : \Omega \to \mathbb{R}^m$ is (of class) C^∞, or **smooth**, if it is C^k for any $k = 0, 1, \ldots$, so one defines

$$C^\infty(\Omega; \mathbb{R}^n) := \bigcap_{k=0,1,\ldots} C^k(\Omega; \mathbb{R}^n) .$$

Again, $C^\infty(\Omega) := C^\infty(\Omega; \mathbb{R})$. Eventually, $f : \Omega \to \mathbb{R}^m$ is C^ω or **real-analytic** if it is C^∞ and it admits a Taylor expansion (in several real variables) at any $p \in \Omega$, on some open ball around p of finite radius, contained in Ω. Usually, when the order k of differentiability is not mentioned explicitly it means that $k = \infty$.

Notation B.1. In this section upper indices denote coordinates of $\mathbb{R}^n$ and components of (contravariant) vectors. Thus the standard coordinates on $\mathbb{R}^n$ will be denoted by $x^1, \ldots, x^n$, instead of $x_1, \ldots, x_n$. ■

B.1 Smooth manifolds, product manifolds, smooth functions

The most general and powerful tool apt to describe the features of spacetime, three-dimensional physical space, and the abstract space of physical systems in classical theories, is the notion of *smooth manifold*. In practice a smooth manifold is a collection of objects, generally called *points*, that admits local coordinates identifying points with n-tuples of $\mathbb{R}^n$.

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7, © Springer-Verlag Italia 2013

Definition B.2. *Let $n = 1,2,3,\cdots$ and $k = 1,2,...,\infty,\omega$ be fixed numbers. A C^k* **manifold of dimension** *n is a set M, whose elements are called* **points***, equipped with the geometric structure defined below.*

(1) *M has a* **differentiable structure** *$\mathscr{A} = \{(U_i,\phi_i)\}_{i\in I}$* **of class** *$C^k$, that is a collection of pairs (U_i,ϕ_i), called* **local charts***, where U_i is a subset in M and ϕ_i a map from U_i to $\mathbb{R}^n$ (the* **local coordinate system** *or* **local frame***) such that:*

- (i) *$\cup_{i\in I} U_i = M$, any ϕ_i is injective and $\phi_i(U_i)$ is open in $\mathbb{R}^n$ (so M is called an n-dimensional manifold, or just n-manifold);*
- (ii) *local charts in $\mathscr{A}$ must be pairwise C^k-compatible. Two injective maps $\phi : U \to \mathbb{R}^n$, $\psi : V \to \mathbb{R}^n$ with $U,V \subset M$ are C^k-***compatible** *if either $U \cap V = \varnothing$, or $U \cap V \neq \varnothing$ and the maps $\phi\circ\psi^{-1} : \psi(U\cap V) \to \phi(U\cap V)$, $\psi\circ\phi^{-1} : \phi(U\cap V) \to \psi(U\cap V)$ are both C^k;*
- (iii) *$\mathscr{A}$ is* **maximal***, i.e.: if $U \subset M$ is open and $\phi : U \to \mathbb{R}^n$ compatible with every local chart of $\mathscr{A}$, then $(U,\phi) \in \mathscr{A}$.*

(2) *Topological requirements:*

- (i) *M is a second-countable Hausdorff space;*
- (ii) *M is, by way of $\mathscr{A}$, locally homeomorphic to $\mathbb{R}^n$. In other terms, if $(U,\phi) \in \mathscr{A}$ then U is open and $\phi : U \to \phi(U)$ is a homeomorphism.*

A smooth C^ω manifold is more often called **real-analytic manifold**.

Remark B.3. **(1)** Every local chart (U,ϕ) enables us to assign n real numbers $(x^1_p,\cdots,x^n_p) = \phi(p)$ bijectively to every point p of U. The entries of the n-tuple are the **coordinates** of p in the local chart (U,ϕ). Points in U are thus in one-to-one correspondence with n-tuples of $\phi(U) \subset \mathbb{R}^n$.
(2) If $U \cap V \neq \varnothing$, the compatibility of local charts (U,ϕ), (V,ψ) implies that the Jacobian matrix of $\phi\circ\psi^{-1}$ is invertible and so has everywhere non-zero determinant. Conversely, if $\phi\circ\psi^{-1} : \psi(U\cap V) \to \phi(U\cap V)$ is bijective, of class C^k, and with non-vanishing Jacobian determinant on $\psi(U\cap V)$, then also $\psi\circ\phi^{-1} : \phi(U\cap V) \to \psi(U\cap V)$ is C^k and the local charts are compatible. The proof can be found in the renowned [CoFr98II].

Theorem B.4 (Implicit function theorem). *Let $D \subset \mathbb{R}^n$ be open, non-empty, and $f : D \to \mathbb{R}^n$ a C^k function for some $k = 1,2,\ldots,\infty$. If the Jacobian of f at $p \in D$ has non-zero determinant there exist open neighbourhoods $U \subset D$ of p and V of $f(p)$ such that: (i) $f\restriction_U : U \to V$ is bijective, (ii) the inverse $f\restriction_U^{-1} : V \to U$ is C^k.*

(3) The topological requirements in (2)(i) (valid for the standard topology of $\mathbb{R}^n$) are technical and guarantee unique solutions to differential equations on M (necessary in physics when the equations describe the evolution of physical systems) and the existence of integrals on M. Condition (2)(ii) intuitively says that M is, around any point, "continuous" like $\mathbb{R}^n$. Standard counterexamples show that the Hausdorff property of $\mathbb{R}^n$ is not carried over to M by local homeomorphisms, so it must be imposed explicitly.

(4) Let M be a second-countable Hausdorff space. A collection of local charts $\mathscr{A}$ on M satisfying (i) and (ii) in (1), but not necessarily (iii), plus (ii) in (2) is called a C^k **atlas** on the n-manifold M. It is not hard to see that any atlas $\mathscr{A}$ on M is contained in some maximal atlas. Two atlases on M such that every chart of one is compatible with any chart of the other induce the same differentiable structure on M. Thus to assign a differentiable structure it suffices to prescribe a non-maximal atlas, one of the many that determine it. The unique differentiable structure associated to a given atlas is said to be **induced** by the atlas.
(5) If $1 \leq k < \infty$ there might be superfluous charts in the differentiable structure (only a finite number!), eliminating which gives a C^∞ atlas. ■

Examples B.5. (1) The simplest examples of differentiable manifolds, of class C^∞ and dimension n, are non-empty open subsets of $\mathbb{R}^n$ (including $\mathbb{R}^n$ itself) with standard differentiable structure determined by the identity map (the inclusion, alone, defines an atlas).
(2) Consider the unit sphere $\mathbb{S}^2$ in $\mathbb{R}^3$ (with topology inherited from $\mathbb{R}^3$) centred at the origin:

$$\mathbb{S}^2 := \left\{(x^1,x^2,x^3) \in \mathbb{R}^3 \;\middle|\; (x^1)^2+(x^2)^2+(x^3)^2 = 1\right\}$$

in canonical coordinates x^1,x^2,x^3 of $\mathbb{R}^3$. It has dimension 2 and a smooth structure induced by $\mathbb{R}^3$ by defining an atlas with 6 local charts $(\mathbb{S}^2_{(i)\pm}, \phi^{(i)}_\pm)$ $(i = 1,2,3)$ as follows. Take the axis x^i $(i = 1,2,3)$ and the pair of open hemispheres $\mathbb{S}^2_{(i)\pm}$ with south-north direction given by x^i, and consider local charts $\phi^{(i)}_\pm : \mathbb{S}^2_{(i)\pm} \to \mathbb{R}^2$ that map $p \in \mathbb{S}^2_{(i)\pm}$ to its coordinates on the plane $x^i = 0$. It can be proved (see below) that $\mathbb{S}^2$ cannot be covered by a single (global) chart, in contrast to $\mathbb{R}^3$ (or any open subspace). This proves that the class of smooth manifolds does not reduce to open non-empty subsets of $\mathbb{R}^n$, and hence is quite interesting. A similar example is the circle in $\mathbb{R}^2$. ■

Given C^k manifolds M and N of respective dimensions m, n, we can construct a third C^k manifold of dimension $m+n$ over the topological product $M \times N$. (The resulting space will be Hausdorff and second-countable.) This is called *product manifold* of M and N, and denoted simply by $M \times N$. The structure described herebelow is called *product structure*. Given local charts (U,ϕ) on M and (V,ψ) on N it is immediate to see

$$U \times V \ni (p,q) \mapsto (\phi(p), \psi(q)) =: \phi \oplus \psi(p,q) \in \mathbb{R}^{m+n} \tag{B.1}$$

is a local homeomorphism. If (U',ϕ') and (V',ψ') are other charts, compatible with the previous ones, the charts $(U \times V, \phi \oplus \psi)$ and $(U' \times V', \phi' \oplus \psi')$ are obviously compatible. As (U,ϕ) and (V,ψ) vary on M and N the charts $(U \times V, \phi \oplus \psi)$ define an atlas on $M \times N$. The structure this atlas generates is, by definition, the product structure.

Definition B.6. *Given C^k manifolds M, N of dimension m, n, the* **product manifold** *is the set $M \times N$ equipped with product topology and C^k structure induced by the local charts $(U \times V, \phi \oplus \psi)$ as of (B.1), when (U,ϕ), (V,ψ) vary on M, N.*

Since a manifold is locally indistinguishable from $\mathbb{R}^n$, the differentiable structure allows to make sense of *differentiable functions* defined on a manifold other than $\mathbb{R}^n$ or subsets. The idea is simple: reduce locally to the standard notion on $\mathbb{R}^n$ using the local charts that cover the manifold.

Definition B.7. *Let M,N be manifolds of dimensions m,n and class C^p, C^q respectively ($p,q \geq 1$). A continuous map $f: M \to N$ is said C^k ($0 \leq k \leq p,q$, possibly $k = \infty$ or ω) if $\psi \circ f \circ \phi^{-1}$ is a C^k map from $\mathbb{R}^m$ to $\mathbb{R}^n$, for any choice of local charts (U,ϕ) on N and (V,ψ) on M.*
The collection of C^k functions from M to N, $k = 0,1,2,\ldots,\infty,\omega$ is denoted $C^k(M;N)$; if $N = \mathbb{R}$ one just writes $C^k(M)$.
A C^k **diffeomorphism** *$f: M \to N$ is a bijective C^k map with C^k inverse. If there is a C^k diffeomorphism f mapping M to N, the two manifolds are called* **diffeomorphic** *(under f).*

Remark B.8. (1) Notice how we allowed for differentiable maps of class C^0, which are actually just continuous maps (like C^0 diffeomorphisms are just homeomorphisms). Every C^k diffeomorphism is clearly a homeomorphism, which explains why there cannot exist any diffeomorphism between $\mathbb{S}^2$ and (a subset of) $\mathbb{R}^2$, for the former is compact, the latter not. Consequently, the sphere $\mathbb{S}^2$ does not admit global charts.
(2) For $f: M \to N$ to be C^p it is enough that $\psi \circ f \circ \phi^{-1}$ is C^k for any local charts $(U,\phi),(V,\psi)$ in the given atlases, without having to check the condition for *every possible* local charts on the manifolds. ∎

A useful notion is that of *embedded submanifold*. $\mathbb{R}^n$ is an embedded submanifold in $\mathbb{R}^m$ if $m > n$. In the canonical coordinates $x^1,\cdots,x^m$ on $\mathbb{R}^m$, $\mathbb{R}^n$ is identified with the subspace given by equations $x^{n+1} = \cdots = x^m = 0$, while the first n coordinates of $\mathbb{R}^m$, $x^1,\cdots,x^n$, are identified with the standard coordinates on $\mathbb{R}^n$. Now the idea is to replace $\mathbb{R}^n$, $\mathbb{R}^m$ using local frames, and generalise to manifolds N, M.

Definition B.9. *Let M be a C^k ($k \geq 1$) manifold of dimension $m > n$. An* **embedded C^k submanifold of M of dimension** *n is the following n-manifold N of class C^k.*

(a) *N is a subset in M with induced topology.*
(b) *The differentiable structure di N is given by the atlas $\{(U_i,\phi_i)\}_{i\in I}$ where:*

(i) *$U_i = V_i \cap N$, $\phi_i = \psi\restriction_{V_i\cap N}$ for a suitable local chart (V_i,ϕ_i) on M;*
(ii) *in the frame $x^1,\cdots,x^m$ associated to (V_i,ϕ_i), the set $V_i \cap N$ is determined by $x^{n+1} = \cdots = x^m = 0$, and the remaining coordinates $x^1,\cdots,x^n$ are the local framing associated to ϕ_i.*

To finish we state an important result (see [doC92, Wes78] for example) to decide when a subset in a manifold is an embedded submanifold. The proof is straightforward from Dini's theorem [CoFr98II].

Theorem B.10 (On regular values). *Let M be a C^k manifold of dimension m. Consider the set*

$$N := \{p \in M \mid f_j(p) = v_j\,,\ j = 1,\cdots,c\}$$

determined by $c(< m)$ constants v_j and c functions $f_j : M \to \mathbb{R}$ of class C^k. Suppose that around each point $p \in N$ there exists a local chart (U, ϕ) on M such that the Jacobian matrix $\partial(f_j \circ \phi^{-1})/\partial x^i|_{\phi(p)}$ ($j = 1, \cdots, c$, $i = 1, \cdots, m$) has rank r. Then N is an embedded C^k submanifold in M of dimension $n := m - c$.
In particular, if the square $c \times c$ matrix

$$\frac{\partial f_j \circ \phi^{-1}}{\partial x^k}, \quad j = 1, \ldots, c, \ k = m-c+1, m-c+2, \ldots, m$$

is non-singular at $\phi(p)$, $p \in N$, then the first n coordinates $x^1, \ldots, x^n$ define a frame system around p in N.

B.2 Tangent and cotangent spaces. Covariant and contravariant vector fields

Ler M be C^k manifold of dimension n ($k \geq 1$). Consider the space $C^k(M)$ as an $\mathbb{R}$-vector space with linear combinations

$$(af + bg)(p) := af(p) + bg(p), \quad \text{for any } p \in M$$

where $a, b \in \mathbb{R}$, $f, g \in C^k(M)$. Given a point $p \in M$, a **derivation** at p is an $\mathbb{R}$-linear map $L_p : C^k(M) \to \mathbb{R}$ satisfying the *Leibniz rule*:

$$L_p(fg) = f(p)L_p(g) + g(p)L_p(f), \quad f, g \in C^k(M). \tag{B.2}$$

A linear combination $aL_p + bL'_p$ of derivations at p ($a, b \in \mathbb{R}$),

$$(aL_p + bL'_p)(f) := aL_p(f) + bL'_p(f), \quad f, g \in C^k(M),$$

is still a derivation. Hence derivations at p form a vector space over $\mathbb{R}$, which we denote $\mathscr{D}^k_p$. Every local chart (U, ϕ) with $U \ni p$ automatically gives n derivations at p, as follows. If $x^1, \ldots, x^n$ are coordinates associated to ϕ, define the kth derivation to be

$$\left.\frac{\partial}{\partial x^k}\right|_p : f \mapsto \left.\frac{\partial f \circ \phi^{-1}}{\partial x^k}\right|_{\phi(p)}, \quad f, g \in C^1(M). \tag{B.3}$$

If 0 is the null derivation and $c^1, c^2, \cdots, c^n \in \mathbb{R}$ satisfy $\sum_{k=1}^n c^k \left.\frac{\partial}{\partial x^k}\right|_p = 0$, we choose a differentiable function conciding with the coordinate map x^l on an open neighbourhood of p (whose closure is in U) and vanishing outside. Then the n derivations $\left.\frac{\partial}{\partial x^k}\right|_p$ at p are *linearly independent*: $\sum_{k=1}^n c^k \left.\frac{\partial}{\partial x^k}\right|_p f = 0$ imples $c^l = 0$. Since we are free to choose l arbitrarily, every coefficient c^r is zero for $r = 1, 2, \ldots, n$. Hence the n derivations $\left.\frac{\partial}{\partial x^k}\right|_p$ form a basis for an n-dimensional subspace of $\mathscr{D}^k_p$ (actually if $k = \infty$ the subspace coincides with $\mathscr{D}_p^\infty$). Changing chart to (V,ψ), $V \ni p$, with frame $y^1,\dots,y^n$, the new derivations are related to the old ones by:

$$\left.\frac{\partial}{\partial y^i}\right|_p = \sum_{k=1}^n \left.\frac{\partial x^k}{\partial y^i}\right|_{\psi(p)} \left.\frac{\partial}{\partial x^k}\right|_p . \tag{B.4}$$

The proof is direct from the definitions. Because the Jacobian $\left.\frac{\partial x^k}{\partial y^i}\right|_{\psi(p)}$ is invertible by definition of chart, the subspace of $\mathscr{D}_p^k$ spanned by the $\left.\frac{\partial}{\partial y^i}\right|_p$ coincides with the span of the $\left.\frac{\partial}{\partial x^k}\right|_p$. The subspace is thus *intrinsically defined*.

Definition B.11. *Let M be an n-dimensional C^k manifold ($k \geq 1$), and fix a point $p \in M$. The vector subspace of derivations at $p \in M$ generated by the n derivations $\left.\frac{\partial}{\partial x^k}\right|_p$, $k = 1,2,\dots,n$, in any local coordinate system (U,ϕ) with $U \ni p$, is called* **tangent space of** *M* **at** *p and is written T_pM. The elements of the tangent space at p are the* **tangent vectors at** *p to M. Tangent vectors are examples of* **contravariant vectors**.

We recall that the space V^* of linear maps from a real vector space V to $\mathbb{R}$ is called **dual space** to V. If the dimension of V is finite, so is the dimension of V^*, for they coincide. In particular, if $\{e_i\}_{i=1,\dots,n}$ is a basis of V, the **dual basis** in V^* is the basis $\{e^{*j}\}_{j=1,\dots,n}$ defined via: $e^{*j}(e_i) = \delta_i^j$, $i,j = 1,\dots,n$, by linearity. With $f \in V^*$, $v \in V$, one uses the notation $\langle v, f\rangle := f(v)$.

Definition B.12. *Let M be an n-dimensional C^k manifold ($k \geq 1$), $p \in M$ a given point. The dual space to T_pM is called* **cotangent space of** *M* **at** *p, written T_p^*M. Points of the cotangent space at p are called* **cotangent vectors at** *p or* **1-forms** *at p, and are instances of* **covariant vectors** *(covectors). For any basis $\left.\frac{\partial}{\partial x^k}\right|_p$ of T_pM, the n elements of the dual basis are indicated by $dx^i|_p$. By definition*

$$\left\langle \frac{\partial}{\partial x^k}|_p, dx^i|_p \right\rangle = \delta_k^i .$$

Let us move on to *vector fields* on a manifold M.

Suppose M is an n-dimensional C^k manifold (including $k = \infty$ and $k = \omega$). A **contravariant C^r vector field**, $r = 0,1,\dots,k$, is a map assigning a vector $v(p) \in T_pM$ to any $p \in M$, so that for any local chart (U,ϕ) with coordinates $x^1,\dots,x^n$ where

$$v(q) = \sum_{i=1}^n v^i(x_q^1,\dots,x_q^n) \left.\frac{\partial}{\partial x^i}\right|_q ,$$

the n functions $v^i = v^i(x^1,\dots,x^n)$ are C^r on $\phi(U)$. Similarly, a **covariant C^r vector field**, $r = 0,1,\dots,k$ is a map sending $p \in M$ to a covector $\omega(p) \in T_p^*M$, so that for any local chart (U,ϕ) with coordinates $x^1,\dots,x^n$ where

$$\omega(q) = \sum_{i=1}^{n} v_i(x^1_q,\dots,x^n_q)\, dx^i\big|_q \,,$$

the n functions $\omega_i = \omega_i(x^1,\dots,x^n)$ are C^r on $\phi(U)$.

Remarks B.13. Take $v \in T_pM$ and two local charts (U,ϕ), (V,ψ) with $U \cap V \ni p$ and respective coordinates $x^1,\dots,x^n$, $x'^1,\dots,x'^n$. Then $v = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}\Big|_p = \sum_{j=1}^n v'^j \frac{\partial}{\partial x'^j}\Big|_p$. Hence $\sum_i^n v^i \frac{\partial}{\partial x^i}\Big|_p = \sum_{j,i=1}^n v'^j \frac{\partial x^i}{\partial x'^j}\Big|_{\psi(p)} \frac{\partial}{\partial x^i}\Big|_p$, so $\sum_{i=1}^n \left(v^i - \sum_{j=1}^n \frac{\partial x^i}{\partial x'^j}\Big|_{\psi(p)} v'^j \right) \frac{\partial}{\partial x^i}\Big|_p = 0$. Since the derivations $\frac{\partial}{\partial x^i}\Big|_p$ are linearly independent, we conclude that the components of a tangent vector in T_pM transform, under coordinate change, as

$$v^i = \sum_{j=1}^{n} \frac{\partial x^i}{\partial x'^j}\bigg|_{\psi(p)} v'^j \,. \tag{B.5}$$

The same argument gives the formula for covariant vectors $\omega = \sum_{i=1}^n \omega_i\, dx^i\big|_p = \sum_{j=1}^n \omega'_j\, dx'^j\big|_p$, namely

$$\omega_i = \sum_{j=1}^{n} \frac{\partial x'^j}{\partial x^i}\bigg|_{\psi(p)} \omega'_j \,. \tag{B.6}$$

■

B.3 Differentials, curves and tangent vectors

Let $f : M \to \mathbb{R}$ be a C^r scalar field on the C^k n-manifold M, and assume $k \geq r > 1$. The **differential** df of f is the covariant vector field of class C^{r-1}

$$df|_p = \sum_{i=1}^{n} \frac{\partial f}{\partial x^i}\bigg|_{\psi(p)} dx^i|_p$$

in any local chart (U,ψ).

Consider a C^r curve inside the C^k manifold M ($r = 0,1,\dots,k$), i.e. a C^r function $\gamma : I \to M$ where $I \subset \mathbb{R}$ is an open interval thought of as a submanifold in $\mathbb{R}$. Assume explicitly that $r > 1$. We can define the *tangent vector* to γ at $p \in \gamma(I)$ by

$$\dot\gamma(p) := \sum_{i=1}^{n} \frac{dx^i}{dt}\bigg|_{t_p} \frac{\partial}{\partial x^i}\bigg|_p \,,$$

where $\gamma(t_p) = p$, in any local chart around p. The definition does *not* depend on the chart. Had we defined

$$\dot\gamma'(p) := \sum_{j=1}^{n} \frac{dx'^j}{dt}\bigg|_{t_p} \frac{\partial}{\partial x'^j}\bigg|_p$$

in another frame system around p, using (B.5) would have given

$$\dot{\gamma}(p) = \dot{\gamma}'(p)\,.$$

So we have this definition.

Definition B.14. *A C^r* **curve***, $r = 0, 1, \ldots, k$, in the n-dimensional C^k manifold M is a C^r map $\gamma : I \to M$, where $I \subset \mathbb{R}$ is an open interval (embedded in $\mathbb{R}$). When $r > 1$, the* **tangent vector** *to γ at $p = \gamma(t_p)$, $t_p \in I$, is the vector $\dot{\gamma}(p) \in T_pM$ given by*

$$\dot{\gamma}(p) := \sum_{i=1}^{n} \left.\frac{dx^i}{dt}\right|_{t_p} \left.\frac{\partial}{\partial x^i}\right|_p, \tag{B.7}$$

in any local framing around p.

B.4 Pushforward and pullback

Let M and N be manifolds of dimensions m and n, and $f : N \to M$ a function (all at least C^1). Given a point $p \in N$ consider local charts (U, ϕ) around p in N and (V, ψ) around $f(p)$ in M. Indicate by $(y^1, \ldots, y^n)$ the coordinates on U, by $(x^1, \ldots, x^m)$ those on V and introduce maps $f^k(y^1, \ldots, y^n) = y^k(f \circ \phi^{-1})$, $k = 1, \ldots, m$. Now define:

(i) the **pushforward** $df_p : T_pN \to T_{f(p)}M$, in coordinates:

$$df_p : T_pN \ni \sum_{i=1}^{n} u^i \left.\frac{\partial}{\partial y^i}\right|_p \mapsto \sum_{j=1}^{m} \left(\sum_{i=1}^{n} \left.\frac{\partial f^j}{\partial y^i}\right|_{\phi(p)} u^i \right) \left.\frac{\partial}{\partial x^j}\right|_p; \tag{B.8}$$

(ii) the **pullback** $f_p^* : T_{f(p)}^*M \to T_p^*N$, in coordinates:

$$f_p^* : T_{f(p)}^*M \ni \sum_{j=1}^{m} \omega_j dx^j|_{f(p)} \mapsto \sum_{i=1}^{n} \left(\sum_{j=1}^{m} \left.\frac{\partial f^j}{\partial y^i}\right|_{\phi(p)} \omega_j \right) dy^i|_p\,. \tag{B.9}$$

It is not hard to see they *do not depend on local frame systems.* The pushforward is also written $f_{p*} : T_pN \to T_{f(p)}M$.

References

[AbCi97] Abbati M.C., Cirelli R.: Metodi Matematici per la fisica, Operatori Linerari negli spazi di Hilbert. Città Studi Edizioni, Torino (1997).

[Alb94] Albert D.-Z.: Quantum Mechanics and Experience. Harvard University Press (1994).

[AnMo12] Annigoni E., Moretti V.: Mass operator and dynamical implementation of mass superselection rule. Hannales H. Poincaré, *in print* (2012).

[Ara09] Araki H.: Mathematical Theory of Quantum Fields. Oxford University Press, Oxford (2009).

[Bar54] Bargmann V.: On Unitary Ray Representations of Continuous groups. Ann. Math. **59**, 1 (1954).

[Bar61] Bargmann V.: On a Hilbert Space of Analytic Functions and an Associated Integral Transform. Communications on Pure and Applied Mathematics **14**, 187–214 (1961).

[BGJKS00] Blanchard Ph., Giulini D., Joos E., Kiefer C., Stamatescu I.-O. (eds.): Decoherence: Theoretical, Experimental, and Conceptual Problems. Lecture Notes in Physics. Springer-Verlag, Heidelberg Berlin (2000).

[BaRa86] Barut A.O., Raczka R.: Theory of group representations and applications (Second Revised Edition). World Scientific, Singapore (1986).

[Bel64] Bell J.S.: On the Einstein Podolsky Rosen Paradox, Physics **1**, 195–200 (1964).

[BeCa81] Beltrametti E.G., Cassinelli G.: The logic of quantum mechanics. Encyclopedia of Mathematics and its Applications, Vol. 15. Addison-Wesley, Reading MA (1981).

[Ber66] Berberian S.K.: Notes on Spectral Theory. Van Nostrand Company, Inc., Princeton (1966).

[BGP07] Bär C., Ginous N., Pfäffe F.: Wave Equations on Lorentzian Manifolds and Quantization. Lecture Notes in Mathematics and Physics. European Mathematical Society (2007).

[BiSo87] Birman M.S., Solomjak M.Z.: Spectral Theory of Self-Adjoint Operators in Hilbert Space. Reidel Publishing Company, Dordrecht (1987).

[Bon97] Boniolo G. (ed.): Filosofia della fisica. Bruno Mondadori, Milano (1997).

[BrKh95] Braginsky V.D., Khalili F.YA.: Quantum measurements. Cambridge University Press, Cambridge (1995).

[BrRo02] Bratteli O., Robinson D.W.: Operator Algebras and Quantum Statistical Mechanics, Vol. I and II (Second Edition). Springer-Verlag, Heidelberg Berlin (2002).

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7,

[Bus03] Busch P.: Quantum states and generalized observables: a simple proof of Gleason's theorem. Physical Review Letters **91**, 120403 (2003).

[BF02] Brunetti R., Fredenhagen K.: Remarks on time energy uncertainty relations. Reviews in Mathematical Physics **14**, 897–906 (2002).

[Bre10] Brezis H.: Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer-Verlag, Heidelberg Berlin (2010).

[BGL95] Busch P., Grabowski M., Lahti P.J.: Operational Quantum Physics. Springer-Verlag, Heidelberg Berlin (1995).

[Coh80] Cohn D.L.: Measure Theory. Birkhäuser, Boston (1980).

[CCP82] Caldirola P., Cirelli R., Prosperi G.M.: Introduzione alla Fisica Teorica. UTET, Torino (1982).

[DA10] Dell'Antonio G.: Aspetti matematici della meccanica quantistica. Vol. 1: Struttura matematica e concettuale. Series on physics and astrophysics 12. Bibliopolis, Napoli (2011).

[Dop09] Doplicher S.: The Principle of Locality. Effectiveness, fate and challenges. Journal of Mathematical Physics **51**, 015218 (2010).

[DCGi02] Dalla Chiara M.L., Giuntini R.: Quantum Logics. In: Gabbay D.M., Guenthner F. (eds.): Handbook of Philosophical Logic, Vol. 6, 129–228. Kluwer Academic Publishers, Netherlands (2002).

[Des99] D'Espagnat B.: Conceptual Foundations Of Quantum Mechanics (Second Edition). Westview Press (1999).

[Dir30] Dirac P.A.M.: The principles of Quantum Mechanics. Oxford University Press, Oxford (1930).

[Dix56] Dixmier J.: Sur la Relation $i(PQ-QP)=I$. Compos. Math. **13**, 263–269 (1956).

[DiUh77] Diestel J., Uhl J.J.Jr.: Vector Measures. American Mathematical Society, Providence (1977).

[doC92] do Carmo M.P.: Riemannian Geometry. Birkhäuser, Boston (1992).

[DI08] Doring A., Isham C.J.: A Topos foundation for theories of physics, I, II, III and IV. Journal of Mathematical Physics **49**, 053515 (2008); **49**, 053516 (2008); **49**, 053517 (2008); **49**, 053518 (2008).

[DS88] Dunford N., Schwartz J.T.: Linear Operators, Part II Spectral Theory, Self Adjoint Operators in Hilbert Space. Interscience Publishing, New York (1988).

[Emc72] Emch G.G.: Algebraic Methods in Statistical Mechanics and Quantum Field Theory. Wiley-Interscience, New York (1972).

[EGL09] Engesser K., Gabbay D.M., Lehmann D. (eds.): Handbook of Quantum Logic and Quantum Structures. Elsevier, Amsterdam (2009).

[EPR35] Einstein A., Podolsky B., Rosen N.: Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?. Physical Review **47**, 777–780 (1935).

[FaMa06] Fasano A., Marmi S.: Analytical Mechanics: An Introduction. Oxford University Press (2006).

[FK09] Friesecke G., Koppen M.: On the Ehrenfest theorem of quantum mechanics. Journal of Mathematical Physics **50**, 082102-082102-6 (2009).

[Fug50] Fuglede B.: A commutativity theorem for normal operators. Proceedings of the National Academy of the Sciences of the United States of America **36**, 35–40 (1950).

[GPS01] Goldstein H., Poole C.P., Safko J.L.: Classical Mechanics (Third Edition). Addison Wesley, Reading MA (2001).

[Giu96] Giulini D.: On Galilei invariance in quantum mechanics and the Bargmann superselection rule. Annals of Physics **249**, 222–235 (1996).

[Giu00] Giulini D.: Decoherence, a Dynamical Approach to Superselection Rules?. In [BGJKS00].

[Ghi97] Ghirardi G.C.: Un'occhiata alle carte di Dio. Il Saggiatore, Milano (1997).

[Gia97] Giannitrapani R.: Positive-operator-valued time observable in quantum mechanics. Internat. J. Theoret. Phys. **36**, 1575–1584 (1997).

[CoFr98I] Courant R., Fritz J.: Introduction to Calculus and Analysis, Vol. I. Springer-Verlag, Heidelberg Berlin (1998).

[CoFr98II] Courant R., Fritz J.: Introduction to Calculus and Analysis, Vol. II/1 and Vol. II/2. Springer-Verlag, Heidelberg Berlin (1998).

[Gle57] Gleason A.M.: Measures on the closed subspaces of a Hilbert space. Journal of Mathematics and Mechanics **6**, 885–893 (1957).

[Gra04] Graffi S.: Alcuni Aspetti Matematici della Meccanica Quantistica. Quaderni dell'INdAM, Gruppo Nazionale di Fisica Matematica, n. 61 (2004).

[Haa96] Haag R.: Local Quantum Physics (Second Revised and Enlarged Edition). Springer-Verlag, Heidelberg Berlin (1996).

[HaKa64] Haag R., Kastler D.: An Algebraic Approach to Quantum Field Theory. Journal of Mathematical Physics **5**, 848 (1964).

[Hal51] Halmos P.R.: Introduction to Hilbert Space and the theory of spectral multiplicity. Chelsea Publishing Company, New York (1951).

[Hal69] Halmos P.R.: Measure Theory. Van Nostrand R. (1969).

[Hua87] Huang K.: Statistical Mechanics (Second Edition). Wiley, New York (1987).

[HLS09] Heunen C., Landsman N.P., Spitters B.: A topos for algebraic quantum theory. Communications in Mathematical Physics **291**(1), 63–110 (2009).

[Jau73] Jauch J.M.: Foundations of Quantum Mechanics. Addison-Wesley Publishing Company, Reading MA (1978).

[JaPi69] Jauch J.M., Piron C.: On the structure of quantal proposition system. Helv. Phys. Acta **42**, 842 (1969).

[Jef04] Jefferies B.: Spectral Properties of Noncommuting Operators. Lecture Notes in Mathematics 1843, Springer, Heidelberg Berlin (2004).

[Kad51] Kadison R.: Isometries of Operator Algebras. Ann. Math. **54**, 325–338 (1951).

[KaRi97] Kadison R., Ringrose J.R.: Fundamentals of the Theory of Operator Algebras, I, II, III and IV. Graduate Studies in Mathematics, AMS (1997).

[KaAk80] Kantorovic L.V., Akilov G.P.: Analisi Funzionale. Editori Riuniti, Roma (1980).

[Kat66] Kato T.: Perturbation for Linear Operators. Springer-Verlag, Heidelberg Berlin (1966).

[KoFo99] Kolmogorov A.N., Fomin S.V.: Elements of the Theory of Functions and Functional Analysis. Dover Publications (1999).

[Kir74] Kirillov A.A.: Éléments de la théorie des représentations. MIR, Moscow (1974).

[KiGv82] Kirillov A.A, Gvisiani A.D.: Theorems and Problems in Functional Analysis. Springer-Verlag, Heidelberg Berlin New York (1982).

[Kup00] Kupsh J.: Mathematical Aspects of Decoherence. In [BGJKS00].

[Lan10] Lang S.: Linear Algebra (Third Edition). Springer-Verlag, Heidelberg Berlin (2010).

[Mac63] Mackey G.: The Mathematical Foundations of Quantum Mechanics. Benjamin, New York (1963).

[Mar82] Martucci G.: Spazi di Hilbert con elementi di meccanica quantistica. Pitagora Editrice, Bologna (1982).

[Mes99] Messiah A.: Quantum Mechanics. Dover Publications, New York (1999).

[Mor99] Moretti V.: Local ζ-function techniques vs point-splitting procedures: a few rigorous results. Communications in Mathematical Physics **201**, 327–364 (1999).

[Mor03] Moretti V.: Comments on the stress-energy tensor operator in curved spacetime. Communications in Mathematical Physics **232**, 232–265 (2003).

[Mrr01] Morrison T.J.: Functional analysis: An introduction to Banach space theory. Wiley-Interscience, New York (2001).

[NaSt82] Najmark M.A., Stern, A.I.: Theory of Group Representations. Springer-Verlag, Heidelberg Berlin New York (1982).

[Nel59] Nelson E.: Analytic Vectors. Ann. Math. **70**, 572–614 (1959).

[Neu32] von Neumann J.: Mathematische Grundlagen der Quantenmechanik. Springer-Verlag, Berlin (1932).

[Neu50] von Neumann J.: Functional Operators, Vol II. Princeton University Press (1950).

[NiCh07] Nielsen M.A., Chuang I.L.: Quantum Computation and Quantum Information (Ninth Printing). Cambridge University Press, Cambridge (2007).

[NiOu82] Nikiforov A., Ouvarov V.: Fonctions Spéciales de la physique mathématicque. MIR, Moscow (1982).

[Pir64] Piron C.: Axiomatique Quantique. Helv. Phys. Acta **37**, 439–468 (1964).

[Pir76] Piron C.: Foundations of Quantum Physics. W.A. Benjamin, Inc., London (1976).

[Prim00] Primas H.: Asymptotically Disjoint Quantum States. In [BGJKS00].

[Pru81] Prugovecki E.: Quantum Mechanics in Hilbert Space (Second Edition). Academic Press, New York (1981).

[RiNa53] Riesz F., Nagy B.Sz.: Functional Analysis. Dover, NewYork (1953).

[ReSi80] Reed M., Simon B.: Methods of Modern Mathematical Physics, Vol. I, II and IV. Academic Press, New York (1980).

[Rob04] Roberts J.E.: More Lectures on Algebraic Quantum Field Theory in Noncommutative Geometry. Lecture Notes in Mathematics 1831 (ed. S. Doplicher, R. Longo). Springer-Verlag, Heidelberg Berlin (2004).

[RR69] Roberts J.E., Roepstorf G.: Some basic concepts of algebraic quantum theory. Communications in Mathematical Physics **11**, 321–338 (1969).

[Rud64] Rudin W.: Principles of Mathematical Analysis. McGraw-Hill, New York (1964).

[Rud82] Rudin W.: Real and Complex Analysis. McGraw-Hill, New York (1986).

[Rud91] Rudin W.: Functional Analysis (Second Edition). McGraw-Hill, Boston (1991).

[Sal08] Salsa S.: Partial Differential Equations in Action. Springer-Verlag, Heidelberg Berlin (2008).

[Sch60] Schatten R.: Norm ideals of completely continuous operators. Springer-Verlag, Heidelberg Berlin New York (1960).

[Ser94I] Sernesi E.: Geometria 1. Bollati Boringheri, Torino (1994).

[Ser94II] Sernesi E.: Geometria 2. Bollati Boringheri, Torino (1994).

[Sim76] Simon B.: Quantum dynamics: From automorphism to Hamiltonian. Studies in Mathematical Physics, Essays in Honor of Valentine Bargmann (ed. E.H. Lieb, B. Simon, A.S. Wightman), Princeton University Press, Princeton, 327–349 (1976).

[Ste08] Stern A.: Anyons and the quantum Hall effect – A pedagogical review. Annals of Physics **323**, 204–249 (2008).

[Stre07] Streater R.F.: Lost Causes in and beyond Physics, Springer-Verlag, Heidelberg Berlin (2007).

[StWi00] Streater R.F., Wightman A.S.: PCT, Spin and Statistics and All That (Revised Edition). Princeton University Press, Princeton (2000).

[Str05a] Strocchi F.: An Introduction To The Mathematical Structure Of Quantum Mechanics: A Short Course For Mathematicians. World Scientific, Singapore (2005).
[Str05b] Strocchi F.: Symmetry breaking. Springer-Verlag, Heidelberg Berlin (2005).
[Str11] Strocchi F.: The Physical Principles of Quantum Mechanics. European Physics Journal Plus **127**, 12 (2012).
[Var07] Varadarajan V.S.: Geometry of Quantum Theory (Second Edition). Springer-Verlag, Heidelberg Berlin (2007).
[Vla81] Valdimirov V.S.: Le distribuzioni nella fisica matematica. MIR, Moscow (1981).
[Wald94] Wald R.M.: Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics. Chicago University Press, Chicago (1994).
[War75] Warner F.W.: Foundation of Differentiable Manifolds and Lie Groups. Springer-Verlag, New York (1975).
[Wei99] Weinberg S.: The Quantum Theory of Fields, Vol. 1, 2 and 3. Cambridge University Press (2005).
[Wes78] von Westenholtz C.: Differential Forms in Mathematical Physics. North-Holland, Amsterdam (1978).
[Wig59] Wigner E.: Group Theory and its Applications to the Quantum Theory of Atomic Spectra. Academic Press, Boston (1959).
[Wigh95] Wightman A.S.: Superselection rules; old and new. Nuovo Cimento **B 110**, 751–769 (1995).
[ZFC05] Zachos C., Fairlie D., Curtright T.: Quantum Mechanics in Phase Space. World Scientific, Singapore (2005).
[Zeh00] Zeh H.D.: The Meaning of Decoherence. In [BGJKS00].

This page intentionally left blank

Index

Moretti V.: *Spectral Theory and Quantum Mechanics*
Unitext – La Matematica per il 3+2
DOI 10.1007/978-88-470-2835-7,

Collana Unitext – La Matematica per il 3+2

As of 2004, the books published in the series have been given a volume number. Titles in grey indicate editions out of print.
As of 2011, the series also publishes books in English.

A. Bernasconi, B. Codenotti
Introduzione alla complessità computazionale
1998, X+260 pp, ISBN 88-470-0020-3

A. Bernasconi, B. Codenotti, G. Resta
Metodi matematici in complessità computazionale
1999, X+364 pp, ISBN 88-470-0060-2

E. Salinelli, F. Tomarelli
Modelli dinamici discreti
2002, XII+354 pp, ISBN 88-470-0187-0

S. Bosch
Algebra
2003, VIII+380 pp, ISBN 88-470-0221-4

S. Graffi, M. Degli Esposti
Fisica matematica discreta
2003, X+248 pp, ISBN 88-470-0212-5

S. Margarita, E. Salinelli
MultiMath – Matematica Multimediale per l'Università
2004, XX+270 pp, ISBN 88-470-0228-1

A. Quarteroni, R. Sacco, F.Saleri
Matematica numerica (2a Ed.)
2000, XIV+448 pp, ISBN 88-470-0077-7
2002, 2004 ristampa riveduta e corretta
(1a edizione 1998, ISBN 88-470-0010-6)

13. A. Quarteroni, F. Saleri
Introduzione al Calcolo Scientifico (2a Ed.)
2004, X+262 pp, ISBN 88-470-0256-7
(1a edizione 2002, ISBN 88-470-0149-8)

14. S. Salsa
Equazioni a derivate parziali - Metodi, modelli e applicazioni
2004, XII+426 pp, ISBN 88-470-0259-1

15. G. Riccardi
Calcolo differenziale ed integrale
2004, XII+314 pp, ISBN 88-470-0285-0

16. M. Impedovo
Matematica generale con il calcolatore
2005, X+526 pp, ISBN 88-470-0258-3

17. L. Formaggia, F. Saleri, A. Veneziani
Applicazioni ed esercizi di modellistica numerica
per problemi differenziali
2005, VIII+396 pp, ISBN 88-470-0257-5

18. S. Salsa, G. Verzini
Equazioni a derivate parziali – Complementi ed esercizi
2005, VIII+406 pp, ISBN 88-470-0260-5
2007, ristampa con modifiche

19. C. Canuto, A. Tabacco
Analisi Matematica I (2a Ed.)
2005, XII+448 pp, ISBN 88-470-0337-7
(1a edizione, 2003, XII+376 pp, ISBN 88-470-0220-6)

20. F. Biagini, M. Campanino
Elementi di Probabilità e Statistica
2006, XII+236 pp, ISBN 88-470-0330-X

21. S. Leonesi, C. Toffalori
Numeri e Crittografia
2006, VIII+178 pp, ISBN 88-470-0331-8

22. A. Quarteroni, F. Saleri
Introduzione al Calcolo Scientifico (3a Ed.)
2006, X+306 pp, ISBN 88-470-0480-2

23. S. Leonesi, C. Toffalori
Un invito all'Algebra
2006, XVII+432 pp, ISBN 88-470-0313-X

24. W.M. Baldoni, C. Ciliberto, G.M. Piacentini Cattaneo
Aritmetica, Crittografia e Codici
2006, XVI+518 pp, ISBN 88-470-0455-1

25. A. Quarteroni
Modellistica numerica per problemi differenziali (3a Ed.)
2006, XIV+452 pp, ISBN 88-470-0493-4
(1a edizione 2000, ISBN 88-470-0108-0)
(2a edizione 2003, ISBN 88-470-0203-6)

26. M. Abate, F. Tovena
Curve e superfici
2006, XIV+394 pp, ISBN 88-470-0535-3

27. L. Giuzzi
Codici correttori
2006, XVI+402 pp, ISBN 88-470-0539-6

28. L. Robbiano
Algebra lineare
2007, XVI+210 pp, ISBN 88-470-0446-2

29. E. Rosazza Gianin, C. Sgarra
Esercizi di finanza matematica
2007, X+184 pp,ISBN 978-88-470-0610-2

30. A. Machì
Gruppi – Una introduzione a idee e metodi della Teoria dei Gruppi
2007, XII+350 pp, ISBN 978-88-470-0622-5
2010, ristampa con modifiche

31 Y. Biollay, A. Chaabouni, J. Stubbe
Matematica si parte!
A cura di A. Quarteroni
2007, XII+196 pp, ISBN 978-88-470-0675-1

32. M. Manetti
Topologia
2008, XII+298 pp, ISBN 978-88-470-0756-7

33. A. Pascucci
Calcolo stocastico per la finanza
2008, XVI+518 pp, ISBN 978-88-470-0600-3

34. A. Quarteroni, R. Sacco, F. Saleri
Matematica numerica (3a Ed.)
2008, XVI+510 pp, ISBN 978-88-470-0782-6

35. P. Cannarsa, T. D'Aprile
Introduzione alla teoria della misura e all'analisi funzionale
2008, XII+268 pp, ISBN 978-88-470-0701-7

36. A. Quarteroni, F. Saleri
Calcolo scientifico (4a Ed.)
2008, XIV+358 pp, ISBN 978-88-470-0837-3

37. C. Canuto, A. Tabacco
Analisi Matematica I (3a Ed.)
2008, XIV+452 pp, ISBN 978-88-470-0871-3

38. S. Gabelli
Teoria delle Equazioni e Teoria di Galois
2008, XVI+410 pp, ISBN 978-88-470-0618-8

39. A. Quarteroni
Modellistica numerica per problemi differenziali (4a Ed.)
2008, XVI+560 pp, ISBN 978-88-470-0841-0

40. C. Canuto, A. Tabacco
Analisi Matematica II
2008, XVI+536 pp, ISBN 978-88-470-0873-1
2010, ristampa con modifiche

41. E. Salinelli, F. Tomarelli
Modelli Dinamici Discreti (2a Ed.)
2009, XIV+382 pp, ISBN 978-88-470-1075-8

42. S. Salsa, F.M.G. Vegni, A. Zaretti, P. Zunino
Invito alle equazioni a derivate parziali
2009, XIV+440 pp, ISBN 978-88-470-1179-3

43. S. Dulli, S. Furini, E. Peron
Data mining
2009, XIV+178 pp, ISBN 978-88-470-1162-5

44. A. Pascucci, W.J. Runggaldier
Finanza Matematica
2009, X+264 pp, ISBN 978-88-470-1441-1

45. S. Salsa
Equazioni a derivate parziali – Metodi, modelli e applicazioni (2a Ed.)
2010, XVI+614 pp, ISBN 978-88-470-1645-3

46. C. D'Angelo, A. Quarteroni
Matematica Numerica – Esercizi, Laboratori e Progetti
2010, VIII+374 pp, ISBN 978-88-470-1639-2

47. V. Moretti
Teoria Spettrale e Meccanica Quantistica – Operatori in spazi di Hilbert
2010, XVI+704 pp, ISBN 978-88-470-1610-1

48. C. Parenti, A. Parmeggiani
Algebra lineare ed equazioni differenziali ordinarie
2010, VIII+208 pp, ISBN 978-88-470-1787-0

49. B. Korte, J. Vygen
Ottimizzazione Combinatoria. Teoria e Algoritmi
2010, XVI+662 pp, ISBN 978-88-470-1522-7

50. D. Mundici
Logica: Metodo Breve
2011, XII+126 pp, ISBN 978-88-470-1883-9

51. E. Fortuna, R. Frigerio, R. Pardini
Geometria proiettiva. Problemi risolti e richiami di teoria
2011, VIII+274 pp, ISBN 978-88-470-1746-7

52. C. Presilla
Elementi di Analisi Complessa. Funzioni di una variabile
2011, XII+324 pp, ISBN 978-88-470-1829-7

53. L. Grippo, M. Sciandrone
Metodi di ottimizzazione non vincolata
2011, XIV+614 pp, ISBN 978-88-470-1793-1

54. M. Abate, F. Tovena
Geometria Differenziale
2011, XIV+466 pp, ISBN 978-88-470-1919-5

55. M. Abate, F. Tovena
Curves and Surfaces
2011, XIV+390 pp, ISBN 978-88-470-1940-9

56. A. Ambrosetti
Appunti sulle equazioni differenziali ordinarie
2011, X+114 pp, ISBN 978-88-470-2393-2

57. L. Formaggia, F. Saleri, A. Veneziani
Solving Numerical PDEs: Problems, Applications, Exercises
2011, X+434 pp, ISBN 978-88-470-2411-3

58. A. Machì
Groups. An Introduction to Ideas and Methods of the Theory of Groups
2011, XIV+372 pp, ISBN 978-88-470-2420-5

59. A. Pascucci, W.J. Runggaldier
Financial Mathematics. Theory and Problems for Multi-period Models
2011, X+288 pp, ISBN 978-88-470-2537-0

60. D. Mundici
Logic: a Brief Course
2012, XII+124 pp, ISBN 978-88-470-2360-4

61. A. Machì
Algebra for Symbolic Computation
2012, VIII+174 pp, ISBN 978-88-470-2396-3

62. A. Quarteroni, F. Saleri, P. Gervasio
Calcolo Scientifico (5ª ed.)
2012, XVIII+450 pp, ISBN 978-88-470-2744-2

63. A. Quarteroni
Modellistica Numerica per Problemi Differenziali (5ª ed.)
2012, XII+630 pp, ISBN 978-88-470-2747-3

64. V. Moretti
Spectral Theory and Quantum Mechanics
With an Introduction to the Algebraic Formulation
2013, XVI+728 pp, ISBN 978-88-470-2834-0

The online version of the books published in this series is available at SpringerLink.
For further information, please visit the following link:
http://www.springer.com/series/5418