Theoretical and Mathematical Physics

Paul Busch
Pekka Lahti
Juha-Pekka Pellonpää
Kari Ylinen

Quantum Measurement

Springer

Quantum Measurement

Theoretical and Mathematical Physics

The series founded in 1975 and formerly (until 2005) entitled *Texts and Monographs in Physics* (TMP) publishes high-level monographs in theoretical and mathematical physics. The change of title to *Theoretical and Mathematical Physics* (TMP) signals that the series is a suitable publication platform for both the mathematical and the theoretical physicist. The wider scope of the series is reflected by the composition of the editorial board, comprising both physicists and mathematicians.
The books, written in a didactic style and containing a certain amount of elementary background material, bridge the gap between advanced textbooks and research monographs. They can thus serve as basis for advanced studies, not only for lectures and seminars at graduate level, but also for scientists entering a field of research.

Editorial Board

W. Beiglböck, Institute of Applied Mathematics, University of Heidelberg, Heidelberg, Germany

P. Chrusciel, Gravitational Physics, University of Vienna, Vienna, Austria

J.-P. Eckmann, Département de Physique Théorique, Université de Genéve, Geneve, Switzerland

H. Grosse, Institute of Theoretical Physics, University of Vienna, Vienna, Austria

A. Kupiainen, Department of Mathematics, University of Helsinki, Helsinki, Finland

H. Löwen, Institute of Theoretical Physics, Heinrich-Heine-University of Düsseldorf, Düsseldorf, Germany

M. Loss, School of Mathematics, Georgia Institute of Technology, Atlanta, USA

N.A. Nekrasov, IHÉS, Bures-sur-Yvette, France

M. Ohya, Tokyo University of Science, Noda, Japan

M. Salmhofer, Institute of Theoretical Physics, University of Heidelberg, Heidelberg, Germany

S. Smirnov, Mathematics Section, University of Geneva, Geneva, Switzerland

L. Takhtajan, Department of Mathematics, Stony Brook University, Stony Brook, USA

J. Yngvason, Institute of Theoretical Physics, University of Vienna, Vienna, Austria

More information about this series at http://www.springer.com/series/720

Paul Busch · Pekka Lahti
Juha-Pekka Pellonpää · Kari Ylinen

Quantum Measurement

Springer

Paul Busch
Department of Mathematics, York Centre for Quantum Technologies
University of York
York
UK

Pekka Lahti
Department of Physics and Astronomy, Turku Centre for Quantum Physics
University of Turku
Turku
Finland

Juha-Pekka Pellonpää
Department of Physics and Astronomy, Turku Centre for Quantum Physics
University of Turku
Turku
Finland

Kari Ylinen
Department of Mathematics and Statistics
University of Turku
Turku
Finland

ISSN 1864-5879 ISSN 1864-5887 (electronic)
Theoretical and Mathematical Physics
ISBN 978-3-319-43387-5 ISBN 978-3-319-43389-9 (eBook)
DOI 10.1007/978-3-319-43389-9

Library of Congress Control Number: 2016946315

© Springer International Publishing Switzerland 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

Quantum Measurement is a book on the mathematical and conceptual foundations of quantum mechanics, with a focus on its measurement theory. It has been written primarily for students of physics and mathematics with a taste for mathematical rigour and conceptual clarity in their quest to understand quantum mechanics. We hope it will also serve as a useful reference text for researchers working in a broad range of subfields of quantum physics and its foundations.

The exposition is divided into four parts entitled *Mathematics* (Chaps. 2–8), *Elements* (Chaps. 9–13), *Realisations* (Chaps. 14–19), and *Foundations* (Chaps. 20–23). An overview of each part is given in the Introduction, Chap. 1, and each chapter begins with a brief non-technical outline of its contents.

A glance through the table of contents shows that different chapters require somewhat different backgrounds and levels of prerequisite knowledge on the part of the reader. The material is arranged in a logical (linear) order, so it should be possible to read the book from beginning to end and gain the relevant skills along the way, either from the text itself or occasionally from other sources cited. However, the reader should also be able to start with any part or chapter of her or his interest and turn to earlier parts where needed.

Part I is designed to be accessible to a reader possessing an undergraduate level of familiarity with linear algebra and elementary metric space theory. Chaps. 2 and 3 can be read as an introduction to the part of Hilbert space theory which does not need measure and integration theory. The latter becomes an essential tool from Chap. 4 onwards, so we give a summary of the key concepts and some relevant results. Starting with Sect. 4.10, and more essentially from Chap. 6 on, we occasionally need the basic notions of general topology and topological vector spaces. Elements of the theory of C^*-algebras and von Neumann algebras are briefly summarised in Chap. 6, but their role is very limited in the sequel.

While prior study of quantum mechanics might be found useful, it is not a prerequisite for a successful study of the book. The essence of the work is the development of tools for a rigorous approach to central questions of quantum mechanics, which are often considered in a more intuitive and heuristic style in the literature. In this way the authors hope to contribute to the clarification of some key issues in the discussions concerning the foundations and interpretation of quantum mechanics.

The bibliography is fairly extensive, but it does not claim to be comprehensive in any sense. It contains many works on general background and key papers in the development of the field of quantum measurement. Naturally, especially most of the more recent references relate to the topics central to this book, in which the authors and their collaborators have also had their share.

The reader will notice that the word *measure* is used in a variety of meanings, which should, however, be clear from the context. A measure as a mathematical concept is a set function which can be specified by giving the value space: we talk about (positive) measures, probability measures, complex measures, operator measures, etc. We also speak about the measures of quantifiable features such as accuracy, disturbance, or unsharpness. The etymologically related word *measurement* may be taken to refer to a process, but it is also given a precise mathematical content that can be viewed as an abstraction of this process.

Much of the material in this book has been extracted and developed from various series of lecture notes for graduate and postgraduate courses in mathematics and theoretical physics held over many years at the universities of Helsinki, Turku and York. In its totality, however, the work is considerably more comprehensive than the union of these courses. It reflects the development of its subject from the early days of quantum mechanics while the selection of topics is inevitably influenced by the authors' research interests. In fact, the book emerged in its present shape from a decade-long collective effort alongside our investigations into quantum measurement theory and its applications. At this point we wish to express our deep gratitude and appreciation to the many colleagues, scientific friends and, not least, our students with whom we have been fortunate to collaborate and discuss fundamental problems of quantum physics.

York, UK
Turku, Finland

Paul Busch
Pekka Lahti
Juha-Pekka Pellonpää
Kari Ylinen

Contents

Chapter 1
Introduction

1.1 Background and Content

The Book of Nature, already according to Galileo Galilei, is written in the language of mathematics. This dictum sounds like a commonplace to scientists today. True, we may qualify its content: we might not know or ever find out the actual writing process, but mathematics appears to be the best, or even only, vehicle into the otherwise impenetrable realm of the microworld. Indeed, the formulation of the theory of quantum mechanics as it emerged in the early 20th century after two or three decades of intense search and debate, frustrations and triumphs, was considered successfully completed only when appropriate mathematical tools had been identified.

Two extraordinarily influential works, Paul Dirac's *The Principles of Quantum Mechanics* (1930) and John von Neumann's *Mathematische Grundlagen der Quantenmechanik* (1932) generalised and crystallised the ideas of the founders into workable methodologies. According to a commonly held view, perhaps Dirac's technique and language were and still are more apt to appeal to (theoretical) physicists and von Neumann's to mathematicians and mathematical physicists. At the outset, von Neumann's work built on the fast growing body of functional analysis, especially the spectral theory of Hilbert space operators. On the surface, Dirac's language is more heuristic, and while there are later theories which can be used to make it mathematically sound, the von Neumann style functional analytic approach still dominates the mathematically oriented research.

The book of von Neumann (with its English translation of 1955) has had an enormous follow-up with a fruitful interplay of physical and mathematical ideas. The present work owes its existence to and emphatically joins this tradition.

The mathematical groundwork for von Neumann's book [1] was laid down in a couple of papers from the year 1927 [2, 3]. There he undertakes an analysis of general statistical aspects of a physical experiment using the concepts of states and observables, with the requirement that these entities determine the respective probabilities for the registration of measurement outcomes. This fundamental investigation led to the following result, summarised here in present-day terminology:

© Springer International Publishing Switzerland 2016

P. Busch et al., *Quantum Measurement*, Theoretical and Mathematical Physics,
DOI 10.1007/978-3-319-43389-9_1

> It is assumed that the description of a physical system is based on a complex separable Hilbert space $\mathcal{H}$ with inner product $\langle \cdot \mid \cdot \rangle$ and that the pure states of the system are represented by the unit vectors φ (modulo a phase factor) of $\mathcal{H}$. It is required further that the measurement outcome probabilities for a given observable are to be given in terms of a *single* (linear) operator acting in $\mathcal{H}$. It follows that this operator must be a selfadjoint operator A, and the probability that a measurement of the observable represented by A in a state described by φ leads to a result in a (real Borel) set X is given by the number $\langle \varphi \mid \mathsf{E}^A(X)\varphi \rangle$, where $\mathsf{E}^A(X)$ is the spectral projection of A associated with the set X.

In addition, von Neumann showed that the most comprehensive representation of states is given in terms of the positive operators ρ of trace one acting on $\mathcal{H}$ called states or density operators; the pure states are the idempotent elements among these operators, $\rho^2 = \rho$, hence the projections onto one-dimensional subspaces of $\mathcal{H}$. Thus he already deduced the trace formula $\mathrm{tr}\big[\rho \mathsf{E}^A(X)\big]$ for the measurement outcome probabilities.

It took until the late 1960's before it was fully recognised that representing measurement outcome probabilities of an observable in a state in terms of a single operator is unnecessarily restrictive. Indeed, from the mathematical point of view the probabilistic analysis leads to the representation of observables as semispectral measures, normalised positive operator measures, thus going beyond the more special spectral measures. This mathematical extension not only broadened the domain of applicability of quantum measurements but also opened new avenues for a quantitative analysis of questions like approximate joint measurability of observables traditionally represented by mutually noncommuting operators (spectral measures) or the unavoidable disturbance caused by a measurement. The monographs [4–7] give an account of this line of development.

These ingredients—states as positive trace one operators, observables as normalised positive operator measures (with all operators acting on a fixed complex separable Hilbert space), and the probability measures they define—form the starting point of the formulation of Hilbert space quantum mechanics discussed in this book. We will mostly adhere to the so-called *minimal interpretation* of quantum mechanics, according to which quantum mechanics is a theory of measurement outcome probabilities defined by states (equivalence classes of preparations) and observables (equivalence classes of measurements). This has the advantage of offering a conceptually clear and mathematically rigorous framework with no immediate need to consider the more philosophical issues in the foundations of quantum mechanics. It is possible, in some sense, to read most of the book as a piece of mathematics, although the choice of topics is dictated by physical applications. This attitude seems to be in line with the actual practice of physicists, who in their collaboration may use the same mathematical language and minimal interpretation whilst maintaining widely diverging philosophical views.

Our book is divided into four parts and 23 chapters: I. Mathematics (2–8); II. Elements (9–13); III. Realisations (14–19); IV. Foundations (20–23). We now give a brief overview of the contents.

Part I. Mathematics. The purpose of this part is to set the stage for a mathematical, or more specifically, Hilbert space based analysis of the physical phenomena generally described as quantum measurements. The choice of the material has been made with readers of diverse backgrounds in mind.

In Chap. 2 we develop the basics of Hilbert space theory. Chapter 3 looks at compact operators from different angles, the main result being the spectral representation of a compact selfadjoint operator. With this we are ready to study operator ideals like the trace class and to consider problems such as the Schmidt decomposition of a vector in a tensor product Hilbert space.

Chapters 4 and 5 contain the spectral representation theory of (generally unbounded) selfadjoint operators and its application to the representation of one-parameter unitary groups. Here we proceed via the case of bounded normal operators and use of the Cayley transform. In anticipation of the physical applications throughout the remaining parts of the book, the treatment adopts the general perspective of positive operator measures (and bimeasures) while it is understood that the spectral theorem only requires the spectral measures.

Chapters 6–8, somewhat less self-contained, introduce various functional analytic techniques including some elements of the theory of C^*-algebras and von Neumann algebras (Chap. 6) and the dilation theories of Naimark and Stinespring (Chap. 7). In the interest of economy, the dilation theorems of Naimark and Stinespring are deduced from a two variable dilation theorem, which has also independent importance in measurement theory. Chapter 8 contains specific physical examples of positive operator measures, which give a glimpse of the kind of material to be expected in the sequel. The technique of direct integral Hilbert spaces and a related elementary approach to a Dirac type treatment are briefly discussed here as well.

Part II. Elements. This part develops the basic notions and structures of Hilbert space quantum mechanics as applied in this monograph. Chapter 9 starts with setting out the associated statistical duality: fixing the set of states of a quantum system to consist of the positive trace one operators on a complex separable Hilbert space, we can then deduce the structure of observables and the measurement outcome probabilities. This chapter also introduces the tools required to describe the changes a physical system may undergo in the course of its time evolution or due to an intervention, such as a measurement. Further we recall the composition rules that lead to the Hilbert tensor product structure as the framework for the theory of compound systems. The chapter concludes with a brief discussion of the important concepts of subsystem states, dynamics, correlations, and entanglement.

The theory of measurement is formulated in Chap. 10 by considering measurements as physical processes subject to the laws of quantum mechanics. We identify a hierarchy of three levels of description: observables–instruments–measurements. Observables are equivalence classes of completely positive instruments, and the latter are equivalence classes of measurement schemes. This hierarchy reflects the options of restricting one's attention to the outcome probabilities at the level of the measured system, or taking into account the system's conditional state changes, or adopting the most comprehensive level of modelling the interaction and information transfer between system and probe.

In Chap. 11 we turn our attention to one of the most striking nonclassical features of quantum mechanics: the existence of sets of observables that are incompatible in the sense that they cannot be measured jointly. We consider several equivalent formulations of the notion of joint measurability of observables. A natural framework is thus obtained for investigating incompatibility and the disturbance of the object system caused by a measurement. Operationally, this disturbance manifests itself in a change of the measurement outcome statistics of some other observables.

The final Chaps. 12 and 13, of Part II make precise such concepts as indeterminacy, uncertainty, approximate measurement, and disturbance caused by a measurement. We also introduce various measures of uncertainty, inaccuracy and disturbance, and show how to quantify the degree of unsharpness of an observable. We use these measures to formulate examples of preparation and measurement uncertainty relations.

Part III. Realisations. In this part the major examples of observables and some of their measurement models are investigated. The list of examples includes qubit observables (Chap. 14), position and momentum (Chap. 15), number and phase (Chap. 16), and time and energy (Chap. 17). The question of approximate joint measurements is taken up once more and examples of model-independent error trade-off relations are given for incompatible pairs of qubit observables and for the position and momentum observables of a particle.

Chapter 18 is devoted to a study of informational completeness and the related problem of state reconstruction. Special attention is given there to the continuous variable case. The key concepts and the basic results will be introduced, including a short discussion of the qubit case. The so-called Pauli problem—the informational incompleteness of the canonical position–momentum pair—and the two basic ways of overcoming this problem are studied.

Part III concludes with Chap. 19 where the tools of measurement theory are put to full use to illustrate the implementation of more or less realistic measurement schemes for typical observables. The focus will be on the realisation of joint approximate measurements of noncommuting pairs of observables, with the Arthurs–Kelly model, homodyne detection schemes and Mach–Zehnder interferometry serving as prototypical examples.

Part IV. Foundations. The final part of the book is devoted to a selection of foundational issues of quantum mechanics insofar as they have some measurement-theoretic significance: Bell inequality violations and their dependence on the use of incompatible measurements (Chap. 20); limitations of measurements due to conservation laws (Chap. 21); the so-called measurement problem (Chap. 22); and finally, an axiomatic justification of the Hilbert space formulation of quantum mechanics based on ontological premises constraining measurement possibilities (Chap. 23).

1.2 Statistical Duality—an Outline

We now give a brief outline of the key mathematical structure that motivates and underlies the developments presented in this book: the duality of states and observables, concepts that are fundamental to the formalisation of any probabilistic physical theory. We also indicate some of the prominent probabilistic features that distinguish quantum mechanics from its classical counterpart. While the Hilbert space realisation of the statistical duality is developed mathematically in Part I and used freely in virtually all applications discussed thereafter, we revisit the abstract duality in the final chapter where it serves as the starting point for an axiomatic basis of quantum mechanics that we will review there.

The Duality

In von Neumann's formulation of quantum mechanics one meets states and observables as positive trace-one operators and general selfadjoint operators (or the associated spectral measures), respectively. The states and the projections that figure in the description of standard observables are elements of the real vector spaces of selfadjoint trace class operators and of selfadjoint bounded operators, respectively, where the latter is the dual space of the former. The extension of the notion of observable towards including general normalised positive operator measures is found to be both natural and comprehensive when considered from the perspective of a general statistical duality.

The dual pair of states and observables can be easily manifested as core elements of a probabilistic description by way of a simple analysis of the general statistical aspects of a physical experiment. In a typical experiment one can distinguish three steps: the *preparation* of a physical system, followed by a *measurement* which is performed on it, and finally the *registration* of a result. In order that an experiment serves its purpose of providing information about the system under investigation, it should meet a requirement of *statistical causality*: the registered outcome should depend, generally in a probabilistic way, on how the system was prepared and what kind of measurement was performed.

The physical system $\mathcal{S}$ under consideration can be prepared in various ways and then subjected to one or more of a range of different measurements. We take the terms *system*, *preparation* and *measurement* to be intuitively understood without trying to explicate them at this stage.

Let π denote a preparation and Π the collection of all possible preparations of the system $\mathcal{S}$. Further let σ stand for a measurement and Σ denote the collection of all conceivable measurements that can be performed on $\mathcal{S}$. By fixing a measurement σ in Σ one also specifies the range of its possible outcomes. We identify these outcomes as members of a set Ω which can typically be thought of as a set of real numbers, and for the purpose of counting statistics a sigma-algebra $\mathcal{A}$ of subsets of Ω will be specified consisting of the test sets, that is, bins within which groups of outcomes are counted. Thus, a measurement is represented as a triple $(\sigma, \Omega, \mathcal{A})$, which we will often simply denote by σ.

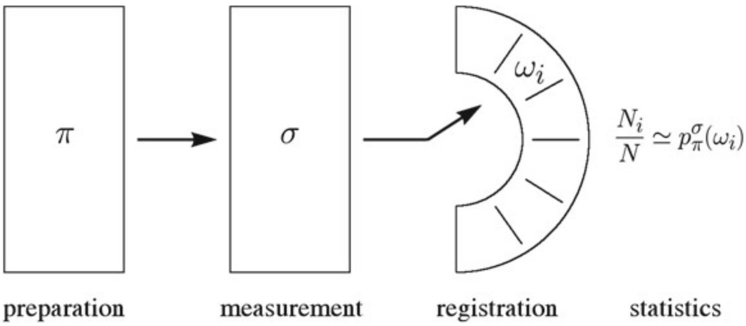

Fig. 1.1 Scheme of a physical experiment

The notion of statistical causality specifies that any preparation π and measurement σ determine a probability distribution for the possible measurement outcomes. Thus there is a probability measure $p_\pi^\sigma : \mathcal{A} \to [0, 1]$, with the heuristic understanding that if one makes a large number, N, of repetitions of the same measurement σ under the same conditions π, and a result $\omega \in \Omega$ is registered $n(X)$ times in a test set X, then[1]

$$\frac{n(X)}{N} \simeq p_\pi^\sigma(X).$$

Two preparations π_1 and π_2 are said to be equivalent, $\pi_1 \equiv \pi_2$, if they give the same measurement outcome probabilities for all measurements, that is, $p_{\pi_1}^\sigma = p_{\pi_2}^\sigma$ for all measurements σ. We may hence consider the collection Π to be divided into equivalence classes $[\pi] = \{\pi' \in \Pi \mid \pi' \equiv \pi\}$. These classes are called *states* of the system. We let $\boldsymbol{S}$ denote the set of states of $\mathcal{S}$. Thus the formal concept of state represents those aspects of a physical process applied as a preparation of a system that determine the outcome probabilities of any subsequent measurement (Fig. 1.1).

Similarly, two measurements σ_1 and σ_2 are equivalent if $p_\pi^{\sigma_1} = p_\pi^{\sigma_2}$ for all preparations π. We may accordingly talk about equivalent classes of measurements as *observables*; we let $\boldsymbol{O}$ denote the collection of all observables. The notion of observable, as delineated here, embodies the idea that a physical quantity is uniquely determined through its probabilistic signature. We shall refer to the pair $(\boldsymbol{S}, \boldsymbol{O})$ as a *statistical duality*.

For any state $s \in \boldsymbol{S}$ and observable $O \in \boldsymbol{O}$ one defines

$$p_s^O = p_\pi^\sigma, \quad \pi \in s, \sigma \in O.$$

[1] We leave aside the problem of justifying the frequency interpretation of probabilities. A lucid account of this problem and a consistent interpretation of probabilities as relative frequencies is given by van Fraassen [8].

This is a well-defined probability measure with the following *minimal interpretation*: the number $p_s^O(X)$ is the probability that a measurement of the observable O leads to a result in the set X when the system is in the state s.

The above discussion leading to this brief statement is there simply to give intuitive background and motivation for the terminology used in our subsequent mathematical work, not to gloss over the inherent problems of the use of mathematical language in physical theories. When we use mathematical structures in the sequel we do not deviate from the usual mathematical parlance.

Elementary Structures

There are two basic structural properties that the statistical duality $(\boldsymbol{S}, \boldsymbol{O})$ may always be assumed to possess. First, since a convex combination of two or more probability measures is a probability measure, the set $\boldsymbol{S}$ of states can be equipped with a convex structure. Indeed, if $s_1, s_2 \in \boldsymbol{S}$ and $0 \leq \lambda \leq 1$, then for any $O \in \boldsymbol{O}$, the convex combination

$$\lambda p_{s_1}^O + (1 - \lambda) p_{s_2}^O$$

is a probability measure. One may thus pose the requirement that there is a (necessarily unique) $s \in \boldsymbol{S}$ such that

$$p_s^O = \lambda p_{s_1}^O + (1 - \lambda) p_{s_2}^O$$

for all $O \in \boldsymbol{O}$. The assumption that $\boldsymbol{S}$ is closed under convex combinations corresponds to the idea that any two preparations $\pi_1 \in s_1,\ \pi_2 \in s_2$ can be combined into a new preparation, for instance by applying π_1 and π_2 in random order with frequencies λN, $(1 - \lambda)N$, respectively; upon measurement one obtains outcome distributions that are given by the convex combination $\lambda p_{s_1}^O + (1 - \lambda) p_{s_2}^O$.

An important feature of the convex structure[2] of the set of states $\boldsymbol{S}$ is the possibility of distinguishing *pure states* as those that cannot be expressed as a convex combinations of other states; all other states are referred to as *mixed states*. Thus, the second assumption concerning the set of states one may adopt is that it contains a sufficiently rich set of pure states, which embody maximal information one may have about the system, so that all other states can be obtained as convex combinations of them (or more generally, as limits of such combinations in a suitable sense). This is realised in the classical and quantum mechanical probabilistic theories.

A classical theory is distinguished by the fact that its set of states is a *simplex*, which means that every mixed state can be expressed as a (generalised) convex combination of pure states in one and only one way. In contrast, a mixed quantum state has infinitely many different decompositions into pure states. (Theorem 9.2 gives a full characterisation of such decompositions.)

[2]The convex structure of the set of states is initially defined abstractly, without first assuming that the set of states is a subset of a real vector space. The underlying linear structure can be deduced by making a simple, innocent additional assumption, namely, that the set of observables allows one to separate distinct states. We return to this point in greater detail in Sect. 23.1.

This formal difference between the quantum and classical statistical dualities is closely related to the fundamental phenomenon of quantum *indeterminacy*, usually referred to by the term *uncertainty principle*. Broadly, this is the statement that there is no state in which all observables would have definite values. A classical statistical duality is typically formulated in terms of a convex set of probability measures on a phase space such that all pure states, given by the point measures (also called Dirac measures), are included. Any phase space point, and hence every point measure, specifies the values of all observables, defined as functions on phase space. Since every mixed state is represented in a unique way as a (generalised) convex combination of pure states, it becomes possible to interpret the associated probabilities as representing a lack of information about the actual value of an observable. In contrast, there is no pure quantum state that could assign probability one to a value of every observable. This fundamental quantum indeterminacy or preparation uncertainty is often quantified by means of the *preparation uncertainty relations*.

Another distinctive feature of classical physical theories, already alluded to above, concerns the joint measurability of observables: in a general statistical duality, one can ask which sets of observables can be measured jointly. In the classical case, there is no restriction to joint measurability, whereas in quantum mechanics, there are severe limitations: according to a theorem due to von Neumann [9], any two observables represented by selfadjoint operators are jointly measurable if and only if they commute with each other. (Theorem 11.3 expresses this result.)

The notion of joint measurability can be readily captured in terms of the general preparation–measurement–registration scheme of a statistical duality $(\boldsymbol{S}, \boldsymbol{O})$. There are several obvious ways of defining the joint measurability of, say, a pair of observables $(O_i, \Omega_i, \mathcal{A}_i)$, $i = 1, 2$. We refer to Chap. 11 for a more comprehensive analysis of this concept and adopt here to the following formulation: assume that there is an observable $(O, \Omega, \mathcal{A})$ with measurable functions $f_i : \Omega \to \Omega_i$, $i = 1, 2$, such that for any state $s \in \boldsymbol{S}$,

$$p_s^{O_1}(X) = p_s^O\big(f_1^{-1}(X)\big), \quad X \in \mathcal{A}_1, \tag{1.1}$$

$$p_s^{O_2}(Y) = p_s^O\big(f_2^{-1}(Y)\big), \quad Y \in \mathcal{A}_2. \tag{1.2}$$

The observable $(O, \Omega, \mathcal{A})$, together with the functions f_i, comprises all probability measures associated with the observables O_1, O_2 and thus serves as their joint measurement.

As already noted, it is a fundamental feature of quantum mechanics that there are observables (represented by noncommuting selfadjoint operators) that cannot be measured jointly. It was a bold idea of Werner Heisenberg, expressed in his seminal paper [10] of 1927, that such observables can, however, be measured jointly in an approximate way if the approximation errors satisfy a *measurement uncertainty relation*. With the tools available at that time, Heisenberg was able to give only intuitive motivations and heuristic arguments for such ideas, essentially on the basis of semiclassical discussions of some thought experiments.

In view of the above notion of joint measurement, an approximate joint measurement of observables $(O_i, \Omega_i, \mathcal{A}_i)$, $i = 1, 2$, is a measurement and thus defines an observable $(\tilde{O}, \Omega, \mathcal{A})$, together with measurable functions $f_i : \Omega \to \Omega_i$, $i = 1, 2$, such that for any state $s \in \boldsymbol{S}$, the measurement outcome distributions of $\tilde{O}_1$ and $\tilde{O}_2$ from (1.1) and (1.2) approximate the corresponding distributions of O_1 and O_2, respectively. It remains to quantify the quality of approximation, that is, to define a 'distance' of $\tilde{O}_i$ from O_i (in terms of a distance between the distributions $p_s^{\tilde{O}_i}$ and $p_s^{O_i}$), and then to analyse the possible measurement uncertainty relations needed for an approximate joint measurement of the two observables. This is the topic of Chap. 13, elaborated further in some examples in Sects. 14.5 and 15.3.

There are several other features which distinguish quantum probabilistic theories from classical theories. These could easily be explained and formalised in terms of the statistical duality $(\boldsymbol{S}, \boldsymbol{O})$. We mention only the possibility of superposing pure states into new pure states, or the phenomenon of entanglement in the case where the system represented by the duality $(\boldsymbol{S}, \boldsymbol{O})$ can be considered to be composed of two (or more) subsystems with the dualities $(\boldsymbol{S}_i, \boldsymbol{O}_i)$. The idea of superposing pure states into new pure states appears naturally in the Hilbert space formulation of quantum mechanics, Sect. 9.1, whereas in Chap. 23 the general notion of superposition of pure states, as given in Definition 23.3, is seen to exclude a classical description. The composition rules of Sect. 9.5 will be seen to lead to the probabilistic dependence known as entanglement between the subsystems, again something that is foreign to classical physical theories. Chapter 20 on Bell inequalities gives further insight into this nonclassical aspect of quantum mechanics.

References

1. von Neumann, J.: Mathematische Grundlagen der Quantenmechanik. Die Grundlehren der mathematischen Wissenschaften, Band 38. Springer, Berlin (1968, 1996) (Reprint of the 1932 original). English translation: Mathematical Foundations of Quantum Mechanics. Princeton University Press, Princeton (1955, 1996)
2. von Neumann, J.: Mathematische Begründung der Quantenmechanik. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1927**, 1–57 (1927)
3. von Neumann, J.: Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1927**, 245–272 (1927)
4. Davies, E.B.: Quantum Theory of Open Systems. Academic Press, London (1976)
5. Helstrom, C.: Quantum Detection and Estimation Theory. Academic Press, New York (1976)
6. Holevo, A.S.: Probabilistic and Statistical Aspects of Quantum Theory. North-Holland Series in Statistics and Probability, vol. 1. North-Holland Publishing Co, Amsterdam (1982) (Translated from the Russian by the author)
7. Ludwig, G.: Foundations of Quantum Mechanics. I. Texts and Monographs in Physics. Springer, New York (1983) (Translated from the German by Carl A. Hein)

8. van Fraassen, B.C.: Foundations of probability: a modal frequency interpretation. Problems in the Foundations of Physics, pp. 344–387 (1979)
9. von Neumann, J.: Über Funktionen von Funktionaloperatoren. Ann. Math. **32**(2), 191–226 (1931)
10. Heisenberg, W.: Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. Z. Physik **43**, 172–198 (1927)

Part I
Mathematics

Chapter 2
Rudiments of Hilbert Space Theory

As the present work is about Hilbert space quantum mechanics, it is mandatory that the reader has sufficient grounding in Hilbert space theory. This short chapter is designed to indicate what sort of basic equipment one needs in the ensuing more sophisticated chapters. At the same time it can be used as an introduction to elementary Hilbert space theory even for the novice. The material is quite standard and appears of course in numerous works, so we do not explicitly specify any references, though some source material can be found in the bibliography.

2.1 Basic Notions and the Projection Theorem

We begin with a key definition. Unless otherwise stated, all vector spaces in this work have the field $\mathbb{C}$ of complex numbers as their field of scalars. We denote by $\mathbb{N}$ the set of positive integers, i.e. $\mathbb{N} = \{1, 2, 3, \ldots\}$, and let $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

Definition 2.1 Let E be a (complex) vector space. We say that a mapping $h : E \times E \to \mathbb{C}$ is an *inner product* (in E) and E (equipped with h) is an *inner product space*, if for all $\varphi, \psi, \eta \in E$ and $\alpha, \beta \in \mathbb{C}$ we have

(IP1) $h(\varphi, \alpha\psi + \beta\eta) = \alpha h(\varphi, \psi) + \beta h(\varphi, \eta)$,
(IP2) $h(\varphi, \psi) = \overline{h(\psi, \varphi)}$,
(IP3) $h(\varphi, \varphi) \geq 0$,
(IP4) $h(\varphi, \varphi) > 0$ if $\varphi \neq 0$.

Unless otherwise stated, in the sequel we write $h(\varphi, \psi) = \langle\, \varphi \,|\, \psi \,\rangle$ in any context described by this definition.

In (b) below there is the *Cauchy–Schwarz inequality*.

© Springer International Publishing Switzerland 2016
P. Busch et al., *Quantum Measurement*, Theoretical and Mathematical Physics,
DOI 10.1007/978-3-319-43389-9_2

Theorem 2.1 *Assume that* E *is an inner product space,* $\varphi, \psi, \eta \in E$, $\alpha, \beta \in \mathbb{C}$. *Then*

(a) $\langle \alpha\varphi + \beta\psi \,|\, \eta \rangle = \overline{\alpha} \langle \varphi \,|\, \eta \rangle + \overline{\beta} \langle \psi \,|\, \eta \rangle$;
(b) $|\langle \varphi \,|\, \psi \rangle|^2 \leq \langle \varphi \,|\, \varphi \rangle \langle \psi \,|\, \psi \rangle$.

Proof Part (a) is an immediate consequence of the definition. To prove (b), note that for any $\alpha \in \mathbb{C}$ we have

$$0 \leq \langle \alpha\varphi + \psi \,|\, \alpha\varphi + \psi \rangle = |\alpha|^2 \langle \varphi \,|\, \varphi \rangle + \overline{\alpha} \langle \varphi \,|\, \psi \rangle + \alpha \langle \psi \,|\, \varphi \rangle + \langle \psi \,|\, \psi \rangle .$$

If $\langle \varphi \,|\, \varphi \rangle \neq 0$, choose $\alpha = -\langle \varphi \,|\, \psi \rangle \langle \varphi \,|\, \varphi \rangle^{-1}$, and then

$$|\langle \varphi \,|\, \psi \rangle|^2 \langle \varphi \,|\, \varphi \rangle^{-1} - |\langle \varphi \,|\, \psi \rangle|^2 \langle \varphi \,|\, \varphi \rangle^{-1} - |\langle \varphi \,|\, \psi \rangle|^2 \langle \varphi \,|\, \varphi \rangle^{-1} + \langle \psi \,|\, \psi \rangle \geq 0,$$

which implies the claim. If $\langle \varphi \,|\, \varphi \rangle = 0$, by multiplying φ with a suitable complex number of modulus one we may assume that $\langle \varphi \,|\, \psi \rangle$ is real, and then it is easy to see that the above inequality can be true for all $\alpha \in \mathbb{R}$ only if $\langle \varphi \,|\, \psi \rangle = 0$. □

Remark 2.1 If (IP1) and (a) above hold for h, then the map h is called *sesquilinear*. If, moreover, (IP3) holds, it is a *positive sesquilinear form*. If h is sesquilinear, then $h(\varphi, \psi) = \overline{h(\psi, \varphi)}$ for all $\varphi, \psi \in \mathcal{H}$ if and only if $h(\varphi, \varphi) \in \mathbb{R}$ for all $\varphi \in \mathcal{H}$. Indeed, the “only if” part is obvious, and to prove the “if” part, write

$$h(\alpha\varphi + \beta\psi, \alpha\varphi + \beta\psi) = |\alpha|^2 h(\varphi, \varphi) + \overline{\alpha}\beta h(\varphi, \psi) + \overline{\beta}\alpha h(\psi, \varphi) + |\beta|^2 h(\psi, \psi)$$

and substitute first $\alpha = \beta = 1$ and then $\alpha = 1$, $\beta = i$ to see that $\operatorname{Im} h(\varphi, \psi) = -\operatorname{Im} h(\psi, \varphi)$ and $\operatorname{Re} h(\varphi, \psi) = \operatorname{Re} h(\psi, \varphi)$. In particular, positive sesquilinear forms automatically satisfy (IP2), and the proof we gave for the Cauchy–Schwarz inequality was so formulated that it is valid without assuming (IP4). This generality will be needed later. ◁

The proof of the next result is an easy exercise.

Theorem 2.2 *Let* E *be an inner product space. Denote* $\|\varphi\| = \sqrt{\langle \varphi \,|\, \varphi \rangle}$, *when* $\varphi \in E$. *Then*

(a) $\|\varphi\| \geq 0$ *for all* $\varphi \in E$;
(b) $\|\varphi\| = 0$, *if and only if* $\varphi = 0$;
(c) $\|\alpha\varphi\| = |\alpha| \|\varphi\|$ *for each* $\alpha \in \mathbb{C}$ *and* $\varphi \in E$;
(d) $\|\varphi + \psi\| \leq \|\varphi\| + \|\psi\|$ *for all* $\varphi, \psi \in E$.

Remark 2.2 The preceding result means that the map $\varphi \mapsto \sqrt{(\varphi|\varphi)}$ is a *norm* on E. Unless otherwise stated, an inner product space will be equipped with this norm. ◁

Example 2.1 The set $\mathbb{C}^n = \{x = (x_1, \ldots, x_n) \,|\, x_k \in \mathbb{C},\ k = 1, \ldots, n\}$ (where $n \in \mathbb{N}$) is an inner product space with respect to its usual operations $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$ and $\alpha(x_1, \ldots, x_n) = (\alpha x_1, \ldots, \alpha x_n)$ and the inner product $\langle x \mid y \rangle = \sum_{k=1}^{n} \overline{x}_k y_k, x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n)$. The Cauchy-Schwarz inequality acquires the form

$$\left| \sum_{k=1}^{n} \overline{x}_k y_k \right| \leq \left(\sum_{k=1}^{n} |x_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} |y_k|^2 \right)^{\frac{1}{2}}.$$

$\triangleleft$

Example 2.2 Denote $\ell^2 = \ell^2_{\mathbb{C}} = \left\{ f : \mathbb{N} \to \mathbb{C} \,\middle|\, \sum_{n=1}^{\infty} |f(n)|^2 < \infty \right\}$ and define as usual $(\alpha f)(n) = \alpha(f(n))$ and $(f+g)(n) = f(n) + g(n)$. Clearly $\alpha f \in \ell^2$ if $\alpha \in \mathbb{C}$ and $f \in \ell^2$. Since $|f(n)g(n)| \leq \frac{1}{2}[|f(n)|^2 + |g(n)|^2]$, we have

$$\begin{aligned} |f(n) + g(n)|^2 &\leq (|f(n)| + |g(n)|)^2 = |f(n)|^2 + 2|f(n)g(n)| + |g(n)|^2 \\ &\leq 2\big[|f(n)|^2 + |g(n)|^2\big], \end{aligned}$$

and so $f + g \in \ell^2$ whenever $f, g \in \ell^2$. Thus ℓ^2 is a vector space. Moreover, the series $\sum_{n=1}^{\infty} \overline{f(n)} g(n)$ converges absolutely if $f, g \in \ell^2$. We define

$$\langle f \mid g \rangle = \sum_{n=1}^{\infty} \overline{f(n)} g(n)$$

for $f, g \in \ell^2$. This defines an inner product in ℓ^2, leading to the norm $\|f\| = \|f\|_2 = \left(\sum_{n=1}^{\infty} |f(n)|^2 \right)^{1/2}$. $\triangleleft$

In the next theorem, the equation in (a) is the inner product space version of the *Pythagorean theorem*. The equation in (b) is called the *parallelogram law* and the one in (c) is the *polarisation identity*. All are proved by straightforward calculations.

Theorem 2.3 *Let E be an inner product space.*

(a) *If $\varphi_1, \ldots, \varphi_n \in E$ are vectors satisfying $\langle \varphi_i \mid \varphi_j \rangle = 0$, whenever $i \neq j$, then*

$$\left\| \sum_{k=1}^{n} \varphi_k \right\|^2 = \sum_{k=1}^{n} \|\varphi_k\|^2.$$

(b) *For all $\varphi, \psi \in E$,*

$$\|\varphi + \psi\|^2 + \|\varphi - \psi\|^2 = 2\|\varphi\|^2 + 2\|\psi\|^2.$$

(c) *If F is any vector space and $B : F \times F \to \mathbb{C}$ a sesquilinear form, then for any $\varphi, \psi \in F$*

$$B(\varphi, \psi) = \frac{1}{4} \sum_{n=0}^{3} i^n B(\psi + i^n \varphi, \psi + i^n \varphi).$$

The polarisation identity, in particular, shows that the norm of an inner product space completely determines the inner product defining it.

Definition 2.2 If the inner product space E is complete with respect to the norm defined by its inner product, i.e. if every Cauchy sequence converges, then E is called a *Hilbert space*.

Example 2.3 The inner product spaces $\mathbb{C}^n$ and ℓ^2 treated in Examples 2.1 and 2.2 are Hilbert spaces. We omit the standard completeness proofs. $\triangleleft$

Unless otherwise stated, throughout the rest of this work we assume that $\mathcal{H}$ is a Hilbert space whose inner product is the mapping $(\varphi, \psi) \mapsto \langle \varphi \,|\, \psi \rangle$. The parallelogram law has a central role in the study of Hilbert space geometry. For example, let $E \neq \emptyset$ be a closed subset of $\mathcal{H}$ and assume that E is also convex (i.e. $t\varphi + (1-t)\psi \in E$ whenever $\varphi, \psi \in E$ and $t \in [0, 1]$). Let d be the infimum of the set $\{\|\varphi\| \mid \varphi \in E\}$. Then there is a sequence (φ_n) in E with $\lim_{n\to\infty} \|\varphi_n\| = d$. The parallelogram law shows that

$$\begin{aligned}\|\varphi_m - \varphi_n\|^2 &= 2\|\varphi_m\|^2 + 2\|\varphi_n\|^2 - 4\left\|\tfrac{1}{2}(\varphi_m + \varphi_n)\right\|^2 \\ &\leq 2\|\varphi_m\|^2 + 2\|\varphi_n\|^2 - 4d^2 \to 0,\end{aligned}$$

when $m, n \to \infty$. Thus (φ_n) is a Cauchy sequence and hence converges to some $\varphi \in E$ (as E is closed), and $\|\varphi\| = \lim_{n\to\infty} \|\varphi_n\| = d$ by the continuity of the norm. Thus φ is an element of E having the smallest possible norm. The parallelogram law can also be applied analogously to the above proof to show that such a φ is uniquely determined.

Suppose now that M is a closed vector subspace of $\mathcal{H}$, $\varphi \in \mathcal{H}$, and $E = \varphi - M$ $(= \{\varphi - \psi \mid \psi \in M\})$. Then, as shown above, in E there is an element $\xi = \varphi - \psi$ having the smallest possible norm. If $\eta \in M$, $\|\eta\| = 1$, the inner product of $\langle \eta \,|\, \xi \rangle \eta$ and $\xi - \langle \eta \,|\, \xi \rangle \eta$ vanishes, and so the Pythagorean theorem shows that

$$|\langle \eta \,|\, \xi \rangle|^2 + \|\xi - \langle \eta \,|\, \xi \rangle \eta\|^2 = \|\xi\|^2.$$

But $\xi - \langle \eta \,|\, \xi \rangle \eta \in E$, so that $\|\xi - \langle \eta \,|\, \xi \rangle \eta\|^2 \geq \|\xi\|^2$, implying $\langle \eta \,|\, \xi \rangle = 0$. Thus ξ belongs to the *orthogonal complement* $M^\perp = \{\theta \in \mathcal{H} \mid \langle \theta \,|\, \eta \rangle = 0 \text{ for all } \eta \in M\}$ of M and $\varphi = \psi + \xi$ where $\psi \in M$, $\xi \in M^\perp$. Since $M \cap M^\perp = \{0\}$ and

$$M^\perp = \cap_{\eta \in M} \{\theta \in \mathcal{H} \mid \langle \theta \,|\, \eta \rangle = 0\}$$

is a closed subspace of $\mathcal{H}$, we have proved the following *projection theorem*:

Theorem 2.4 *If M is a closed subspace of $\mathcal{H}$, then $\mathcal{H}$ is the direct sum of M and the closed subspace $M^\perp$, that is, $\mathcal{H} = M \oplus M^\perp$.*

The statement $\mathcal{H} = M \oplus M^\perp$ means that every $\varphi \in \mathcal{H}$ can be uniquely expressed as $\varphi = \psi + \xi$ with $\psi \in M$ and $\xi \in M^\perp$. Denoting $P_M\varphi = \psi$, we thus obtain a mapping $P_M : \mathcal{H} \to M$, which we call the *(orthogonal) projection* of $\mathcal{H}$ onto M. The definition immediately shows that P_M is linear. Since $\|\varphi\|^2 = \|\psi\|^2 + \|\xi\|^2$ by the Pythagorean theorem, we have $\|P_M\varphi\| = \|\psi\| \le \|\varphi\|$.

2.2 The Fréchet–Riesz Theorem and Bounded Linear Operators

Let $\mathcal{H}$ be a Hilbert space. A linear map $T : \mathcal{H} \to \mathcal{H}$ is called a *bounded (linear) operator* or a *bounded linear map* if there is a constant $C \in [0, \infty)$ satisfying $\|T\varphi\| \le C\,\|\varphi\|$ for all $\varphi \in \mathcal{H}$. The existence of such a constant C is equivalent to the norm continuity of T. We let $\mathcal{L}(\mathcal{H})$ denote the set of all bounded linear maps $T : \mathcal{H} \to \mathcal{H}$. If $T \in \mathcal{L}(\mathcal{H})$ we write $\|T\| = \sup\{\|T\varphi\| \mid \varphi \in \mathcal{H},\ \|\varphi\| \le 1\}$. It is easy to see that $\mathcal{L}(\mathcal{H})$ is a vector space, and $\|\cdot\|$ is a norm. Moreover, $\|T\varphi\| \le \|T\|\,\|\varphi\|$, which implies that $\|ST\| \le \|S\|\,\|T\|$ for all $S,\ T \in \mathcal{L}(\mathcal{H})$.

In general, we denote by G^* the *dual* of a normed space G (over $\mathbb{C}$), i.e. G^* is the space of continuous linear functionals $f : G \to \mathbb{C}$. Here continuity is equivalent to the condition

$$\|f\| = \sup\{|f(x)| \mid \|x\| \le 1\} < \infty,$$

and the function $\|\cdot\|$ is a norm on G^*. The following key result is called the *Fréchet–Riesz representation theorem*.

Theorem 2.5 *For each $f \in \mathcal{H}^*$ there is a unique $\psi \in \mathcal{H}$ satisfying $f(\varphi) = \langle\, \psi \mid \varphi\,\rangle$ for all $\varphi \in \mathcal{H}$. Moreover, $\|\psi\| = \|f\|$.*

Proof Let $f \in \mathcal{H}^*$. We may assume that $f \ne 0$, so that $M = \{\varphi \mid f(\varphi) = 0\}$ is a proper closed subspace of $\mathcal{H}$. It follows from Theorem 2.4 that there is a $\xi \in M^\perp$ such that $\|\xi\| = 1$. If $\varphi \in \mathcal{H}$, then

$$\varphi - \frac{f(\varphi)}{f(\xi)}\xi \in M,$$

since $f\big(\varphi - \big(f(\varphi)/f(\xi)\big)\,\xi\big) = f(\varphi) - \big(f(\varphi)/f(\xi)\big)\,f(\xi) = 0$. This means that $\big\langle\, \xi \mid \varphi - \big(f(\varphi)/f(\xi)\big)\,\xi\,\big\rangle = 0$, implying $\langle\, \xi \mid \varphi\,\rangle = f(\varphi)/f(\xi)\ \langle\, \xi \mid \xi\,\rangle = f(\varphi)/f(\xi)$. Therefore we may choose $\psi = \overline{f(\xi)}\xi$. The uniqueness part is clear, since if $\big\langle\, \psi \mid \psi - \psi'\,\big\rangle = \big\langle\, \psi' \mid \psi - \psi'\,\big\rangle$, then $\big\|\psi - \psi'\big\|^2 = 0$. As $|f(\varphi)| \le \|\psi\|\,\|\varphi\|$, we have $\|f\| \le \|\psi\|$, and on the other hand $\|\psi\|^2 = f(\psi) \le \|f\|\,\|\psi\|$, so that $\|\psi\| \le \|f\|$. □

A straightforward consequence is that the mapping $\psi \mapsto f_\psi$ where $f_\psi(\varphi) = \langle\, \psi \mid \varphi\,\rangle$ for all $\varphi \in \mathcal{H}$ is a conjugate-linear isometric bijection from $\mathcal{H}$ onto $\mathcal{H}^*$. Another consequence is the following result.

Proposition 2.1 *Let* $B : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ *be a* bounded sesquilinear form, *i.e. a mapping satisfying the following conditions:*

(i) $B(\alpha\varphi + \beta\psi, \xi) = \overline{\alpha}B(\varphi, \xi) + \overline{\beta}B(\psi, \xi)$ *and*
(ii) $B(\varphi, \alpha\psi + \beta\xi) = \alpha B(\varphi, \psi) + \beta B(\varphi, \xi)$ *for all* $\alpha,\ \beta \in \mathbb{C},\ \varphi,\ \psi,\ \xi \in \mathcal{H}$;
(iii) $\sup\big\{|B(\varphi, \psi)| \,\big|\, \|\varphi\| \le 1,\ \|\psi\| \le 1\big\} < \infty.$

Then there is a unique $S \in \mathcal{L}(\mathcal{H})$ *such that* $B(\varphi, \psi) = \langle\, S\varphi \,|\, \psi \,\rangle$ *for all* $\varphi,\ \psi \in \mathcal{H}$. *Moreover,* $\|S\| = \sup\big\{|B(\varphi, \psi)| \,\big|\, \|\varphi\| \le 1,\ \|\psi\| \le 1\big\}$.

Proof Let C denote the supremum in (iii). If $\varphi \in \mathcal{H}$, we get a linear functional f_φ on $\mathcal{H}$ by setting $f_\varphi(\psi) = B(\varphi, \psi)$, and since $|f_\varphi(\psi)| \le C\,\|\varphi\|\,\|\psi\|$, f_φ is continuous. Theorem 2.5 yields a unique $\xi_\varphi \in \mathcal{H}$ such that $f_\varphi(\psi) = \big\langle\, \xi_\varphi \,|\, \psi \,\big\rangle$ for all $\psi \in \mathcal{H}$. We define $S\varphi = \xi_\varphi$. Since $B(\alpha\varphi_1 + \beta\varphi_2, \psi) = \overline{\alpha}B(\varphi_1, \psi) + \overline{\beta}(\varphi_2, \psi) = \overline{\alpha}\,\langle\, S\varphi_1 \,|\, \psi \,\rangle + \overline{\beta}\,\langle\, S\varphi_2 \,|\, \psi \,\rangle = \langle\, \alpha S\varphi_1 + \beta S\varphi_2 \,|\, \psi \,\rangle$, S is linear. Since

$$\|S\varphi\|^2 = \langle\, S\varphi \,|\, S\varphi \,\rangle = B(\varphi, S\varphi) \le C\,\|\varphi\|\,\|S\varphi\|$$

we have $\|S\varphi\| \le C\,\|\varphi\|$, and so S is bounded. The uniqueness of S follows from that of ξ_φ. The proof of the norm equality is an easy exercise. □

The above result can be used to define for each $T \in \mathcal{L}(\mathcal{H})$ its *adjoint* as the map $T^* \in \mathcal{L}(\mathcal{H})$ which is characterised by the equation $\langle\, \varphi \,|\, T\psi \,\rangle = \langle\, T^*\varphi \,|\, \psi \,\rangle$ for all $\varphi,\ \psi \in \mathcal{L}(\mathcal{H})$: we simply take $B(\varphi, \psi) = \langle\, \varphi \,|\, T\psi \,\rangle$ in Proposition 2.1. Since $\|T^*\varphi\|^2 \le \langle\, \varphi \,|\, TT^*\varphi \,\rangle \le \|\varphi\|\,\|T\|\,\|T^*\varphi\|$, it is clear that $\|T^*\| \le \|T\|$. Using (a) in the next theorem, we see that on the other hand $\|T\| = \|T^{**}\| \le \|T^*\|$, and so $\|T^*\| = \|T\|$.

Theorem 2.6 *If* $S,\ T \in \mathcal{L}(\mathcal{H})$ *and* $\alpha \in \mathbb{C}$, *then*

(a) $T^{**} = T$;
(b) $(S + T)^* = S^* + T^*$;
(c) $(\alpha T)^* = \overline{\alpha}T^*$;
(d) $(ST)^* = T^*S^*$;
(e) $\|T^*T\| = \|T\|^2$.

We omit the simple proof. We still mention some notions defined in terms of the adjoint of $T \in \mathcal{L}(\mathcal{H})$. If $T^* = T$, T is *selfadjoint*. If $T^*T = TT^*$, T is *normal*. If $T^*T = TT^* = I$, where I (or $I_{\mathcal{H}}$) is the identity map of $\mathcal{H}$, T is *unitary*. If $\|T\varphi\| = \|\varphi\|$ for all $\varphi \in \mathcal{H}$, T is *isometric*. Using the polarisation identity it is easy to see that T is unitary if and only if it is an isometric surjection.

The norm of a selfadjoint operator has the following property.

Proposition 2.2 *If* $T \in \mathcal{L}(\mathcal{H})$ *is selfadjoint, then*

$$\|T\| = \sup_{\|\varphi\| \le 1} |\,\langle\, \varphi \,|\, T\varphi \,\rangle\,|.$$